



Article

Mapping the Invasive Species *Stellera chamaejasme* in Alpine Grasslands Using Ecological Clustering, Spectral Separability and Image Classification

Nianzhao Hu ¹, Yongmei Liu ^{1,2,*}, Xinghua Ge ¹, Xingzhi Dong ¹, Huaiyu Wang ¹, Yongqing Long ^{1,2} 
and Lei Wang ^{1,2}

¹ College of Urban and Environmental Sciences, Northwest University, Xi'an 710127, China

² Shaanxi Key Laboratory of Earth Surface System and Environmental Carrying Capacity, Xi'an 710127, China

* Correspondence: liuym@nwu.edu.cn

Abstract: *Stellera chamaejasme* (Thymelaeaceae) is amongst the worst invasive species of the alpine grasslands on the Qinghai–Tibet Plateau; timely and effective monitoring is critical for its prevention and control. In this study, by using high spatial resolution Planet imagery, an optimal approach was explored to improve the discrimination of *S. chamaejasme* from surrounding communities, integrated with TWINSAPN technique, Transformed divergence and image classification algorithms. Results demonstrated that there were obvious spectral conflicts observed among the TWINSAPN ecological communities, owing to the inconsistency of *S. chamaejasme* coverage within the communities. By determining the threshold of spectral separability, the adjustment of ecological classification produced spectrally separated *S. chamaejasme* communities and native species communities. The sensitive index characterizing the spectra of *S. chamaejasme* contributes to its discrimination; moderate or good classification accuracy was obtained by using four machine learning algorithms, of which Random Forest achieved the highest accuracy of *S. chamaejasme* classification. Our study suggests the distinct phenological feature of *S. chamaejasme* provides a basis for the detection of the toxic weed, and the establishment of communities using the rule of spectral similarity can assist the accurate discrimination of invasive species.



Citation: Hu, N.; Liu, Y.; Ge, X.; Dong, X.; Wang, H.; Long, Y.; Wang, L. Mapping the Invasive Species *Stellera chamaejasme* in Alpine Grasslands Using Ecological Clustering, Spectral Separability and Image Classification. *Agronomy* **2023**, *13*, 593. <https://doi.org/10.3390/agronomy13020593>

Received: 26 December 2022

Revised: 2 February 2023

Accepted: 15 February 2023

Published: 18 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: plant community; TWINSAPN; Transformed divergence; machine learning algorithms; Planet multispectral image

1. Introduction

Alpine grassland is one of the most representative ecosystems on the Qinghai–Tibet Plateau (QTP), and plays an irreplaceable role in maintaining regional ecological balance and animal husbandry sustainability [1,2]. During the last several decades, alpine grasslands have been overgrazed and suffer on-going degradation; the trends indicate remarkable increases of forbs and poisonous plants [3–5]. *Stellera chamaejasme*, a toxic perennial weed, has become established in alpine grasslands on the QTP since the 1960s, and become a major invasive species, seriously threatening a wide range of grasslands [6]. With broad ecological adaptability, *S. chamaejasme* competed and encroached the niches of native forage species, gradually dominating alpine grasslands, and serving as an indicator of moderate to severe grassland degeneration [7]. The toxic weed rapidly spreads on the QTP, especially in alpine meadows and alpine steppes, leading to reasonably adverse impacts on alpine ecosystem functions and local economies [8]. This issue has not yet been paid sufficient attention.

Field surveys of ecological invasion are time-consuming and laborious, particularly if large areas are involved [9,10]. Remote sensing technologies provide efficient methods for invasive plant monitoring [11,12]. High resolution panchromatic and color aerial photography acquired in early stages achieved good results in studying the invasion

