remote sensing

MDPI

Article

Assessment of Invasive and Weed Species by Hyperspectral Imagery in Agrocenoses Ecosystem

Pavel A. Dmitriev [1,*], Boris L. Kozlovsky [1], Denis P. Kupriushkin [1], Anastasia A. Dmitrieva [1], Vishnu D. Rajput [1], Vasily A. Chokheli [1], Ekaterina P. Tarik [1], Olga A. Kapralova [1], Valeriy K. Tokhtar [2], Tatiana M. Minkina [1] and Tatiana V. Varduni [1]

[1] Botanical Garden, Academy of Biology and Biotechnologies, Southern Federal University, 344006 Rostov-on-Don, Russia; blk@sfedu.ru (B.L.K.); dkupryushkin@sfedu.ru (D.P.K.); admit@sfedu.ru (A.A.D.); rvishnu@sfedu.ru (V.D.R.); vachokheli@sfedu.ru (V.A.C.); tarik@sfedu.ru (E.P.T.); oakapralova@sfedu.ru (O.A.K.); minkina@sfedu.ru (T.M.M.); varduny@sfedu.ru (T.V.V.)

[2] Botanical Garden, Belgorod State National Research University, 308015 Belgorod, Russia; tokhtar@bsu.edu.ru

* Correspondence: pdmitriev@sfedu.ru

Abstract: The present study aimed to investigate the possibility of using hyperspectral imaging data to identify the invasive and weed species in agrocenoses ecosystem. The most common weeds in grain agrocenoses, i.e., *Ambrosia artemisiifolia* L., *Euphorbia seguieriana* Neck., *Atriplex tatarica* L., *Glycyrrhiza glabra* L., *Setaria pumila* (Poir.) Roem. and Schult, served as objects. The population of weeds, especially *Ambrosia artemisiifolia* is invasive for the selected region of study. Therefore, the shooting of objects was carried out with a hyperspectral camera, Cubert UHD185, and the values of 100 spectral channels were obtained from hyperspectral images. The values of 80 vegetation indices (VIs) were calculated. The material was processed using mathematical statistics (analysis of variance, *t*-test) and search methods of data analysis (principal component analysis, decision tree, and random forest). Using statistical methods, the simultaneous use of several VIs differentiated between species more deliberately and precisely. The combination of VIs Derivative index (D1), Chlorophyll content index (Datt3), and Pigment specific normalized difference (PSND) can be used for weeds identification. Using the decision tree method, VIs established a good division of weeds into groups; (1) perennial rhizomatous weeds (*Euphorbia seguieriana*, and *Glycyrrhiza glabra*), and (2) annual weeds (*A. artemisiifolia*, *A. tatarica*, and *S. pumila*); These Vis are Chlorophyll index (CI), D1, and Datt3. Using the random forest method, the VIs that have the greatest impact on Mean Decrease Accuracy and Mean Decrease Gini are D1, Datt3, PSND, and Double Peak Index (DPI). The use of spectral channel values for the identification of plant species using the principal component analysis, decision tree, and random forest methods showed worse results than when using VIs. A great similarity of the results was obtained with the help of statistical and search methods of data analysis.

Keywords: invasive species; weeds; hyperspectral imaging; vegetation indices; species identification; agrocenosis

Citation: Dmitriev, P.A.; Kozlovsky, B.L.; Kupriushkin, D.P.; Dmitrieva, A.A.; Rajput, V.D.; Chokheli, V.A.; Tarik, E.P.; Kapralova, O.A.; Tokhtar, V.K.; Minkina, T.M.; et al. Assessment of Invasive and Weed Species by Hyperspectral Imagery in Agrocenoses Ecosystem. *Remote Sens.* **2022**, *14*, 2442. https://doi.org/10.3390/rs14102442

Academic Editors: Barbara Tokarska-Guzik and Sylwia Szporak-Wasilewska

Received: 8 April 2022
Accepted: 17 May 2022
Published: 19 May 2022

1. Introduction

The effect of alien organisms on the flora, fauna, and society is gaining global significance, since the problems associated with their distribution in the world currently can be addressed only at the international level [1–5]. The global agricultural sector is facing increasing challenges posed by a range of stressors, including a rapidly growing population, the depletion of natural resources, environmental pollution, crop diseases, and climate change [6]. Under the conditions of agrarially developed regions, invasive species become an integral component of the weed flora of agrophytocenoses. An important task of preventing invasions within vast plowed areas is the identification of invasive and weedy plants.

Weed vegetation is an inherent natural component of agrocenoses which grows and spreads rapidly within fields. It becomes a rival to grain crops when it requires physical space, nutrients, sunlight, and water [7]. This rivalry has a great impact on harvest. Weed spread is usually controlled mechanically or chemically with herbicides. A large proportion of herbicides end up on crops or soil, and only a small percentage of herbicides reach the weeds themselves [8]. Apart from being potentially harmful to the environment and human health, herbicides and their usage cost a lot of money in terms of the production of agricultural products. That is why the latest technologies aimed to control and fight with weeds are very important in ecological and economical spheres. Due to that fact, the understanding of weed spreading and their timely identification within a field for using plant protection products is crucial. Earth remote sensing provides great prospects for weed vegetation foci detection [9,10].

Earth remote sensing is one of the most important tools in the study of agricultural lands. Its advantage is the possibility to obtain profound information about the qualitative and quantitative state of the agricultural land quickly and accurately [11,12]. Earth remote sensing is a valuable tool in predicting the occurrence of possible stress conditions and monitoring the state of the agricultural crops, soil, identifying plant diseases, identifying water shortages, weeds, damage to plants caused by insects, hail, and wind [13,14].

The works which connect the hyperspectral method study of plants and agrocenosis ecosystems usually analyze the main following topics: (a) crop management and productivity, (b) disease detection, (c) weed and alien species detection, (d) crop (seed) quality, (e) and environment management [15].

Crop management and productivity is readily reduced by competition from weeds. It is particularly important to control weeds early to prevent yield losses. Limited herbicide choices and increasing costs of weed management are threatening the profitability of crops. The most important thing for an automatic system to remove weeds within crop rows is to utilize reliable sensing technology to achieve accurate differentiation of weeds and crops at specific locations in a field. Currently, weeds within crop rows still rely on manual removal in many cases, but manual weeding is a less efficient method. Its cost can be more than five times the cost of the former [16].

The literature related to the research direction present data on the use of sensing methods including spectroscopy, color imaging, and hyperspectral imaging in the discrimination of crops and weeds [17]. Image segmentation between crop and weed is useful for selective weeding using a pixel discriminant model generated from hyperspectral images [17,18].

Hyperspectral imaging combines the main features of imaging and spectroscopy to collect spectral information over the full wavelength range for each pixel of the acquired image [19]. Smart agriculture can use intelligent technology to accurately measure the distribution of weeds in the field and perform weed control tasks in selected areas, which can not only improve the effectiveness of pesticides, but also increase the economic benefits of agricultural products [17,20]. Related food safety risks have aroused growing interest in organic foods. The total production of organic foods has increased significantly in recent years, especially in Europe [21].

Hyperspectral imaging is expected to become a powerful technology for the detection and control of weeds and alien species. To perform this, it is necessary to differentiate weeds from agricultural plants. For this, different approaches are used [22]. Noble and Crowe (2001) classified six plant species using the UVVis-NIR (250–2500 nm) spectrophotometer [23]. Terawaki et al. (2002) discriminated between sugar beet and weeds using RGB color and leaf shapes [24]. Some studies proposed a portable hyperspectral imaging system using; this technique is expected to be applied to future automatic mechanical weeding systems [25]. Results of the research were carried out depend on image processing for different species weed detection [26]. There are a lot of very perspective and reliable results which testify the new possibilities in identification of many alien plant species [3,27–29].

The first step required to distinguish spectrally between invasive species, weeds, and crops is to obtain a spectrum of individual plants for each species or group of species. This

can be performed with high spatial and spectral resolutions [12,30]. To combat invasive and weedy species, it is important to have a method for identifying plants for their simultaneous recognition and destruction using unmanned aerial vehicles (UAVs). In this regard, the purpose of the study was to identify alien species and weeds according to the data of a hyperspectral survey of agricultural fields after the harvest of winter wheat. Thus, the research objectives were: (1) conduct hyperspectral imaging of such species as *Ambrosia artemisiifolia* L., *Atriplex tatarica* L., (invasive species), *Euphorbia seguieriana* Neck., *Glycyrrhiza glabra* L., and *Setaria pumila* (Poir.) Roem. and Schult (weeds), (2) calculate the values of vegetation indices, (3) evaluate the efficiency of using spectral indices and spectral channels to identify different types of weeds, (4) and use analysis of variance (ANOVA), *t*-test, random forest (RF), decision tree (DT), and principal component analysis (PCA) methods to select the most informative VIs for identifying weed species.

2. Materials and Methods

2.1. Research Region

The study of samples of weed species was carried out in the Rostov region 7°16′25.63″N; 39°19′13.59″E on the territory of the educational and experimental farm "Nedvigovka" of the Southern Federal University (SFedU), Rostov-on-Don, Russia (Figure 1). The climate of this region is moderately continental and dry. The average annual precipitation is 548 mm, most of it falls during the frost-free period. Summer is hot, the average temperature in July is + 22 to + 23 °C, the maximum temperature is +40 °C. Winter is moderately mild, the average temperature in January is −5 °C, the average absolute minimum air temperature is −20 to −25 °C, the absolute minimum temperature is −32 °C. The growing season lasts 216 days (from 1 April to 4 November), the frost-free period lasts 258 days [31].

Figure 1. Location of the study site.

2.2. Research Methods

Five of the most common weed species found in agrocenoses of grain crops were selected as objects of research: *A. artemisiifolia* L., *E. seguieriana* Neck., *A. tatarica* L., *G. glabra* L., and *S. pumila* (Poir.) Roem. and Schult. (Figure 2).

The spectral characteristics of the selected objects were studied using a Cubert UHD185 frame hyperspectral camera [32,33] (Figure 3 and Table 1).

Weeds were removed in 2020 a month after harvesting winter wheat. The hyperspectral survey was carried out in the daytime in sunny and cloudless weather. The objects were mostly illuminated by the sun and were selected for shooting, for which the camera was located from the southeast side of the object at a distance of 90 cm. The reflected

electromagnetic radiation from the range in the range of 450–950 nm was recorded. A white reference panel was used to calibrate the reflectance. Each species is represented by 20 samples randomly selected from the group and each sample is represented by 5 images.

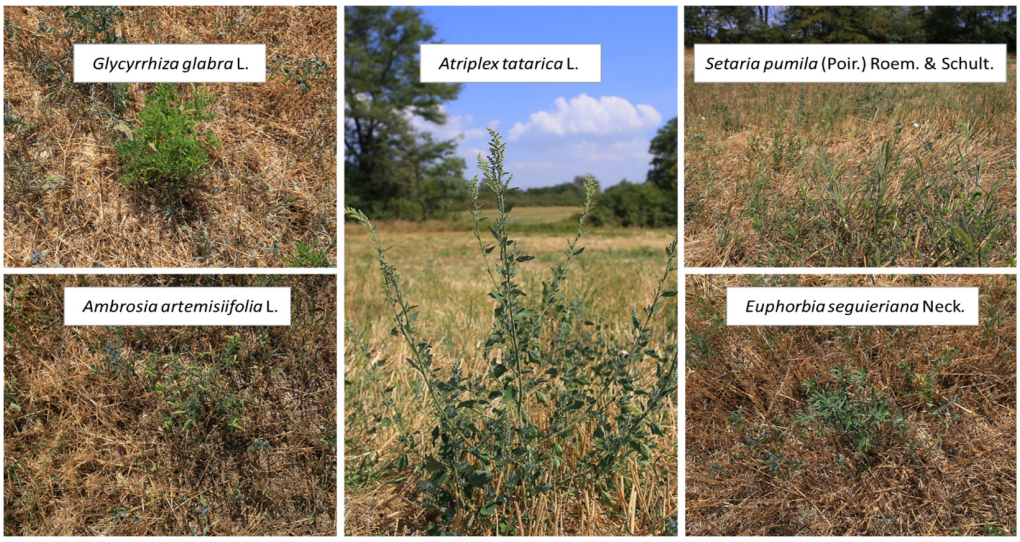

Figure 2. Photo of the studied plant species in the field. Educational and experimental farm "Nedvigovka" of the Southern Federal University, Rostov-on-Don, Russia.