process [13,14]. Taking the advantages of extensive coverage and repeated observation, Landsat TM/ETM/OLI multispectral images are successfully used to extract the distribution of woody plants over large areas [15,16]. Hyperspectral images are an ideal source of information for spectrally discriminating invasive species [17–19], but their use is limited by the difficult accessibility. Most recently, unmanned aerial vehicle (UAV) imaging has been increasingly applied in identifying ecological invasive patterns at small scale, benefiting from rapidly providing accurate information of affected areas [20,21]. Compared to invasive trees and shrubs, spectral discrimination of herbaceous plant species is more challenging due to the overlapping of spectral signature. High resolution multispectral satellite imagery provides a spectrally and spatially optimal choice for invasion mapping, achieving a balance between accuracy and scale [13]. QuickBird and SPOT6 images were used to recognize *Avena sterilis* and *Parthenium hysterophorus* with high accuracy [15,22]; *Heracleum mantegazzianum* in large patches could be discerned with moderate accuracy based on RapidEye images [13]. IKONOS images were adopted to effectively monitor the distribution area of leafy spurge with greater than 30% coverage [23]. The studies indicate that high resolution multispectral images hold great potential in monitoring invasive grasses at large scale.

Ecological classification groups species/quadrats data into plant communities, conceptualizing the communities and characterizing their environmental gradients [24,25]. Vegetation mapping mainly specifies the spatial distribution of plant communities based on their spectral differences, but the ecological significance (composition, structure, and pattern, etc.) of the divided classes is ignored, which may result in misclassification. Scholars have discussed how to integrate the techniques of quantitative ecology with image classification to improve vegetation mapping accuracy. Combined canonical correspondence analysis (CCA) and two-way indicator species analysis (TWINSPAN) with spectral separability analysis and maximum likelihood classification (MLC) technique, optimized classification of boreal forest peatland was achieved by using 6 bands CASI images in northern Canada [24]. The integration of TWINSPAN, principal component analysis (PCA) and CCA were applied to examine the spectral correspondence between SPOT multispectral images and the shrub forest communities in the Himalayan foothills [26]. Furthermore, the non-metric multidimensional scaling (MNDs) in combination with imaging spectroscopy provided an expedient approach for mapping floristic compositional patterns in Bavaria swamps and meadows [27]. The findings indicate that there is a lack of obvious relationship between plant species and image spectra, suggesting the difficulty of classification at species level; however, the communities composed of ecologically related species showed definite spectral separability [24]. The conjunction of ecological and remote sensing techniques helps to classify plant communities with clear composition and structure, thus providing an accurate approach for the management of toxic weeds invasion in alpine grasslands.

Previous attempts have been concentrated on using spectral reflectance as the main metric to map vegetation types. The correspondence between spectral signature and plant community provides an important basis for the identification of invasive plants, has not yet been well studied. Our previous study extracted the distribution area of *S. chamaejasme* using IKONOS images on the basis of spectral difference between *S. chamaejasme* and the occurring species, but the ecological significance of the classification was not clarified [28]. Therefore, the main goal of this study is to explore the link between ecological classification and image classification, and propose an combined procedure to map the regional extent of *S. chamaejasme*. The species /community sampling data were used in conjunction with Planet imagery, our aims are as follows: (1) to examine the spectral separability of TWINSPAN classes (*S. chamaejasme* communities and the surrounding communities), and present an image classification scheme defined by spectrally separated communities, still maintaining their ecological significance; (2) to determine the optimal classification algorithm for *S. chamaejasme* discrimination, and evaluate the applicability of high-resolution Planet multispectral imagery in the large-scale monitoring of invasive toxic weeds.

2. Materials and Methods

2.1. Study Area

S. chamaejasme is a perennial herbaceous plant with the height of 20–50 cm, characterized by lanceolate leaves, dark brown capsules, and white-pink inflorescences of capitula [7]. *S. chamaejasme* is distributed in patchy patterns over large areas, its white-pink flowers are distinct from surrounding communities during the flowering phase of late June to late July (Figure 1). The whole plant of *S. chamaejasme* is toxic, and vomiting, abdominal pain, diarrhoea, convulsion, and even death may occur when livestock ingest it [5]. Populations of the toxic plant have been increasing constantly due to long-term degraded succession of alpine grassland, making it a persistent weed problem on the Qinghai–Tibet Plateau, forming typical degraded grasslands invaded by *S. chamaejasme* [29]. In Qinghai Province, *S. chamaejasme* is mainly distributed in Haibei Tibetan Autonomous Prefecture, Hainan Tibetan Autonomous Prefecture, Huangnan Tibetan Autonomous Prefecture, Yushu Tibetan Autonomous Prefecture, and Qinghai Lake area, covering an area of 1.4 million hm² [30].

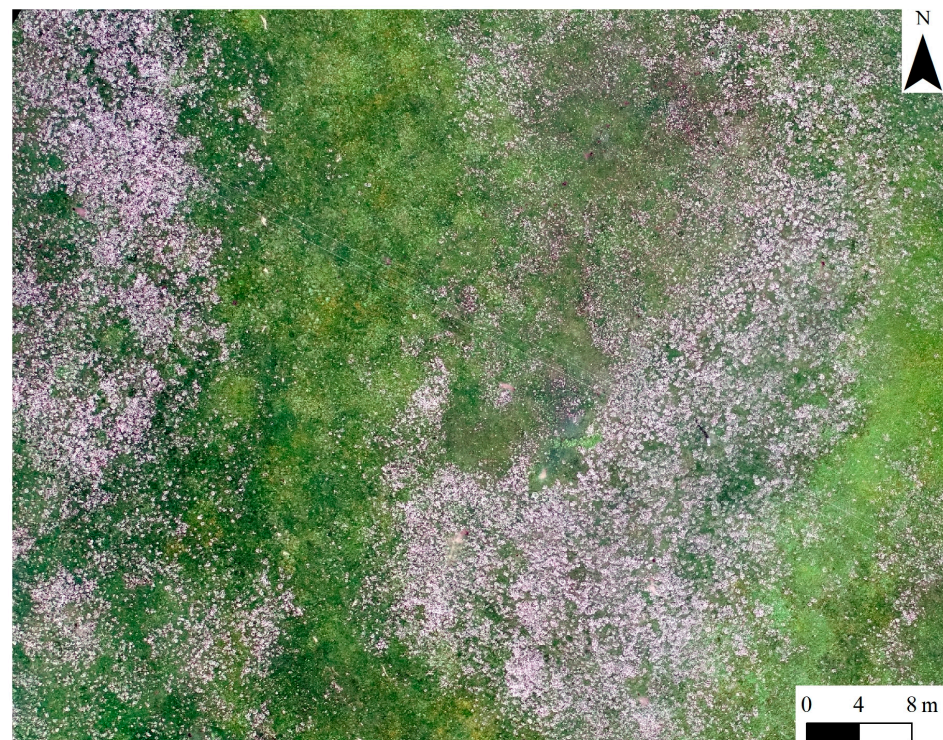


Figure 1. UAV RGB photo *S. chamaejasme* distribution in the study area. Clumped patches of a white-pink color are *S. chamaejasme* communities.

The study took place in Qilian County, Haibei Prefecture, Qinghai Province (at 100°11′15.69″–101°0′34.21″ E, 37°45′41.84″–38°16′30.49″ N, Figure 2). The study area is about 4000 km² and the altitude is 3169 m. It is a typical plateau climate in the region, with annual average temperature of 1 °C and annual precipitation of 420 mm. Vegetation is typical of the degraded alpine meadow, which is used as winter rangeland for the heavy grazing of yaks and sheep, yearly from November to May. Dominant species mainly include *S. chamaejasme*, *Poa annua*, *Kobresia pygmaea*, *Elymus nutans*, *Potentilla fruticosa*, and *Thermopsis lanceolata*. Co-occurring species include *Anemone rivularis*, *Potentilla chinensis*, *Medicago ruthenica*, *Gentiana straminea*, *Taraxacum mongolicum*, and *Plantago asiatica*.