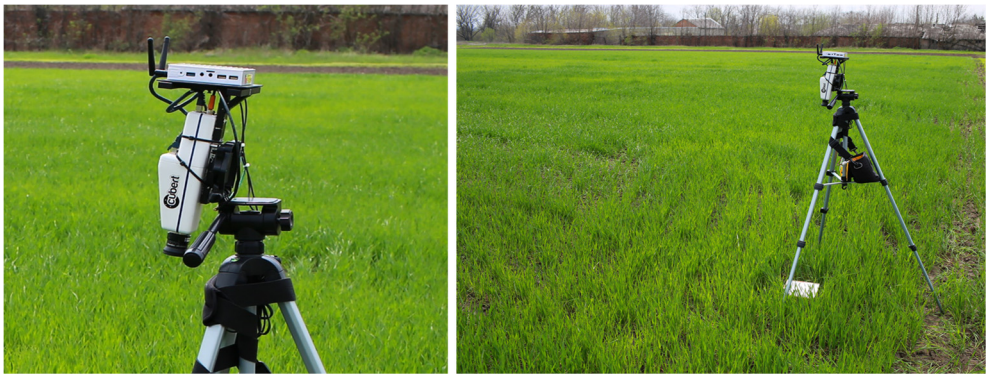

Figure 3. Collecting spectral characteristics of weed plant species (**right**) using a Cubert UHD 185 hyperspectral camera (**left**).

Table 1. Characteristics of the Cubert UHD 185 hyperspectral camera.

Technology	MPS
The number of spectral channels	125
Spectral productivity	2500 spectra
Spectral range	450–950 nm
Spectral resolution	4 nm
Type of the camera	Frame
Sensor type	Si CCD
Signal width	12 bit
Signal to noise ratio	58 dB
Dynamic range	68 dB
Focal length	10 mm

Data are recorded in the form of 1000 px × 1000 px panchromatic images and 50 px × 50 px images for each hyperspectral channel. The spectral profile of the sample was calculated using the Cube-Pilot program, developed specifically for the hyperspectral camera Cubert UHD-185.

To assess the possibilities of using VIs, calculated according to hyperspectral survey data for remote detection of certain invasive and quarantine plant species, a list of 80 VIs was compiled from literature data (Table 2).

Table 2. VIs tested for their ability to distinguish target weeds of this study.

Index Name	Formula for Calculating	References
Boochs	D_{703}	[34]
Boochs2	D_{720}	[34]
CARI	$R_{700} \times abs(a \times 670 + R_{670} + b)/R_{670} \times (a2 + 1) \times 0.5$ a = $(R_{700} \times R_{550})/150$, b = $R_{550} - (a \times 550)$	[35]
Carter2	R_{695}/R_{760}	[36]
Carter3	R_{605}/R_{760}	[36]
Carter4	R_{710}/R_{760}	[36]
Carter5	R_{695}/R_{670}	[36]
Carter6	R_{550}	[36]
CI	$R_{675} \times R_{690}/R^2_{683}$	[37]
CI2	$R_{760}/R_{700} - 1$	[38]
ClAInt	$\int_{600nm}^{735nm} R$	[39]
CRI1	$1/R_{515} - 1/R_{550}$	[38]
CRI2	$1/R_{515} - 1/R_{770}$	[38]
CRI3	$1/R_{515} - 1/R_{550} \times R_{770}$	[38]
CRI4	$1/R_{515} - 1/R_{700} \times R_{770}$	[38]
D1	D_{730}/D_{706}	[37]
D2	D_{705}/D_{722}	[37]
Datt	$(R_{850} - R_{710})/(R_{850} - R_{680})$	[40]
Datt2	R_{850}/R_{710}	[40]
Datt3	D_{754}/D_{704}	[40]
Datt4	$R_{672}/(R_{550} \times R_{708})$	[41]
Datt5	R_{672}/R_{550}	[41]
Datt6	$R_{860}/R_{550} \times R_{708}$	[41]
DD	$(R_{749} - R_{720}) - (R_{701} - R_{672})$	[42]
DDn	$2 \times (R_{710} - R_{660} - R_{760})$	[42]
DPI	$D_{688} \times D_{710}/D^2_{697}$	[37]
DWSI4	R_{550}/R_{680}	[43]
EGFN	$(max(D_{650:750}) + max(D_{500:550}))/(max(D_{650:750}) + max(D_{500:550}))$	[44]
EGFR	$max(D_{650:750})/max(D_{500:550})$	[44]
EVI	$2.5 \times ((R_{800} - R_{670})/(R_{800} - 6 \times R_{670} - 7.5 \times R_{475} + 1))$	[45]
GI	R_{554}/R_{677}	[46]
Gitelson	$1/R_{700}$	[47]
Gitelson2	$(R_{750} - R_{800}/R_{695} - R_{740}) - 1$	[38]
GMI1	R_{750}/R_{550}	[38]
GMI2	R_{750}/R_{700}	[38]
Green NDVI	$(R_{800} - R_{550})/(R_{800} + R_{550})$	[48]
Maccioni	$(R_{780} - R_{710})/(R_{780} - R_{680})$	[49]
MCARI	$((R_{700} - R_{670}) - 0.2 \times (R_{700} - R_{550})) \times (R_{700} - R_{670})$	[50]
MCARI2	$((R_{700} - R_{670}) - 0.2 \times (R_{700} - R_{550})) \times (R_{700}/R_{670})$	[50]
MPRI	$(R_{515} - R_{530})/(R_{515} + R_{530})$	[51]
MSAVI	$0.5 \times (2 \times R_{800} + 1 - ((2 \times R_{800} + 1)2 - 8 \times (R_{800} - R_{670}))0.5)$	[52]
mSR2	$(R_{750}/R_{705}) - 1/(R_{750}/R_{705} + 1) \times 0.5$	[53]
MTCI	$(R_{754} - R_{709})/(R_{709} - R_{681})$	[54]
MTVI	$1.2 \times (1.2 \times (R_{800} - R_{550}) - 2.5 \times (R_{670} - R_{550}))$	[55]
NDVI	$(R_{800} - R_{680})/(R_{800} + R_{680})$	[56]
NDVI2	$(R_{750} - R_{705})/(R_{750} + R_{705})$	[57]
NDVI3	$(R_{682} - R_{553})/(R_{682} + R_{553})$	[58]
OSAVI	$(1 + 0.16) \times (R_{800} - R_{670})/(R_{800} + R_{670} + 0.16)$	[59]
OSAVI2	$(1 + 0.16) \times (R_{750} - R_{705})/(R_{750} + R_{705} + 0.16)$	[60]
PARS	R_{746}/R_{513}	[61]
PRI	$(R_{531} - R_{570})/(R_{531} + R_{570})$	[62]

Table 2. *Cont.*

Index Name	Formula for Calculating	References
PRI_norm	$PRI \times (-1)/(RDVI \times R_{700}/R_{670})$	[63]
PRI*CI2	$PRI \times CI2$	[64]
PSRI	$(R_{678} - R_{500})/R_{750}$	[65]
PSSR	R_{800}/R_{635}	[66]
PSND	$(R_{800} - R_{470})/(R_{800} - R_{470})$	[66]
RDVI	$(R_{800} - R_{670})/(R_{800} + R_{670})0.5$	[67]
REP_Li	$700 + 40 \times ((Rre - R_{700})/(R_{740} - R_{700})Rre = (R_{670} - R_{780})/2$	[68]
SAVI	$(1 + L)/(R_{800} - R_{670})/(R_{800} + R_{670} + L)$	[69]
SPVI	$0.4 \times 3.7 \times (R_{800} - R_{670}) - 1.2 \times ((R_{530} - R_{670})2) \times 0.5$	[70]
SR	R_{800}/R_{680}	[71]
SR1	R_{750}/R_{700}	[72]
SR2	R_{752}/R_{690}	[72]
SR3	R_{750}/R_{550}	[72]
SR4	R_{700}/R_{670}	[73]
SR5	R_{675}/R_{700}	[61]
SR6	R_{750}/R_{710}	[74]
SR8	R_{515}/R_{550}	[75]
Sum_Dr1	$\sum_{i=626}^{795} D1_i$	[76]
Sum_Dr2	$\sum_{i=680}^{780} D1_i$	[77]
TCARI	$3 \times ((R_{700} - R_{670}) - 0.2 \times (R_{700} - R_{550}) \times (R_{700}/R_{670}))$	[55]
TCARI/OSAVI	TCARI/OSAVI	[55]
TCARI2	$3 \times ((R_{750} - R_{705}) - 0.2 \times (R_{750} - R_{550}) \times (R_{750}/R_{705}))$	[60]
TCARI2/OSAVI2	TCARI2/OSAVI2	[60]
TGI	$-0.5 \times (190 \times (R_{670} - R_{550}) - 120 \times (R_{670} - R_{480}))$	[78]
TVI	$0.5 \times (120 \times (R_{750} - R_{550}) - 200 \times (R_{670} - R_{550}))$	[79]
Vogelmann	R_{740}/R_{720}	[80]
Vogelmann2	$(R_{734} - R_{747})/(R_{715} + R_{726})$	[80]
Vogelmann3	D_{715}/D_{705}	[80]
Vogelmann4	$(R_{734} - R_{747})/(R_{715} + R_{720})$	[80]

Rxxx: reflectance at the wavelength "xxx"; Dxxx: first derivation of reflectance values at the wavelength "xxx".

For the analysis, the values of the VIs were calculated for each specimen from the studied species, as shown in Table 2. A Savitsky–Golay filter (length 12 nm) was used as a preprocessing step to reduce the measurement error and remove artifacts in the spectral data. The processing of the hyperspectral survey results was carried out in the environment for statistical calculations R (R Core Team), using the hsdar package [81]. The normal VIs distribution was tested using Shapiro–Wilk, Pearson's chi-square, Lilliefors, and Cramer von Mises tests (Supplementary Table S1) to select the method of statistical data analysis methods (parametric and non-parametric). To determine VIs, the values of which to the greatest extent and reliability depend on the species of weeds, a one-way analysis of variance (ANOVA) and an independent two-sample *t*-test were used. Weeds were identified using data analysis methods: principal component analysis (PCA), decision tree (DT), and random forest (RF).

3. Results

3.1. Statistical Methods of Data Analysis

It was found that the values of all VIs of all studied species were distributed according to the value as per normal law (Figure 4).

The one-way analysis of variance showed that the value of 76 VIs significantly depends on the species of plant samples; the intergroup variance significantly exceeds the intragroup variance (Figure 5, Supplementary Table S2). The values of the four VIs (MPRI, Gitelson, Carter6, and ClAInt) do not depend significantly on the "species" factor and are associated with the influence of random factors (Figure 6, Supplementary Table S2).

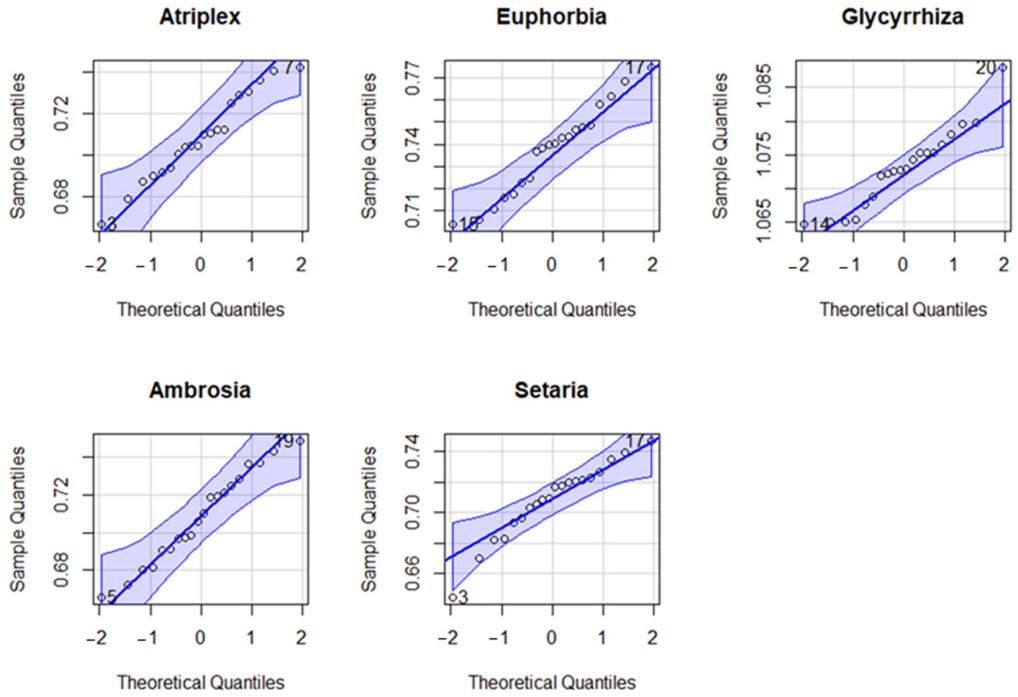

Figure 4. QQ-plot of PSND values.