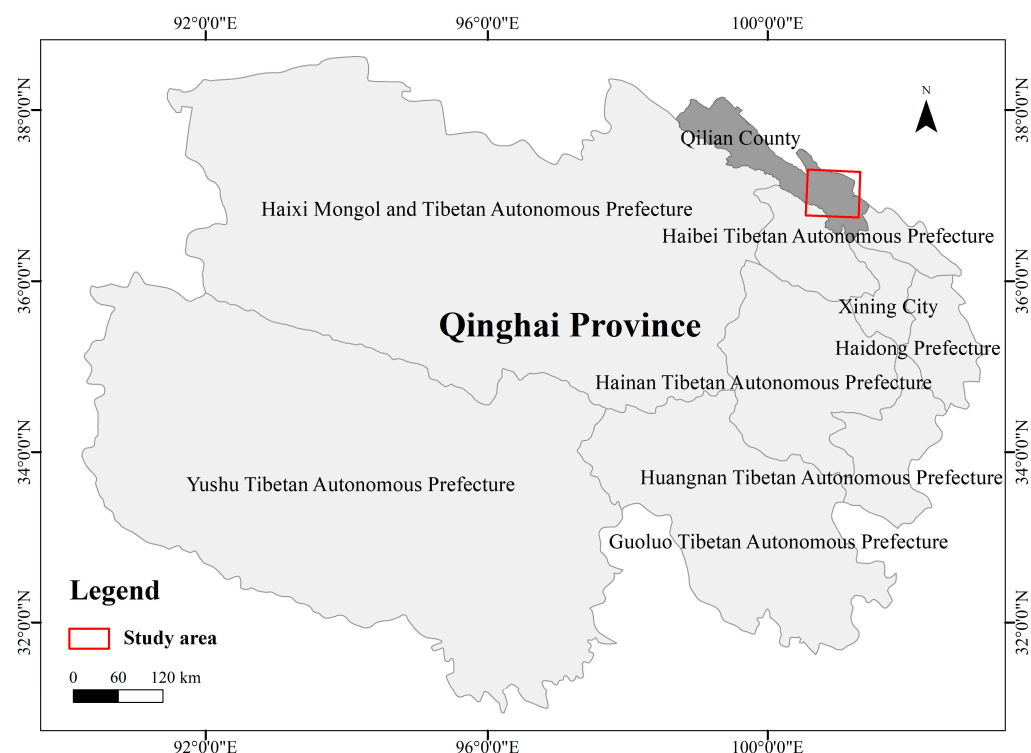


Figure 2. Location of the study area.

2.2. Data Collection

2.2.1. Remote Sensing Data

Planet image is used as the data source of *S. chamaejasme* discrimination. The image is obtained by Planet Labs satellites, which consist of roughly 200 Dove satellites in orbit, and provide high-resolution and complete view of the world with the daily frequency. Five channels were collected with 3 m spatial resolution and 16 bits depth, ranged from visible to near-infrared wavelengths (Table 1). Two images in the flowering phase of *S. chamaejasme* were spliced together for covering the study area, which were taken on 12 and 17 July 2021. The acquired images have been radiometric calibrated, further atmospheric correction and orthorectification were performed by using ENVI 5.3 software.

Table 1. Planet channel wavelengths.

Channel	Waveband (nm)	Range
B1	465–515	blue
B2	547–583	green
B3	650–680	red
B4	697–713	red edge
B5	845–885	near infrared

2.2.2. Field Vegetation Data

In mid-July 2020 and 2021, Field campaigns were carried out during the flowering phase of *S. chamaejasme*. By combining vegetation and terrain with accessibility, two transects were placed along Qingyang Gully and Babao River Valley. The floristic data were collected in 37 plots, each consisting of 2–3 quadrats (1 m × 1 m); 93 quadrats were totally obtained (Figure 3 and Table 2). When sampling, the homogeneous community patch of larger than 3 m × 3 m was selected, and the quadrat was placed in the center of the region, in order to ensure the spectral accuracy of 3 m Planet image representing the community characteristics of 1 m quadrat. In each quadrat, the photograph was taken vertically downward and the species/community coverage were estimated by image interpretation.

The location (lat/lon) of the quadrat was specified using GNSS RTK. Then, plant species were recorded, and the height and abundance of each species were measured. Finally, the aboveground biomass of each species in the quadrat was measured by clipping the plant samples at the soil-surface level, the samples were weighed after Kill-green at 105 °C for 0.5 h and oven-drying at 80 °C for 8 h.

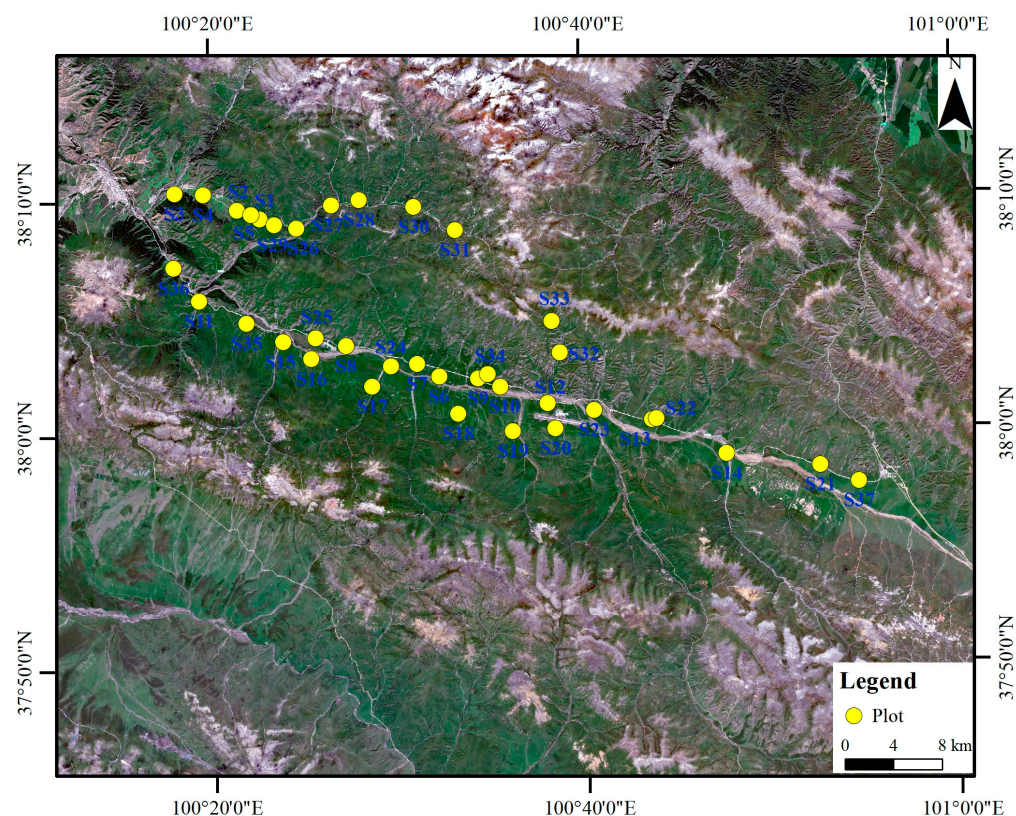


Figure 3. Sampling plots distribution in the study area.

Table 2. Sampling plot and quadrat list.

Sampling Plot/Quadrat	Sampling Plot/Quadrat	Sampling Plot/Quadrat
S1/1, 2, 3	S14/37, 38	S27/67, 68
S2/4, 5, 6	S15/39, 40	S28/69, 70, 71
S3/7, 8, 9	S16/41, 42	S29/72, 73, 74
S4/10, 11, 12	S17/43, 44	S30/75, 76
S5/13, 14, 15	S18/45, 46	S31/77, 78, 79
S6/16, 17, 18	S19/47, 48, 49	S32/80, 81
S7/19, 20, 21	S20/50, 51, 52	S33/82, 83
S8/22, 23, 24	S21/53, 54	S34/84, 85, 86
S9/25, 26, 27	S22/55, 56, 57	S35/87, 88, 89
S10/28, 29, 30	S23/58, 59	S36/90, 91
S11/31, 32	S24/60, 61	S37/92, 93
S12/33, 34	S25/62, 63	
S13/35, 36	S26/64, 65, 66	

Field tested points were collected in the study area, located in the center of large and homogeneous sites. There were 160 test sites for *S. chamaejasme* and 166 test sites for non-*S. chamaejasme* (native grassland, residential area, farmland, road, and river), for a total of 326 obtained ground test points. For the sites, information of vegetation type and coverage were collected, and their locations recorded as the above.