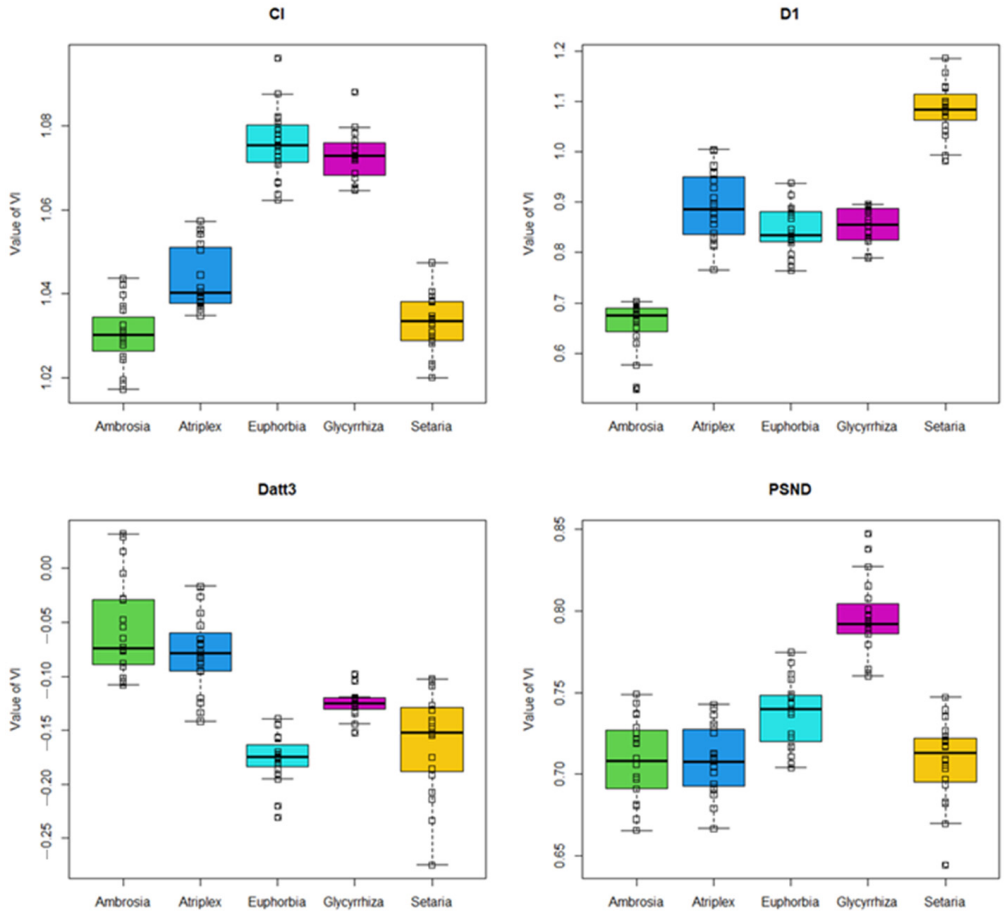

Figure 5. Boxplot of significantly different Vis for *Ambrosia artemisiifolia*, *Atriplex tatarica*, *Glycyrrhiza glabra*, *Euphorbia seguieriana*, and *Setaria pumila*.

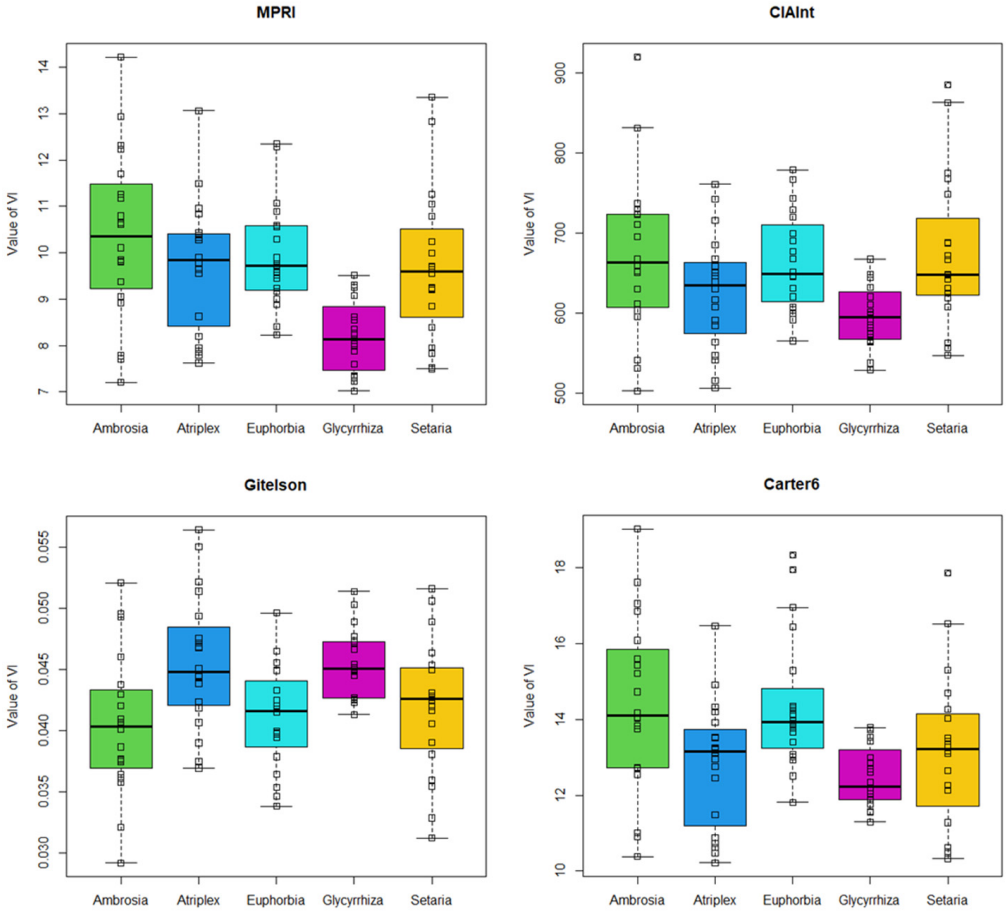

Figure 6. Boxplot of VIs MPRI, Gitelson, Carter6 and ClAInt for *Ambrosia artemisiifolia*, *Atriplex tatarica*, *Glycyrrhiza glabra*, *Euphorbia seguieriana*, and *Setaria pumila*.

The VIs group, the average values of which significantly differed in the compared species according to an independent two-sample *t*-test, is presented in Tables 3, 4 and S3.

Table 3. Significantly different VIs between compared pairs of species according to an independent two-sample *t*-test.

Species	*Ambrosia*	*Atriplex*	*Euphorbia*	*Glycyrrhiza*	*Setaria*
Ambrosia	-	50	69	76	40
Atriplex	50	-	57	65	54
Euphorbia	69	57	-	54	61
Glycyrrhiza	76	65	54	-	68
Setaria	40	54	61	68	-

Table 4. VIs suitable to distinguish species.

Compared Species	Vegetation Index (VI)
Ambrosia vs. *Atriplex*	Boochs2, CARI, Carter2, Carter3, Carter4, Carter5, CI, CI2, CRI3, CRI4, D1, D2, Datt, Datt2, Datt4, Datt6, DD, EGFN, EGFR, Gitelson, GMI1, GMI2, Green_NDVI, Maccioni, MCARI, MCARI2, mSR2, MTCI, NDVI, NDVI2, OSAVI2, PRI.CI2, PSRI, PSSR, REP_Li, SR, SR1, SR2, SR3, SR4, SR6, SR8, TCARI, TCARI.OSAVI, TCARI2, TGI, Vogelmann, Vogelmann2, Vogelmann3, and Vogelmann4

Table 4. *Cont.*

Compared Species	Vegetation Index (VI)
Ambrosia vs. *Euphorbia*	Boochs, Boochs2, Carter2, Carter3, Carter4, CI, CI2, CRI1, CRI3, CRI4, D1, D2, Datt, Datt2, Datt3, Datt4, Datt5, Datt6, DD, DDn, DPI, DWSI4, EGFN, EGFR, EVI, GI, Gitelson2, GMI1, GMI2, Green_NDVI, Maccioni, MCARI, MCARI2, MSAVI, mSR2, MTCI, MTVI, NDVI, NDVI2, NDVI3, OSAVI, OSAVI2, PARS, PRI, PRI.CI2, PRI_norm, PSND, PSRI, PSSR, RDVI, REP_Li, SAVI, SPVI, SR, SR1, SR2, SR3, SR4, SR5, SR6, SR8, Sum_Dr1, Sum_Dr2, TCARI2, TCARI2.OSAVI2, TVI, Vogelmann, Vogelmann2, and Vogelmann4
Ambrosia vs. *Glycyrrhiza*	Boochs, Boochs2, Carter2, Carter3, Carter4, Carter5, Carter6, CI, CI2, ClAInt, CRI1, CRI2, CRI3, CRI4, D1, D2, Datt, Datt2, Datt3, Datt4, Datt5, Datt6, DD, DDn, DPI, DWSI4, EGFN, EGFR, EVI, GI, Gitelson, Gitelson2, GMI1, GMI2, Green_NDVI, Maccioni, MCARI, MCARI2, MPRI, MSAVI, mSR2, MTCI, MTVI, NDVI, NDVI2, NDVI3, OSAVI, OSAVI2, PARS, PRI, PRI.CI2, PRI_norm, PSND, PSRI, PSSR, RDVI, REP_Li, SAVI, SPVI, SR, SR1, SR2, SR3, SR4, SR5, SR6, SR8, Sum_Dr1, Sum_Dr2, TCARI.OSAVI, TCARI2, TVI, Vogelmann, Vogelmann2, Vogelmann3, and Vogelmann4
Ambrosia vs. *Setaria*	Boochs, CARI, Carter5, D1, D2, Datt, Datt2, Datt3, Datt4, Datt5, Datt6, DD, DPI, DWSI4, EGFN, EGFR, GI, Maccioni, MCARI, MTCI, MTVI, NDVI3, PRI, PRI.CI2, PRI_norm, PSRI, RDVI, REP_Li, SPVI, SR4, SR5, Sum_Dr2, TCARI, TCARI.OSAVI, TCARI2.OSAVI2, TGI, TVI, Vogelmann, Vogelmann2, and Vogelmann4
Atriplex vs. *Euphorbia*	Boochs, Boochs2, CARI, Carter3, Carter5, Carter6, CI, CRI1, D1, D2, Datt, Datt3, Datt4, Datt5, DD, DPI, DWSI4, EGFN, EGFR, EVI, GI, Gitelson, Gitelson2, Maccioni, MCARI, MCARI2, MSAVI, MTCI, MTVI, NDVI, NDVI3, OSAVI, PARS, PRI, PRI.CI2, PRI_norm, PSND, PSRI, PSSR, RDVI, REP_Li, SAVI, SPVI, SR, SR2, SR4, SR5, SR8, Sum_Dr1, Sum_Dr2, TCARI, TCARI. OSAVI, TCARI2, TCARI2.OSAVI2, TGI, TVI, and Vogelmann3
Atriplex vs. *Glycyrrhiza*	Boochs, Boochs2, CARI, Carter2, Carter3, Carter4, Carter5, CI, CI2, CRI1, CRI2, CRI3, CRI4, D2, Datt2, Datt3, Datt4, Datt5, DPI, DWSI4, EGFN, EGFR, EVI, GI, Gitelson2, GMI1, GMI2, Green_NDVI, MCARI, MCARI2, MPRI, MSAVI, mSR2, MTVI, NDVI, NDVI2, NDVI3, OSAVI, OSAVI2, PARS, PRI, PRI.CI2, PRI_norm, PSND, PSRI, PSSR, RDVI, SAVI, SPVI, SR, SR1, SR2, SR3, SR4, SR5, SR6, SR8, Sum_Dr1, Sum_Dr2, TCARI, TCARI2, TGI, TVI, Vogelmann, and Vogelmann3
Atriplex vs. *Setaria*	Boochs, Boochs2, Carter2, Carter3, Carter4, Carter5, CI, CI2, CRI3, CRI4, D1, Datt3, Datt4, Datt5, DPI, DWSI4, EGFN, EGFR, EVI, GI, GMI1, GMI2, MCARI, MCARI2, MSAVI, mSR2, MTVI, NDVI, NDVI2, NDVI3, OSAVI, OSAVI2, PRI, PRI.CI2, PRI_norm, PSRI, PSSR, RDVI, SAVI, SPVI, SR, SR1, SR2, SR3, SR4, SR5, SR6, SR8, Sum_Dr1, Sum_Dr2, TCARI, TCARI2, TVI, and Vogelmann3
Euphorbia vs. *Glycyrrhiza*	Boochs2, Carter2, Carter3, Carter4, Carter5, Carter6, CI2, ClAInt, CRI1, CRI2, CRI3, CRI4, D2, Datt, Datt2, Datt3, Datt6, DD, DPI, EVI, Gitelson, GMI1, GMI2, Green_NDVI, Maccioni, MCARI2, MPRI, MSAVI, mSR2, MTCI, NDVI, NDVI2, OSAVI, OSAVI2, PARS, PSND, PSSR, RDVI, REP_Li, SAVI, SR, SR1, SR2, SR3, SR4, SR5, SR6, SR8, TCARI.OSAVI, TCARI2.OSAVI2, Vogelmann, Vogelmann2, Vogelmann3, and Vogelmann4
Euphorbia vs. *Setaria*	Boochs, Boochs2, CARI, Carter2, Carter3, Carter4, Carter5, CI, CI2, CRI1, CRI3, CRI4, D1, Datt, Datt4, Datt5, DDn, DWSI4, EVI, GI, Gitelson2, GMI1, GMI2, Green_NDVI, MCARI, MCARI2, MSAVI, mSR2, MTVI, NDVI, NDVI2, NDVI3, OSAVI, OSAVI2, PARS, PRI, PRI.CI2, PRI_norm, PSND, PSRI, PSSR, RDVI, SAVI, SPVI, SR, SR1, SR2, SR3, SR4, SR5, SR6, SR8, Sum_Dr1, Sum_Dr2, TCARI, TCARI.OSAVI, TCARI2, TCARI2.OSAVI2, TGI, TVI, and Vogelmann
Glycyrrhiza vs. *Setaria*	Boochs, Boochs2, CARI, Carter2, Carter3, Carter4, Carter5, CI, CI2, ClAInt, CRI1, CRI2, CRI3, CRI4, D1, Datt, Datt2, Datt3, Datt4, Datt5, Datt6, DDn, DPI, DWSI4, EVI, GI, Gitelson2, GMI1, GMI2, Green_NDVI, MCARI, MCARI2, MPRI, MSAVI, mSR2, MTVI, NDVI, NDVI2, NDVI3, OSAVI, OSAVI2, PARS, PRI, PRI.CI2, PRI_norm, PSND, PSRI, PSSR, RDVI, SAVI, SPVI, SR, SR1, SR2, SR3, SR4, SR5, SR6, SR8, Sum_Dr1, Sum_Dr2, TCARI, TCARI2, TCARI2.OSAVI2, TGI, TVI, Vogelmann, and Vogelmann3