2.3. Methods

2.3.1. Two-Way Indicator Species Analysis (TWINSpan)

TWINSpan is a hierarchical clustering method with a “top-down” split strategy, modified from indicator species analysis [31]. The approach adopts discriminant function analysis and weighting algorithms to divide species dataset into small subsets, based on the indicator species that contributes the most variation [24]. By dichotomy, the division is repeated until the specified criteria are reached. TWINSpan divides the entire dataset by stepwise refinement, constructing an ordered two-way table to group the objects (quadrats) and variables (species), and finally obtains a classification of quadrats in hierarchical structure and forms a dendrogram of classes. In the study, taking the importance value of the species as the metric, all quadrats were grouped into the plant communities by TWINSpan analyses in PC-ORD 5.0 software. The importance value (IV) is a parameter used to represent the relative importance of plant species in a community, and the calculation formula is as follows [32]:

$$IV = (R_C + R_H + R_{AB})/3 \quad (1)$$

where R_C is the relative coverage, referring to the ratio of the coverage of each species to the sum of the coverage of all species in the quadrat; R_H and R_{AB} are relative height and relative aboveground biomass, respectively, and have the same meaning as R_C .

2.3.2. Spectral Index Conduction

During the flowering phase, *S. chamaejasme* forms dense white-pink inflorescences of capitula. Such distinct features make the toxic weed distinguishable from the surrounding green vegetation. For the sake of enhancing the spectral difference among *S. chamaejasme* communities and other communities, spectral indices were conducted on the principle of difference index (DI), ratio index (RI), multiplicative index (MI), and normalized difference index (NDI). For each kind of the 4 indexes, 10 combinations can be derived by pairwise calculation using the 5 bands of Planet imagery. In total, 40 combinations were obtained. The indices are as follows:

$$DI_{mn} = \rho_m - \rho_n \quad (2)$$

$$RI_{mn} = \rho_m / \rho_n \quad (3)$$

$$MI_{mn} = \rho_m \times \rho_n \quad (4)$$

$$NDI_{mn} = (\rho_m + \rho_n) / (\rho_m - \rho_n) \quad (5)$$

where ρ_m and ρ_n are the 5 bands of Planet imagery (listed in Table 1), m, n are band number, and $m \neq n$.

2.3.3. Spectral Separability Analysis

Transformed divergence (TD) was selected as a spectral separability criterion, the method offers advantages of simple rules and strong robustness [33]. TD algorithm calculates the separability of two classes based on their spectral signatures, which is as follows [34]:

$$D_{ij} = \frac{1}{2} \text{tr} \left[\left(\sum_i - \sum_j \right) \left(\sum_i^{-1} - \sum_j^{-1} \right) \right] + \frac{1}{2} \text{tr} \left[\left(\sum_i - \sum_j \right) (U_i - U_j) (U_i - U_j)^T \right] \quad (6)$$

$$TD_{ij} = 2 \left[1 - \exp \left(-\frac{D_{ij}}{8} \right) \right] \quad (7)$$

where TD_{ij} is the degree of separation between class i and class j ; $\sum_i$ and $\sum_j$ are covariance matrices of class i and j , respectively, and U_i and U_j are mean vectors of class i and j , respectively. The range of TD value is 0–2. TD values approaching to 2 indicate high degree of separability between classes. As TD value decreases, the degree of separability between classes decreases. TD values approaching 0 indicate low degree of separability between classes. Spectral separability analysis was performed through Python script programming.

Spectral separability measurement was used in two aspects. First, TD values of all pairs of the 93 quadrats were determined by using each index, according to Equations (2)–(5), then the mean of the TD values (hereinafter referred to as TD Mean) was calculated, which characterized the capability of each index in spectrally discriminating the quadrats. The indices with high values of TD Mean were chosen as the sensitive indices. Furthermore, the same processing was performed on the integration of original five bands of Planet image and the sensitive indices, and the optimal feature combination were selected for the remote sensing classification of plant communities. Second, spectral confusions among the TWINSpan communities were examined, and TD threshold of spectral separability was determined by measuring the TD value of the quadrats between and within the communities. Then, the quadrats of TWINSpan communities were adjusted and a spectrally separated scheme of plant communities were derived.