Nineteen VIs significantly differed between *A. artemisiifolia* and all other species, 20 differed between *A. tatarica* and *E. seguieriana* and all other species, 40 differed between *G. glabra* and all other species, 21 differed between *S. pumila* and all other species, and only 1 vegetation index (VI) significantly differed between species in their values simultaneously in all pairs; which can be observed in SR4 (Figure 7). The critical confidence level for the *Ambrosia* vs. *Atriplex* and *Euphorbia* vs. *Glycyrrhiza* pairs is 0.04 (Supplementary Table S3).

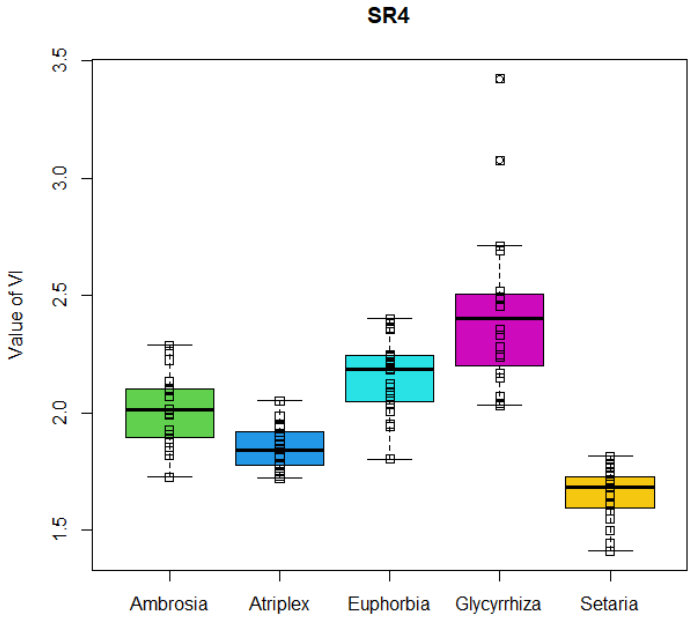

Figure 7. Boxplot of VI SR4.

The result of the analysis showed that the usage of several values of the vegetation VIs for different species at the same time makes it possible to divide them into groups more deliberately than when using only one VI (Figures 6 and 8).

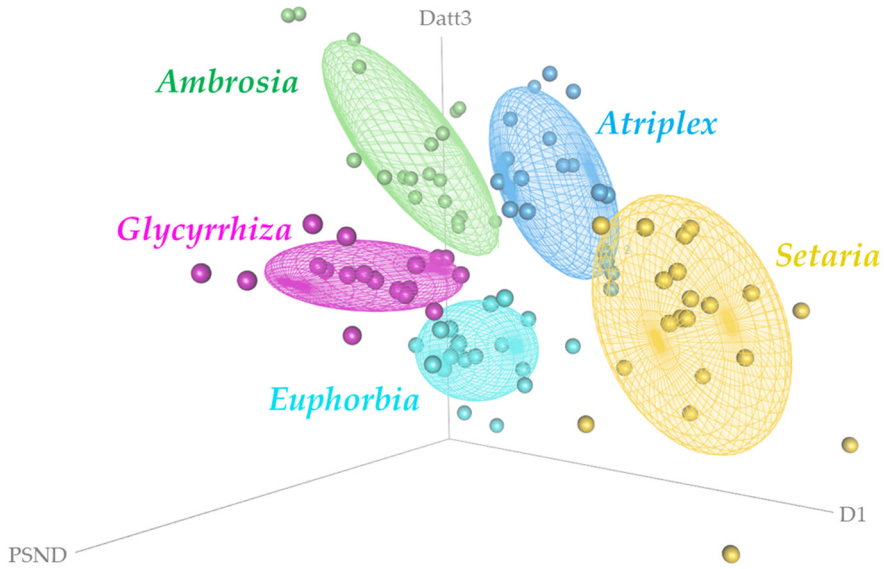

Figure 8. *Ambrosia artemisiifolia*, *Euphorbia seguieriana*, *Atriplex tatarica*, *Glycyrrhiza glabra* и *Setaria pumila* in the range of VIs D1, Datt3, and PSND.

It should be noted that a clear division of weeds by species using combinations of VIs was obtained after harvesting an agricultural crop against a stubble background. This makes it possible to conduct a relative assessment of the degree of weediness in the fields and determine the weed control strategy for the next season.

3.2. Search Methods of Data Analysis

When using probabilistic-statistical methods, it is possible to establish the accuracy of the results and obtained conclusions. However, they are not well-suited for solving problems of identifying objects, especially with a large number of features. Therefore, additional methods of data analysis were used in the study: PCA, DT, and RF. The VIs values and the values of the spectral channels were used as data (the spectral range from 470 to 866 nm was used in the calculations).

3.2.1. Principal Component Analysis (PCA)

The PCA results for the spectral channels are presented in Figure 9.

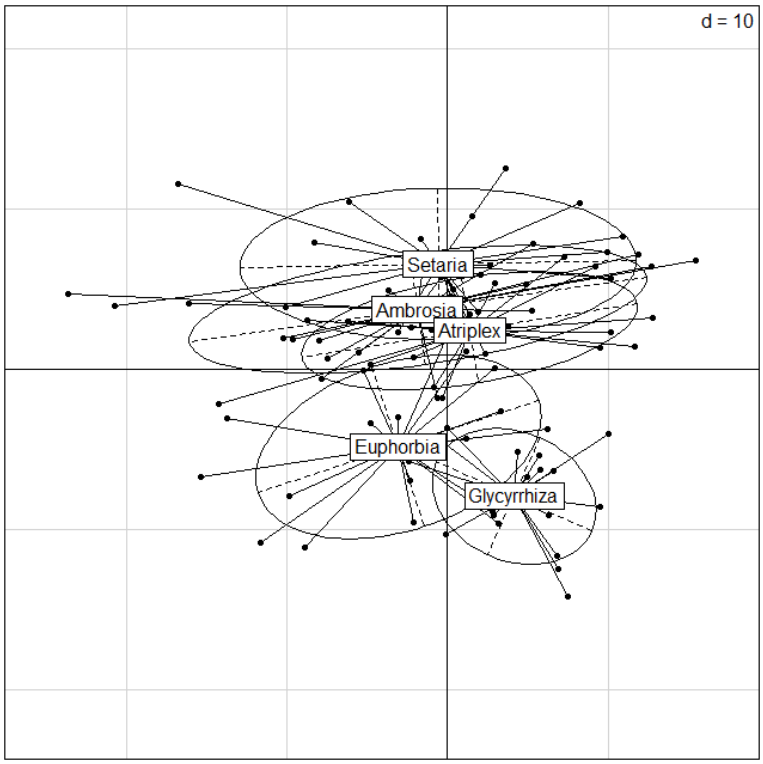

Figure 9. Projection of 100 spectral bands for *Ambrosia artemisiifolia*, *Atriplex tatarica*, *Glycyrrhiza glabra*, *Euphorbia seguieriana*, and *Setaria pumila*.

The projection of spectral channel values shows that *G. glabra* separates from *S. pumila*, *A. artemisiifolia,* and *A. tatarica*. At the same time, *Setaria pumila* is separated from *E. seguieriana* and *G. glabra*. The projection of the values of the spectral channels is extended along the second component, the extreme positions are occupied by *S. pumila* and *G. glabra*. This is due to the fact that at the time of the survey, *G. glabra* was in the stage of active regeneration after mowing, and *S. pumila* was at the stage of an old generative individual.

It was noted that the first two principal components account for 96% of the dispersion of values (Table 5 and Table S4). The factors are limited to four components (in accordance with the Kaiser criterion). However, the factor loads of the spectral channels for all selected components are very low and practically do not differ. It is possible to single out a group of channels from 510 to 594 with conditionally high factor loadings. These channels correspond to the green and yellow parts of the visible region of the electromagnetic spectrum.

PCA results for VIs are similar to those for spectral channels (Figure 10, Tables 6 and S5). Differences are observed in the proportion of dispersions of the first and second components; in the case of VI, the proportion of the dispersion of the first component is higher. VIs have a low factor load. For the first principal component, the factor loading does not exceed 0.142.

Table 5. The value of the dispersion of the main components of the PCA of the spectral channels for *Ambrosia artemisiifolia*, *Atriplex tatarica*, *Glycyrrhiza glabra*, *Euphorbia seguieriana*, and *Setaria pumila*.

Statistics	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5
Standard deviation	7.52233	6.207426	1.448625	1.175973	0.51599
Proportion of variance	0.57157	0.389214	0.021197	0.013969	0.002689
Cumulative proportion	0.57157	0.960784	0.981981	0.99595	0.998639

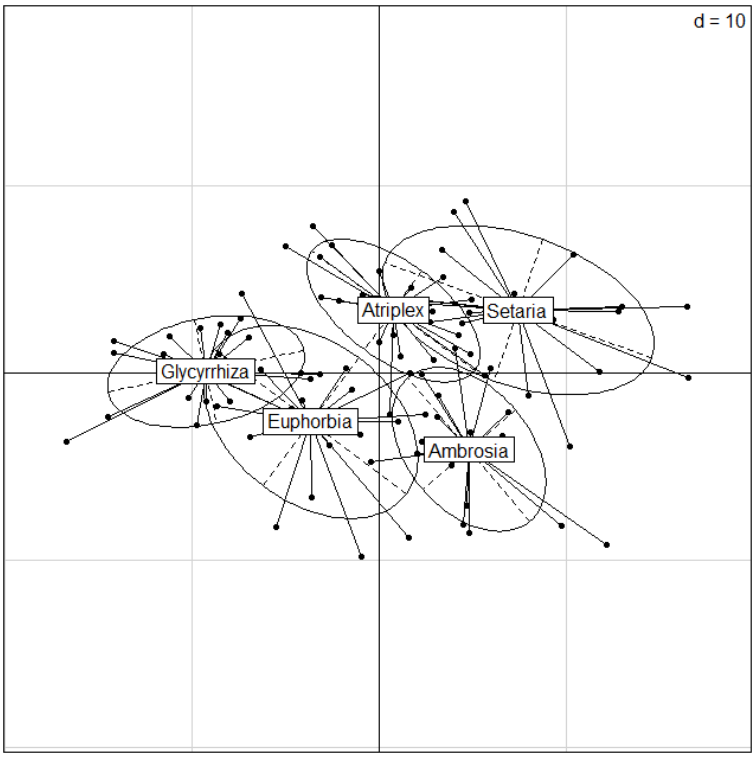

Figure 10. Projection of 80 VI values for *Ambrosia artemisiifolia*, *Atriplex tatarica*, *Glycyrrhiza glabra*, *Euphorbia seguieriana*, and *Setaria pumila*.

Table 6. The value of the dispersion of the main components of the VIs *Ambrosia artemisiifolia*, *Atriplex tatarica*, *Glycyrrhiza glabra*, *Euphorbia seguieriana*, and *Setaria pumila*.

Statistics	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5	Comp.6
Standard deviation	6.96777	4.118203	2.448039	2.019743	1.087375	0.979837
Proportion of Variance	0.613003	0.214136	0.075668	0.051507	0.014929	0.012122
Cumulative Proportion	0.613003	0.827139	0.902807	0.954314	0.969243	0.981365

It is possible to single out VIs that simultaneously have a factor load on the first and second principal components. These are Boochs, Datt2, MCARI, PRI*CI2, SR8, TVI, Vogelmann2, and Vogelmann4. The PCA of these VIs allows us to more clearly divide weeds into two groups: *G. glabra*, *A. artemisiifolia*, and *E. seguieriana;* and *S. pumila* and *A. tatarica* (Figure 11, Tables 7 and 8). In the first group, long-term vegetative species are capable of active regeneration after mowing, among them are two perennial weeds (*G. glabra*, *E. seguieriana*); in the second group are short-vegetating annual weed *S. pumila* and long-term vegetative annual weed *A. tatarica*. The *A. tatarica* has a reflective pubescence of shoots, which probably affects its spectral characteristics.

It was shown that the low and slightly different factor loadings of VIs and spectral channels on the main components do not allow using PCA to solve the main problem; to identify the most effective VIs types for separation.

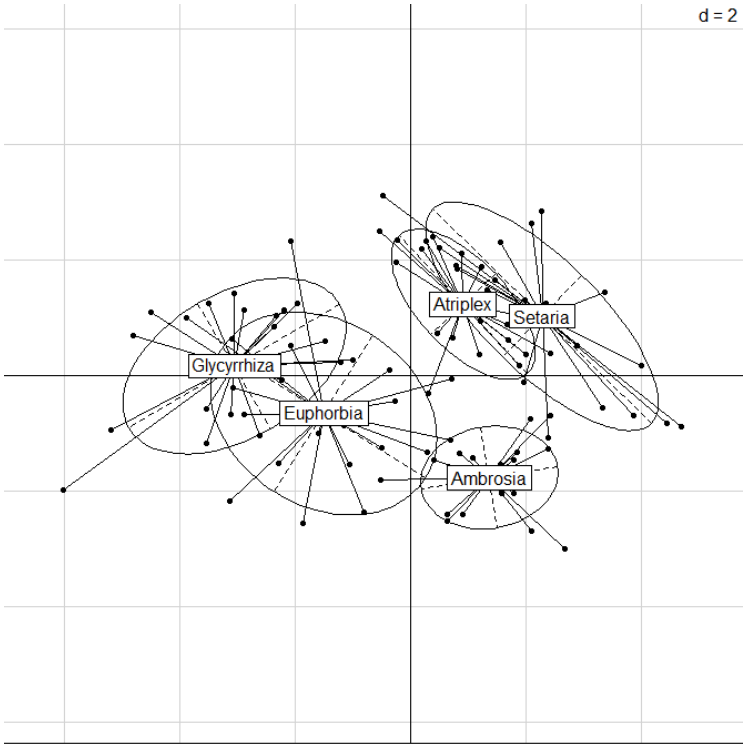

Figure 11. Projection of Boochs, Datt2, MCARI, PRI*CI2, SR8, TVI, Vogelmann2, Vogelmann4 values for *Ambrosia artemisiifolia*, *Atriplex tatarica*, *Glycyrrhiza glabra*, *Euphorbia seguieriana*, and *Setaria pumila*.

Table 7. Variance value of the principal components of Boochs, Datt2, MCARI, PRI*CI2, SR8, TVI, Vogel-mann2, Vogelmann4 *Ambrosia artemisiifolia*, *Atriplex tatarica*, *Glycyrrhiza glabra*, *Euphorbia seguieriana*, and *Setaria pumila*.

Statistics	Comp.1	Comp.2	Comp.3
Standard deviation	2.27126	1.503265	0.521665
Proportion of variance	0.651341	0.285329	0.03436
Cumulative proportion	0.651341	0.93667	0.97103

Table 8. VIs factor loadings on the main components for *Ambrosia artemisiifolia*, *Atriplex tatarica*, *Glycyrrhiza glabra*, *Euphorbia seguieriana*, and *Setaria pumila*.

VI	Comp.1	Comp.2
Boochs	0.405	0.191
Datt2	0.331	−0.427
MCARI	0.367	0.336
PRI*CI2	0.379	0.233
SR8	−0.364	−0.219
TVI	0.417	0.124
Vogelmann2	−0.263	0.528
Vogelmann4	−0.269	0.521

3.2.2. Decision Tree (DT)

To avoid a subjective approach when choosing VIs and spectral channels, the DT method was used (Figure 12). This method, according to VIs, without alternative (dichotomously), divided weeds into two groups in three stages: perennial weeds (*G. glabra*, *E. seguieriana*) and annual weeds (*A. artemisiifolia*, *S. pumila*, *A. tatarica*). The division of species occurred according to the values of only three Vis: CI, D1, and Datt3.

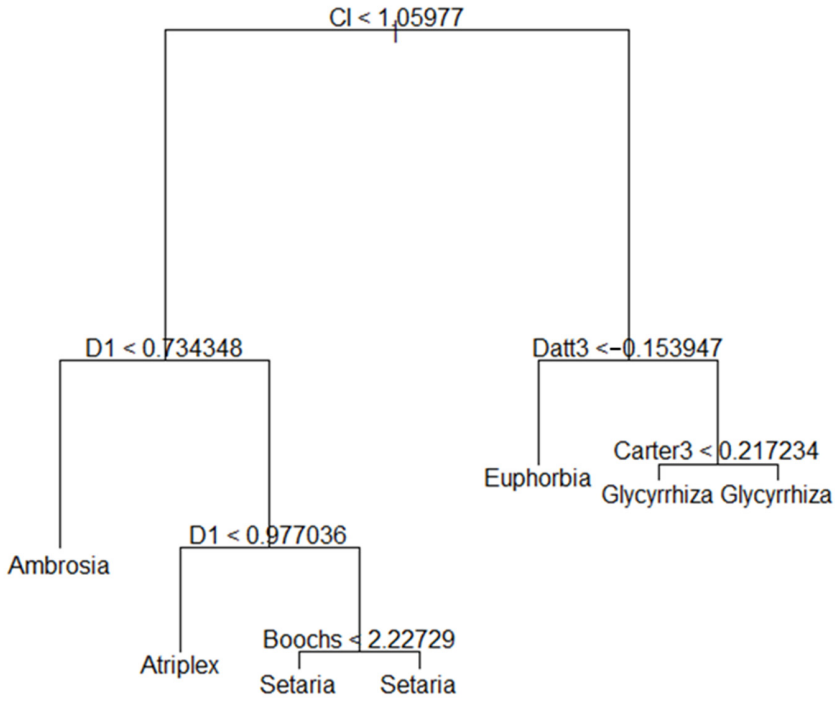

Figure 12. VI DT for *Ambrosia artemisiifolia*, *Atriplex tatarica*, *Glycyrrhiza glabra*, *Euphorbia seguieriana*, and *Setaria pumila*.

The PCA for these VIs provides good results (Figure 13, Tables 9 and 10). The variance is relatively evenly distributed over two main components and amounts to 89%; VIs have medium and high factor loadings (Table 10). The projection of VIs values shows a good differentiation of the group of perennial weeds from the group of annual weeds (Figure 13).

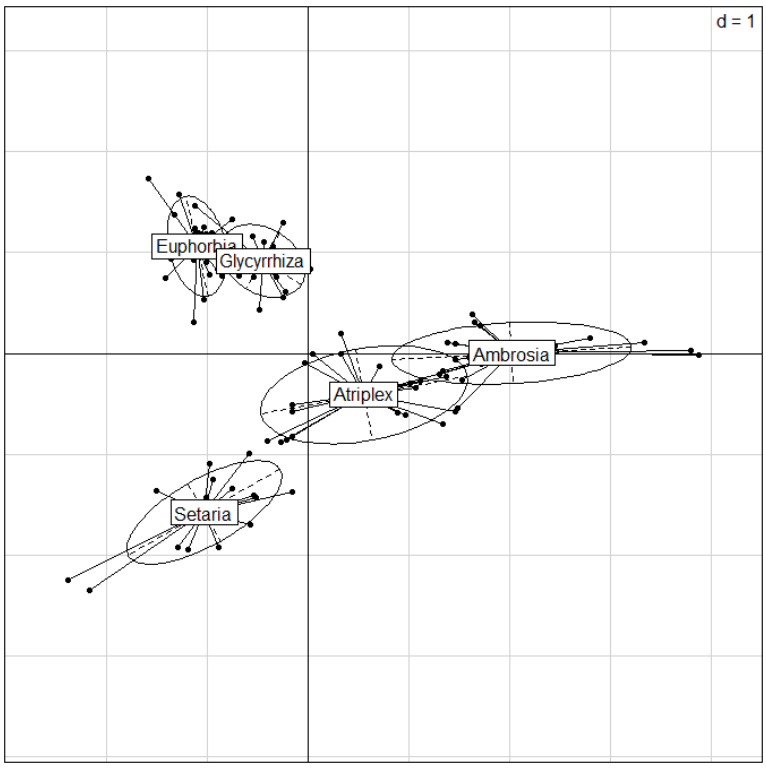

Figure 13. Projection of CI, D1, Datt3 values for *Ambrosia artemisiifolia*, *Atriplex tatarica*, *Glycyrrhiza glabra*, *Euphorbia seguieriana*, and *Setaria pumila*.

Table 9. The value of the dispersion of the principal components of the projection CI, D1, Datt3 *Ambrosia artemisiifolia*, *Atriplex tatarica*, *Glycyrrhiza glabra*, *Euphorbia seguieriana*, and *Setaria pumila*.

Statistics	Comp.1	Comp.2	Comp.3
Standard deviation	1.282888	1.00116	0.567342
Proportion of variance	0.554142	0.337482	0.108376
Cumulative proportion	0.554142	0.891624	1

Table 10. Factor loads on the main components for *Ambrosia artemisiifolia*, *Atriplex tatarica*, *Glycyrrhiza glabra*, *Euphorbia seguieriana*, and *Setaria pumila*.

VI	Comp.1	Comp.2
CI	0.401	0.818
D1	0.578	−0.576
Datt3	−0.710	0

The DT method divided the weed species according to the values of the spectral channels in four stages (Figure 14). The species were separated according to wavelengths: 454 nm, 510 nm, 514 nm, 526 nm, 738 nm, 886 nm, 902 nm, 914 nm, and 962 nm. These channels correspond to the blue (2 channels), green (3 channels), far red (1 channel), and near infrared (3 channels) parts of the electromagnetic spectrum. When divided into clusters, weed species had an alternative. Thus, *E. seguieriana* simultaneously fell into different clusters of the higher hierarchy, while other species (*A. artemisiifolia* and *A. tatarica*) ended up simultaneously in different clusters of the lower hierarchy. This nature of the separation of objects does not correspond to the purpose of the experiment; the differentiation of weed species. Visualization of the result of species differentiation using PCA by spectral channels is shown in Figure 15 and in Tables 11 and 12. Provided that the variance of the first two principal components is high (91%), the factor loadings of the spectral channels on the components are low. This result does not meet the requirements of the experimental work.

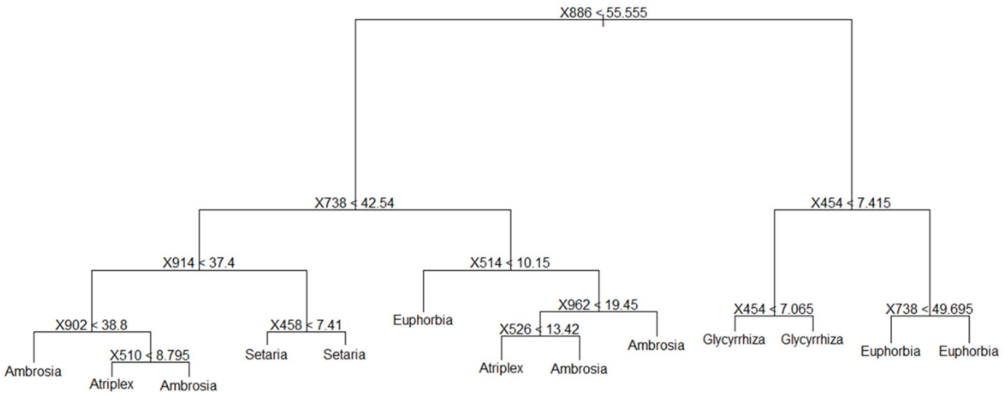

Figure 14. Wavelength DT for *Ambrosia artemisiifolia*, *Atriplex tatarica*, *Glycyrrhiza glabra*, *Euphorbia seguieriana*, and *Setaria pumila*.

Table 11. The value of the dispersion of the main components of the projection at wavelengths of 454 nm, 510 nm, 514 nm, 526 nm, 738 nm, 886 nm, 902 nm, 914 nm, 962 nm for *Ambrosia artemisiifolia*, *Atriplex tatarica*, *Glycyrrhiza glabra*, *Euphorbia seguieriana*, and *Setaria pumila*.

Statistics	Comp.1	Comp.2	Comp.3
Standard deviation	2.2567781	1.7478754	0.77501095
Proportion of variance	0.5716103	0.3428809	0.06741212
Cumulative proportion	0.5716103	0.9144911	0.98190325

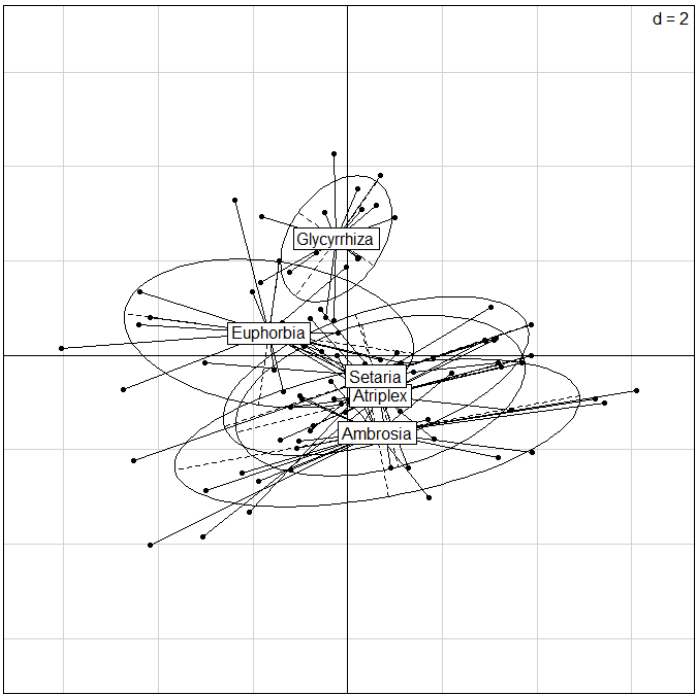

Figure 15. Wavelength projection 454 nm, 510 nm, 514 nm, 526 nm, 738 nm, 886 nm, 902 nm, 914 nm, 962 nm for *Ambrosia artemisiifolia*, *Atriplex tatarica*, *Glycyrrhiza glabra*, *Euphorbia seguieriana*, and *Setaria pumila.*

Table 12. Factor loadings by spectral channels on significant components for *Ambrosia artemisiifolia*, *Atriplex tatarica*, *Glycyrrhiza glabra*, *Euphorbia seguieriana*, and *Setaria pumila*.

Wavelength, nm	Comp.1	Comp.2
886	0.357	0.313
738	0.355	0.208
454	0.309	−0.384
914	0.356	0.325
514	0.301	−0.413
902	0.358	0.328
962	0.315	0.195
510	0.287	−0.429
526	0.351	−0.326

3.2.3. Random Forest (RF)

The next method used to separate species is random forest. The compilation of the confusion matrix for the separation of weeds was carried out by specimens (each species is represented by 20 specimens). Each weed specimen was characterized by 80 VIs values. Weed identification prediction errors based on 80 VIs values are presented in Table 13.

Table 13. Confusion matrix for weed identification, 80 VI.

Confusion Matrix	*Ambrosia*	*Atriplex*	*Euphorbia*	*Glycyrrhiza*	*Setaria*	Class Error
Ambrosia	20	0	0	0	0	0
Atriplex	0	20	0	0	0	0
Euphorbia	0	1	16	3	0	0.2
Glycyrrhiza	0	0	1	19	0	0.05
Setaria	0	1	0	0	19	0.05

Number of trees: 500; No. of variables tried at each split: 8; OOB estimate of error rate: 6%.

In this case, out-of-bag (OOB) error is 6%. The RF method also makes it possible to determine VIs that are most suitable for species identification. In Figure 16 VIs are arranged depending on their influence on the error value and the Gini criterion.

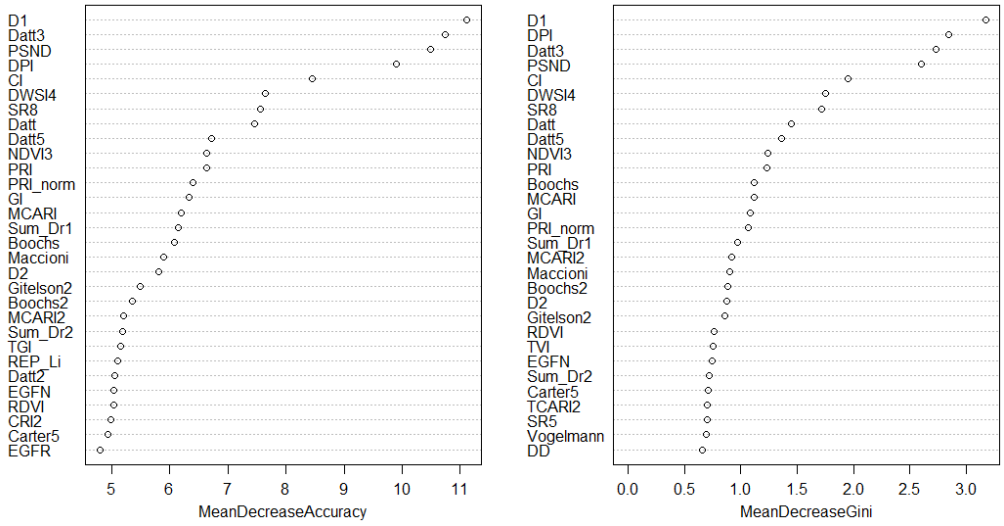

Figure 16. Mean Decrease Accuracy and Mean Decrease Gini.

The most significant VIs are D1, Datt3, PSND, and DPI.

Next, a confusion matrix was compiled for weed separation using the most significant VIs.

The first option is D1, Datt3, and CI (according to the results of DT) (Table 14).

The second option is D1, Datt3, PSND, and DPI (based on RF results) (Table 15).

The third option is D1, Datt3, and PSND (according to the results of paired comparison using *t*-test) (Table 16).

Table 14. Confusion matrix for weed identification by D1, Datt3, and CI.

Confusion Matrix	*Ambrosia*	*Atriplex*	*Euphorbia*	*Glycyrrhiza*	*Setaria*	Class Error
Ambrosia	20	0	0	0	0	0
Atriplex	0	20	0	0	0	0
Euphorbia	0	0	18	2	0	0.10
Glycyrrhiza	0	0	2	18	0	0.10
Setaria	0	0	0	0	20	0

Number of trees: 500; No. of variables tried at each split: 1; OOB estimate of error rate: 4%.

Table 15. Confusion matrix for weed identification by D1, Datt3, PSND, and DPI.

Confusion Matrix	*Ambrosia*	*Atriplex*	*Euphorbia*	*Glycyrrhiza*	*Setaria*	Class Error
Ambrosia	20	0	0	0	0	0
Atriplex	0	20	0	0	0	0
Euphorbia	0	0	18	2	0	0.10
Glycyrrhiza	0	0	1	19	0	0.05
Setaria	0	0	1	0	19	0.05

Number of trees: 500; No. of variables tried at each split: 2; OOB estimate of error rate: 4%.

Table 16. Confusion matrix for weed identification by D1, Datt3, and PSND.

Confusion Matrix	*Ambrosia*	*Atriplex*	*Euphorbia*	*Glycyrrhiza*	*Setaria*	Class Error
Ambrosia	20	0	0	0	0	0
Atriplex	0	20	0	0	0	0
Euphorbia	0	0	19	1	0	0.10
Glycyrrhiza	0	0	1	19	0	0.05
Setaria	0	0	2	0	18	0.05

Number of trees: 500; No. of variables tried at each split: 1; OOB estimate of error rate: 4%.

Thus, the number of VIs for separating *A. artemisiifolia*, *A. tatarica*, *G. glabra*, *E. seguieriana*, and *S. pumila* can be significantly reduced.

4. Discussion

The use of vegetation indices calculated from hyperspectral survey data has great prospects for precision agriculture [82]. It is possible to detect perennial weeds in real time and selectively destroy them with herbicides [9].

4.1. Using Spectral Channels and Vis

In several studies, it is recommended to use spectral channels to identify weeds according to hyperspectral survey data [83,84]. In the present study, a comparison was made of the efficiency of using spectral channels and VIs for the identification of weeds. The results obtained showed better results when using VIs than when using spectral channels. This can be explained by the fact that VIs carry more information than spectral channels. In addition, VIs values correlate with specific physiological characteristics of plants. [85].

4.2. Comparison of the Result Obtained by Different Methods

The results obtained by data analysis methods (PCA, DT, and RF) were compared with the results obtained by mathematical statistics methods (Supplementary Table S3, Figure 13). To perform this, consider a paired comparison of weed species using the *t*-test (Supplementary Table S3). All pairs of species are significantly different in VI CI value («*Ambrosia* vs. *Atriplex*», «*Ambrosia* vs. *Euphorbia*», «*Ambrosia* vs. *Glycyrrhiza*», «*Atriplex* vs. *Euphorbia*», «*Atriplex* vs. *Glycyrrhiza*», «*Atriplex* vs. *Setaria*», «*Euphorbia* vs. *Setaria*», and «*Glycyrrhiza* vs. *Setaria*»), except for «*Ambrosia* vs. *Setaria*» and «*Euphorbia* vs. *Glycyrrhiza*». In terms of Datt3 VI, the difference between the «*Ambrosia* vs. *Atriplex*» and «*Euphorbia* vs. *Setaria*» pairs is not statistically significant; however, the difference in the «*Euphorbia* vs. *Glycyrrhiza*» pair is highly significant. According to the D1 value, the pairs «*Ambrosia* vs. *Atriplex*», «*Ambrosia* vs. *Glycyrrhiza*», and «*Atriplex* vs. *Setaria*» are statistically significantly different, and the pairs «*Atriplex* vs. *Glycyrrhiza*» and «*Euphorbia* vs. *Glycyrrhiza*» do not differ. This is in complete agreement with the results of the DT analysis. Thus, it was shown that with the help of statistical methods, the fineness of division into clusters by the DT method can be estimated.

Good results for the identification of species according to hyperspectral survey data are obtained by the RF method [86–89]. It is shown that if VIs are selected that significantly affect Mean Decrease Accuracy and Mean Decrease Gini (D1, Datt3, PSND, DPI), then OOB error decreases compared with using all 80 VIs. This method distinguishes well all species except *E. seguieriana*. The methods used, with the exception of PCA, showed good agreement between the results of selecting VIs that are effective for identifying weeds in agrocenoses of grain crops when surveying from a short distance (Table 17).

Table 17. Best VIs for weeds identification.

Analysis Method	VI
ANOVA, t-критерий	D1, Datt3, PSND
Principal component analysis	Boochs, Datt2, MCARI, PRI*CI2, SR8, TVI, Vogelmann2, Vogelmann4
Decision tree	CI, D1, Datt3
Random forest	D1, Datt3, PSND, DPI.

Reducing the number of VIs reduces the number of spectral channels. This makes it possible to create simpler and cheaper sensors for solving problems of weeds identification.

4.3. Significance of Weed and Invasive Species Identification Results

The obtained results can be of practical use. It is advisable to treat fields with herbicides after harvesting cultivated plants only against perennial rhizomatous (*G. glabra*) and

rhizomatous weeds (*E. seguieriana*). Annual segetal weeds, as a rule, have time to form and spread many seeds during the growing season of a cultivated plants (*A. artemisiifolia*, *S. pumila*). They are characterized by perennial soil seed banks. Controlling them with herbicides after harvesting crops is not effective. Rhizome and root offspring perennial weeds produce few seeds under culture conditions (or do not have time to form seeds at all) and spread mainly vegetatively, often with mechanized tillage. When using remote sensing means, it becomes possible to identify perennial weeds in real time and selectively destroy them with herbicides.

Hyperspectral cameras are currently expensive devices and cannot be widely used in agriculture. It is necessary to design multispectral cameras with specified characteristics. So, when designing multispectral cameras for remote sensing, one can use spectral channels that are used to calculate VIs that best separate weed species.

5. Conclusions

The results of the study showed the high efficiency of using VIs calculated from hyperspectral survey data to identify various types of weeds in agrocenoses of grain crops when shooting from a short distance. Using the methods of mathematical statistics (ANOVA and *t*-test) and search methods of analysis (principal component analysis, decision tree, and random forest), a group of VIs was determined that allow a good separation of weeds.

The combination of PCA and DT methods allows us to well-differentiate weeds into groups: perennial weeds (*Glycyrrhiza glabra* and *Euphorbia seguieriana*) and annual weeds (*Ambrosia artemisiifolia*, *Atriplex tatarica*, and *Setaria pumila*).

Using the RF method, the Vis that have the greatest impact on the Mean Decrease Accuracy and Mean Decrease Gini were identified. These are D1, Datt3, PSND, and DPI. A great similarity of the results obtained with the help of statistical methods and search methods of data analysis was noted.

When differentiating weed species, VIs showed a better result than spectral channels. This can be explained by the fact that VIs carries more information than spectral channels. VIs values are associated with specific physiological characteristics of plants. Weeds were in different qualitative states: different stages of ontogenesis and regeneration. We believe that it was for this reason that weeds were successfully differentiated into species. The use of VIs for species differentiation may not be effective for plants of the same life form, similar ontogenesis, and phenology.

The ability to divide weeds into perennial rhizomatous, rhizomatous plants, and annual weeds is of great practical importance. After harvesting cultivated plants, it is advisable to control only perennial weeds. The use of remote sensing tools, if possible, to identify perennial weeds in real time makes it possible to selectively destroy them with herbicides. Additional experiments are necessary to determine the repeatability obtained across seasons and the optimum distance to the imaging target (i.e., altitude of a UAV). Additionally, we will increase the number of species under consideration and create a spectral library of weed species.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/rs14102442/s1, Supplementary Table S1. Shapiro–Wilk (1), Pearson's chi-squared (2), Lilliefors (3), Cramér–von Mises (4) norm tests of invasive and weed species; Supplementary Table S2. ANOVA test of invasive and weed species; Supplementary Table S3. Paired comparison of invasive and weed species using t-test; Supplementary Table S4. Factor loads on 4 significant components for Ambrosia artemisiifolia, Atriplex tatarica, Glycyrrhiza glabra, Euphorbia seguieriana, Setaria pumila; Supplementary Table S5. Factor loads on 5 significant components for Ambrosia artemisiifolia, Atriplex tatarica, Glycyrrhiza *glabra, Euphorbia seguieriana, Setaria pumila*.

Author Contributions: Experiment planning, P.A.D., B.L.K., V.K.T., T.M.M. and T.V.V.; collecting data, P.A.D., D.P.K., A.A.D., E.P.T. and O.A.K.; processing data, P.A.D., B.L.K. and A.A.D.; analysis of results, P.A.D. and B.L.K.; publication preparation, A.A.D., V.D.R., V.A.C., O.A.K., V.K.T. and T.M.M. All authors have read and agreed to the published version of the manuscript.

Funding: The research was financially supported by the Ministry of Science and Higher Education of the Russian Federation within the framework of the state task in the field of scientific activity (No. 0852-2020-0029).

Data Availability Statement: Not applicable.

Acknowledgments: The research was performed with the equipment of Multiaccess Center 'Biotechnology, biomedicine and environmental monitoring' and Multiaccess Center 'High technologies' of Southern Federal University (Rostov-on-Don).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tokhtar, V.; Groshenko, S.A. Differentiation of the Climatic Niches of the Invasive *Oenothera* L. (Subsect. Oenothera, Onagraceae) Species in the Eastern Europe. *Adv. Environ. Biol.* **2014**, *8*, 529–531. [CrossRef]
2. Tokhtar, V.K. Advanced Approaches to the Visualization of Data Characterizing Distribution Features of Alien Plant Species. *Russ. J. Biol. Invasions* **2018**, *9*, 263–269. [CrossRef]
3. Bzdęga, K.; Zarychta, A.; Urbisz, A.; Szporak-Wasilewska, S.; Ludynia, M.; Fojcik, B.; Tokarska-Guzik, B. Geostatistical models with the use of hyperspectral data and seasonal variation—A new approach for evaluating the risk posed by invasive plants. *Ecol. Indic.* **2021**, *121*, 107204. [CrossRef]
4. McGeoch, M.A.; Butchart, S.H.M.; Spear, D.; Marais, E.; Kleynhans, E.J.; Symes, A.; Chanson, J.; Hoffmann, M. Global indicators of biological invasion: Species numbers, biodiversity impact and policy responses. *Divers. Distrib.* **2010**, *16*, 95–108. [CrossRef]
5. Van Kleunen, M.; Weber, E.; Fischer, M. A meta-analysis of trait differences between invasive and non-invasive plant species. *Ecol. Lett.* **2010**, *13*, 235–245. [CrossRef]
6. Lu, B.; Dao, P.; Liu, J.; He, Y.; Shang, J. Recent Advances of Hyperspectral Imaging Technology and Applications in Agriculture. *Remote Sens.* **2020**, *12*, 2659. [CrossRef]
7. Kok, L. Classical Biological Control of Nodding and Plumeless Thistles. *Biol. Control* **2001**, *21*, 206–213. [CrossRef]
8. Gerhards, R.; Oebel, H. Practical experiences with a system for site-specific weed control in arable crops using real-time image analysis and GPS-controlled patch spraying. *Weed Res.* **2006**, *46*, 185–193. [CrossRef]
9. Slaughter, D.; Giles, D.; Downey, D. Autonomous robotic weed control systems: A review. *Comput. Electron. Agric.* **2008**, *61*, 63–78. [CrossRef]
10. Sheffield, K.J.; Dugdale, T.M. Supporting Urban Weed Biosecurity Programs with Remote Sensing. *Remote Sens.* **2020**, *12*, 2007. [CrossRef]
11. Everitt, J.H.; Alaniz, M.A.; Escobar, D.E.; Davis, M.R. Using Remote Sensing to Distinguish Common (*Isocoma coronopifolia*) and Drummond Goldenweed (*Isocoma drummondii*). *Weed Sci.* **1992**, *40*, 621–628. [CrossRef]
12. Lamb, D.W.; Brown, R.B. PA—Precision Agriculture: Remote-sensing and mapping of weeds in crops. *J. Agric. Eng. Res.* **2001**, *78*, 117–125. [CrossRef]
13. Brown, R.B.; Steckler, J.-P.G.A.; Anderson, G.W. Remote sensing for identifcation of weeds in no-till corn. *Trans. ASAE* **1994**, *37*, 297–302. [CrossRef]
14. Shearer, S.A.; Holmes, R.G. Plant identication using color co occurrence matrices. *Trans. ASAE* **1990**, *33*, 2037–2044. [CrossRef]
15. Liakos, K.G.; Busato, P.; Moshou, D.; Pearson, S.; Bochtis, D. Machine learning in agriculture: A review. *Sensors* **2018**, *18*, 2674. [CrossRef]
16. Chandler, J.; Cooke, F. Economics of cotton losses caused by weeds. In *Weeds Cotton: Characterization and Control*; The Cotton Foundation: Memphis, TN, USA, 1992; pp. 85–116.
17. Su, W.-H. Advanced Machine Learning in Point Spectroscopy, RGB- and Hyperspectral-Imaging for Automatic Discriminations of Crops and Weeds: A Review. *Smart Cities* **2020**, *3*, 767–792. [CrossRef]
18. Suzuki, Y.; Okamoto, H.; Kataoka, T. Image Segmentation between Crop and Weed using Hyperspectral Imaging for Weed Detection in Soybean Field Environ. *Control Biol.* **2008**, *46*, 163–173. [CrossRef]
19. Bayer, B. Color Imaging Array. U.S. Patent 3,971,065, 20 July 1976.
20. Manh, A.-G.; Rabatel, G.; Assemat, L.; Aldon, M.-J. AE—Automation and Emerging Technologies: Weed Leaf Image Segmentation by Deformable Templates. *J. Agric. Eng. Res.* **2001**, *80*, 139–146. [CrossRef]
21. Su, W.-H.; Sun, D.-W. Facilitated wavelength selection and model development for rapid determination of the purity of organic spelt (*Triticum spelta* L.) flour using spectral imaging. *Talanta* **2016**, *155*, 347–357. [CrossRef]
22. El-Faki, M.S.; Zhang, N.; Peterson, D.E. Factors affecting color-based weed detection. *Trans. ASAE* **2000**, *43*, 1001–1009. [CrossRef]
23. Noble, S.; Crowe, T. Plant discrimination based on leaf reflectance. In Proceedings of the 2001 ASAE Annual International Meeting, Meeting Paper No. 011150, Sacramento, CA, USA, 30 July–1 August 2001; ASAE: St Joseph, MI, USA, 2001. [CrossRef]
24. Terawaki, M.; Kataoka, T.; Okamoto, H.; Hata, S. Distinction between sugar beet and weeds based on shape characteristics using image processing technique. *J. JSAM* **2002**, *64*, 93–101.
25. Okamoto, H.; Murata, T.; Kataoka, T.; Hata, S.-I. Plant classification for weed detection using hyperspectral imaging with wavelet analysis. *Weed Biol. Manag.* **2007**, *7*, 31–37. [CrossRef]

26. Burks, T.; Shearer, S.; Payne, F. Classification of weed species using color texture features and discriminant analysis. *Trans. ASAE* **2000**, *43*, 441–448. [CrossRef]
27. Peerbhay, K.; Mutanga, O.; Ismail, R. The identification and remote detection of alien invasive plants in commercial forests: An Overview. *S. Afr. J. Geomat.* **2016**, *5*, 49–67.
28. Ishii, J.; Washitani, I. Early detection of the invasive alien plant *Solidago altissima* in moist tall grassland using hyperspectral imagery. *Int. J. Remote Sens.* **2013**, *34*, 5926–5936. [CrossRef]
29. Andrew, M.E.; Susan, L. Ustin The role of environmental context in mapping invasive plants with hyperspectral image data. *Remote Sens. Environ.* **2008**, *112*, 4301–4317. [CrossRef]
30. Maes, W.H.; Steppe, K. Estimating evapotranspiration and drought stress with ground-based thermal remote sensing in agriculture: A review. *J. Exp. Bot.* **2012**, *63*, 4671–4712. [CrossRef]
31. Panov, V.D.; Lurie, P.M.; Larionov, Y.A. *The Climate of the Rostov Region: Yesterday, Today, Tomorrow*; Donskoy Publishing House: Rostov-on-Don, Russia, 2006; p. 488.
32. Bareth, G.; Aasen, H.; Bendig, J.; Gnyp, M.L.; Bolten, A.; Jung, A.; Michels, R.; Soukkamäki, J. Low-weight and UAV-based hyperspectral full-frame cameras for monitoring crops: Spectral comparison with portable spectroradiometer measurements. *Photogr. Fernerkund. Geoinf.* **2015**, *1*, 69–79. [CrossRef]
33. Aasen, H.; Burkart, A.; Bolten, A.; Bareth, G. Generating 3D hyperspectral information with lightweight UAV snapshot cameras for vegetation monitoring: From camera calibration to quality assurance. *ISPRS J. Photogramm. Remote Sens.* **2015**, *108*, 245–259. [CrossRef]
34. Boochs, F.; Kupfer, G.; Dockter, K.; Kühbauch, W. Shape of the red edge as vitality indicator for plants. *Int. J. Remote Sens.* **1990**, *11*, 1741–1753. [CrossRef]
35. Kim, M.; Daughtry, C.; Chappelle, E.; McMurtrey, J.; Walthall, C. The use of high spectral resolution bands for estimating absorbed photosynthetically active radiation (Apar). In Proceedings of the Sixth Symposium on Physical Measurements and Sig-natures in Remote Sensing, Val d'Isère, France, 17–21 January 1994; Volume 17, pp. 299–306.
36. Carter, G.A. Ratios of leaf reflectances in narrow wavebands as indicators of plant stress. *Int. J. Remote Sens.* **1994**, *15*, 697–703. [CrossRef]
37. Zarco-Tejada, P.J.; Pushnik, J.C.; Dobrowski, S.; Ustin, S.L. Steady-state chlorophyll a fluorescence detection from canopy derivative reflectance and double-peak red-edge effects. *Remote Sens. Environ.* **2003**, *84*, 283–294. [CrossRef]
38. Gitelson, A.A.; Gritz, Y.; Merzlyak, M.N. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *J. Plant Physiol.* **2003**, *160*, 271–282. [CrossRef] [PubMed]
39. Oppelt, N.; Mauser, W. Hyperspectral monitoring of physiological parameters of wheat during a vegetation period using AVIS data. *Int. J. Remote Sens.* **2004**, *25*, 145–159. [CrossRef]
40. Datt, B. Visible/near infrared reflectance and chlorophyll content in Eucalyptus leaves. *Int. J. Remote Sens.* **1999**, *20*, 2741–2759. [CrossRef]
41. Datt, B. Remote Sensing of Chlorophyll a, Chlorophyll b, Chlorophyll a+b, and Total Carotenoid Content in Eucalyptus Leaves. *Remote Sens. Environ.* **1998**, *66*, 111–121. [CrossRef]
42. Le Maire, G.; François, C.; Dufrêne, E. Towards universal broad leaf chlorophyll indices using PROSPECT simulated database and hyperspectral reflectance measurements. *Remote Sens. Environ.* **2004**, *89*, 1–28. [CrossRef]
43. Apan, A.; Held, A.; Phinn, S.; Markley, J. Detecting sugarcane 'orange rust' disease using EO-1 Hyperion hyperspectral imagery. *Int. J. Remote Sens.* **2004**, *25*, 489–498. [CrossRef]
44. Penuelas, J.; Gamon, J.A.; Fredeen, A.L.; Merino, J.; Field, C.B. Reflectance indices associated with physiological-changes in nitrogen-limited and water-limited sun ower leaves. *Remote Sens. Environ.* **1994**, *48*, 135–146. [CrossRef]
45. Huete, A.; Liu, H.Q.; Batchily, K.; van Leeuwen, W. A comparison of vegetation indices over a global set of TM images for EOS–MODIS. *Remote Sens. Environ.* **1997**, *59*, 440–451. [CrossRef]
46. Smith, R.; Adams, J.; Stephens, D.; Hick, P. Forecasting wheat yield in a Mediterranean-type environment from the NOAA satellite. *Aust. J. Agric. Res.* **1995**, *46*, 113–125. [CrossRef]
47. Gitelson, A.; Buschmann, C.; Lichtenthaler, H. The chlorophyll fluorescence ratio F735/F700 as an accurate measure of the chlorophyll content in plants—Experiments with autumn chestnut and maple leaves. *Remote Sens. Environ.* **1999**, *69*, 296–302. [CrossRef]
48. 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]
49. Maccioni, A.; Agati, G.; Mazzinghi, P. New vegetation indices for remote measurement of chlorophylls based on leaf directional reflectance spectra. *J. Photochem. Photobiol. B Biol.* **2001**, *61*, 52–61. [CrossRef]
50. Daughtry, C.; Walthall, C.; Kim, M.; Brown de Colstoun, E.; McMurtrey, J.E., III. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sens. Environ.* **2000**, *74*, 229–239. [CrossRef]
51. Hernández-Clemente, R.; Navarro-Cerrillo, R.M.; Suárez, L.; Morales, F.; Zarco-Tejada, P.J. Assessing structural effects on PRI for stress detection in conifer forests. *Remote Sens. Environ.* **2011**, *115*, 2360–2375. [CrossRef]
52. Qi, J.; Chehbouni, A.; Huete, A.R.; Kerr, Y.H.; Sorooshian, S. A modified soil adjusted vegetation index. *Remote Sens. Environ.* **1994**, *48*, 119–126. [CrossRef]

53. Chen, J.M. Evaluation of Vegetation Indices and a Modified Simple Ratio for Boreal Applications. *Can. J. Remote Sens.* **1996**, *22*, 229–242. [CrossRef]
54. Dash, J.; Curran, P.J. The MERIS terrestrial chlorophyll index. *Int. J. Remote Sens.* **2004**, *25*, 5403–5413. [CrossRef]
55. Haboudane, D.; Miller, J.R.; Tremblay, N.; Zarco-Tejada, P.J.; Dextraze, L. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sens. Environ.* **2002**, *81*, 416–426. [CrossRef]
56. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [CrossRef]
57. 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]
58. Gandia, S.; Fernandez, G.; Garcia, J.; Moreno. Retrieval of vegetation biophysical variables from CHRIS/PROBA data in the SPARC campaign. *Esa SP* **2004**, *578*, 40–48.
59. Rondeaux, G.; Steven, M.; Baret, F. Optimization of soil-adjusted vegetation indices. *Remote Sens. Environ.* **1996**, *55*, 95–107. [CrossRef]
60. Wu, C.; Niu, Z.; Tang, Q.; Huang, W. Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation. *Agric. For. Meteorol.* **2008**, *148*, 1230–1241. [CrossRef]
61. Chappelle, E.W.; Kim, M.S.; McMurtrey, J.E. Ratio analysis of reflectance spectra (rars)—An algorithm for the remote esti-mation of the concentrations of chlorophyll-a, chlorophyll-b, and carotenoids in soybean leaves. *Remote Sens. Environ.* **1992**, *39*, 239–247. [CrossRef]
62. Gamon, J.; Penuelas, J.P.; Field, C. A narrow-waveband spectral index that tracks diurnal changes in photosynthetic effi-ciency. *Remote Sens. Environ.* **1992**, *41*, 35–44. [CrossRef]
63. Zarco-Tejada, P.J.; Gonzalez-Dugo, V.; Williams, L.E.; Suárez, L.; Berni, J.A.J.; Goldhamer, D.; Fereres, E. A PRI-based water stress index combining structural and chlorophyll effects: Assessment using diurnal narrow-band airborne imagery and the CWSI thermal index. *Remote Sens. Environ.* **2013**, *138*, 38–50. [CrossRef]
64. Garrity, S.R.; Eitel, J.U.; Vierling, L.A. Disentangling the relationships between plant pigments and the photochemical re-flectance index reveals a new approach for remote estimation of carotenoid content. *Remote Sens. Environ.* **2011**, *115*, 628–635. [CrossRef]
65. Merzlyak, M.N.; Gitelson, A.A.; Chivkunova, O.B.; Rakitin, V.Y. Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. *Physiol. Plant.* **1999**, *106*, 135–141. [CrossRef]
66. Blackburn, G.A. Quantifying Chlorophylls and Caroteniods at Leaf and Canopy Scales: An Evaluation of Some Hyperspectral Approaches. *Remote Sens. Environ.* **1998**, *66*, 273–285. [CrossRef]
67. Roujean, J.-L.; Breon, F.-M. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sens. Environ.* **1995**, *51*, 375–384. [CrossRef]
68. Guyot, G.; Baret, F. Utilisation de la haute resolu-tion spectrale pour suivre l'etat des couverts vegetaux. *Spectr. Signat. Objects Remote Sens.* **1988**, *287*, 279–286.
69. Huete, A.R. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [CrossRef]
70. Vincini, M.; Frazzi, E.; D'Alessio, P. Angular dependence of maize and sugar beet VIs from directional CHRIS/PROBA data. In Proceedings of the Fourth ESA CHRIS PROBA Workshop ESRIN, Frascati, Italy, 19–21 September 2006.
71. Jordan, C.F. Derivation of Leaf-Area Index from Quality of Light on the Forest Floor. *Ecology* **1969**, *50*, 663–666. [CrossRef]
72. Gitelson, A.A.; Merzlyak, M.N. Remote estimation of chlorophyll content in higher plant leaves. *Int. J. Remote Sens.* **1997**, *18*, 2691–2697. [CrossRef]
73. McMurtrey, J.E.; Chappelle, E.W.; Kim, M.S.; Meisinger, J.J.; Corp, L.A. Distinguishing nitrogen fertilization levels in field corn (*Zea mays* L.) with actively induced fluorescence and passive reflectance measurements. *Remote Sens. Environ.* **1994**, *47*, 36–44. [CrossRef]
74. Zarco-Tejada, P.J.; Miller, J.R. Land cover mapping at BOREAS using red edge spectral parameters from CASI imagery. *J. Geophys. Res. Earth Surf.* **1999**, *104*, 27921–27933. [CrossRef]
75. Hernandez-Clemente, R.; Navarro-Cerrillo, R.M.; Zarco-Tejada, P.J. Carotenoid content estimation in a heterogeneous co-nifer forest using narrowband indices and PROSPECT + DART simulations. *Remote Sens. Environt.* **2012**, *127*, 298–315. [CrossRef]
76. Elvidge, C.D.; Chen, Z.K. Comparison of broad-band and narrow-band red and near-infrared vegetation indices. *Remote Sens. Environ.* **1995**, *54*, 38–48. [CrossRef]
77. Filella, I.; Penuelas, J. The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status. *Int. J. Remote Sens.* **1994**, *15*, 1459–1470. [CrossRef]
78. Hunt, E.R., Jr.; Doraiswamy, P.C.; McMurtrey, J.E.; Daughtry, C.S.T.; Perry, E.M.; Akhmedov, B. A visible band index for remote sensing leaf chlorophyll content at the canopy scale. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *21*, 103–112. [CrossRef]
79. Broge, N.H.; Leblanc, E. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sens. Environ.* **2001**, *76*, 156–172. [CrossRef]
80. Vogelmann, J.E.; Rock, B.N.; Moss, D.M. Red edge spectral measurements from sugar maple leaves. *Int. J. Remote Sens.* **1993**, *14*, 1563–1575. [CrossRef]
81. Lehnert, L.W.; Meyer, H.; Obermeier, W.A.; Silva, B.; Regeling, B.; Thies, B.; Bendix, J. Hyperspectral data analysis in R: The hsdar package. *J. Stat. Softw.* **2019**, *89*, 1–23. [CrossRef]

82. Candiago, S.; Remondino, F.; De Giglio, M.; Dubbini, M.; Gattelli, M. Evaluating Multispectral Images and Vegetation Indices for Precision Farming Applications from UAV Images. *Remote Sens.* **2015**, *7*, 4026–4047. [CrossRef]
83. Farooq, A.; Hu, J.; Jia, X. Analysis of Spectral Bands and Spatial Resolutions for Weed Classification via Deep Convolutional Neural Network. *IEEE Geosci. Remote Sens. Lett.* **2018**, *16*, 183–187. [CrossRef]
84. Farooq, A.; Jia, X.; Hu, J.; Zhou, J. Multi-Resolution Weed Classification via Convolutional Neural Network and Superpixel Based Local Binary Pattern Using Remote Sensing Images. *Remote Sens.* **2019**, *11*, 1692. [CrossRef]
85. Ronay, I.; Ephrath, J.E.; Eizenberg, H.; Blumberg, D.G.; Maman, S. Hyperspectral Reflectance and Indices for Characterizing the Dynamics of Crop–Weed Competition for Water. *Remote Sens.* **2021**, *13*, 513. [CrossRef]
86. De Sá, N.C.; Castro, P.; Carvalho, S.; Marchante, E.; López-Núñez, F.A.; Marchante, H. Mapping the flowering of an invasive plant using unmanned aerial vehicles: Is there potential for biocontrol monitoring? *Front. Plant Sci.* **2018**, *9*, 1–13. [CrossRef]
87. Franklin, S.E.; Ahmed, O.S. Deciduous tree species classification using object-based analysis and machine learning with unmanned aerial vehicle multispectral data. *Int. J. Remote Sens.* **2018**, *39*, 5236–5245. [CrossRef]
88. Miyoshi, G.T.; Imai, N.N.; Tommaselli, A.M.G.; de Moraes, M.V.A.; Honkavaara, E. Evaluation of hyperspectral multitem-poral information to improve tree species identification in the highly diverse atlantic forest. *Remote Sens.* **2020**, *12*, 244. [CrossRef]
89. Dash, J.P.; Watt, M.S.; Paul, T.S.H.; Morgenroth, J.; Pearse, G.D. Early Detection of Invasive Exotic Trees Using UAV and Manned Aircraft Multispectral and LiDAR Data. *Remote Sens.* **2019**, *11*, 1812. [CrossRef]