2.3.4. Image Classification and Assessment Metrics

Based on the ASTER GDEM (<https://earthdata.nasa.gov>, accessed on 11 October 2021) and Vegetation map of the People’s Republic of China (1:1,000,000) [35], the snow/ice, bare land with the altitude above 4000 m, and mountain forest/shrub were hierarchically excluded from the image. Then, using the TWINSpan-adjusted groups and 40 samples of residential area, farmland, road, and river chosen from the 166 non-*S. chamaejasme* sites, classification was performed based on the optimal feature combination, applying four machine learning algorithms. By grouping the resulted categories into the *S. chamaejasme* area and the non-*S. chamaejasme* area, the infestation of *S. chamaejasme* communities was identified.

The four algorithms are: random forest (RF), support vector machines (SVM), naive Bayes classifier (NBC), and k-nearest neighbor (KNN). Cross-validation was used for model training in order to obtain robust estimates of classification accuracy. Among the 93 quadrats, 70% of the samples were used as the training set and 30% as the test set. The algorithms were implemented using Python language based on PyCharm platform. Detailed parameters for each classification model are as follows:

Random forest: the ensemble package in the sklearn library is used, the number of decision trees is 10, the maximum depth is 50, and the minimum leaf size is 4.

Support vector machines: the radial basis function (RBF) is selected as the kernel function, with a penalty coefficient of 10 and a gamma of 0.1.

Naive Bayes classifier: The Gaussian Naive Bayes classifier in the sklearn library was adopted.

K-nearest neighbor: the Euclidean distance is used as the distance parameter, and the *n_neighbors* parameter is set to 5.

To accurately assess accuracy, the overall accuracy (OA), user’s accuracy (UA), producer’s accuracy (PA), and Kappa coefficient, in combination with Quantity disagreement (QD) and Allocation disagreement (AD), were adopted as assessment metrics. Quantity disagreement is defined as the amount of difference between reference map and predicted map, which is caused by the imperfect matching in the proportion of the classes. Allocation disagreement is defined as the amount of difference between reference map and predicted map, which is caused by the imperfect matching in the spatial allocation of the classes [36]. A lower value of quantity disagreement and allocation disagreement indicates good classification accuracy. The two metrics were calculated as follows:

$$P_{ij} = \left(\frac{n_{ij}}{\sum_{j=1}^J n_{ij}} \right) \left(\frac{N_i}{\sum_{i=1}^I N_i} \right) \quad (8)$$

$$QD = \frac{1}{2} \sum_{g=1}^J q_g = \frac{1}{2} \sum_{g=1}^J \left| \left(\sum_{i=1}^I P_{ig} \right) - \left(\sum_{j=1}^J P_{gj} \right) \right| \quad (9)$$

$$AD = \frac{1}{2} \sum_{g=1}^J a_g = \frac{1}{2} \sum_{g=1}^J \times 2 \min \left[\left(\sum_{i=1}^J P_{ig} \right) - p_{gg}, \left(\sum_{j=1}^J P_{gj} \right) - p_{gg} \right] \quad (10)$$

where P_{ij} is the estimated proportion of the study area; i, j represents the class i in the predicted map and the class j in the reference map, respectively. J is the number of classes. N_i is the number of samples of class i in the predicted map; n_{ij} is the number of samples classified as i and referenced as j . g is an arbitrary class, q_g is the quantitative disagreement of an arbitrary class g , and QD is the total quantity disagreement of all classes; a_g is the allocation disagreement of an arbitrary class g , and AD is the total allocation disagreement of all classes.

3. Results

3.1. Ecological Classification by TWINSpan

The 93 quadrats collected in the study area were classified into six communities by TWINSpan, the dendrogram and final communities are shown in Figure 4. Among the six groups, *P. annua*–*K. pygmaea* community (Group I), *E. nutans* community (Group V), and *P. fruticosa* community (Group VI) are typical native vegetation types in alpine meadows. Three types of *S. chamaejasme* communities were separated by TWINSpan, including Group II, III, and IV (Table 3). As the invasion aggravated, the species dominance of *S. chamaejasme* communities changed significantly; *S. chamaejasme* replaced *P. annua* (Gramineae) as the major dominant species in Group II and III. The coverage of three *S. chamaejasme* communities exhibited a gradient increase (14.3%, 22.3%, 29.5%), which indicated gradually serious grassland degradation.

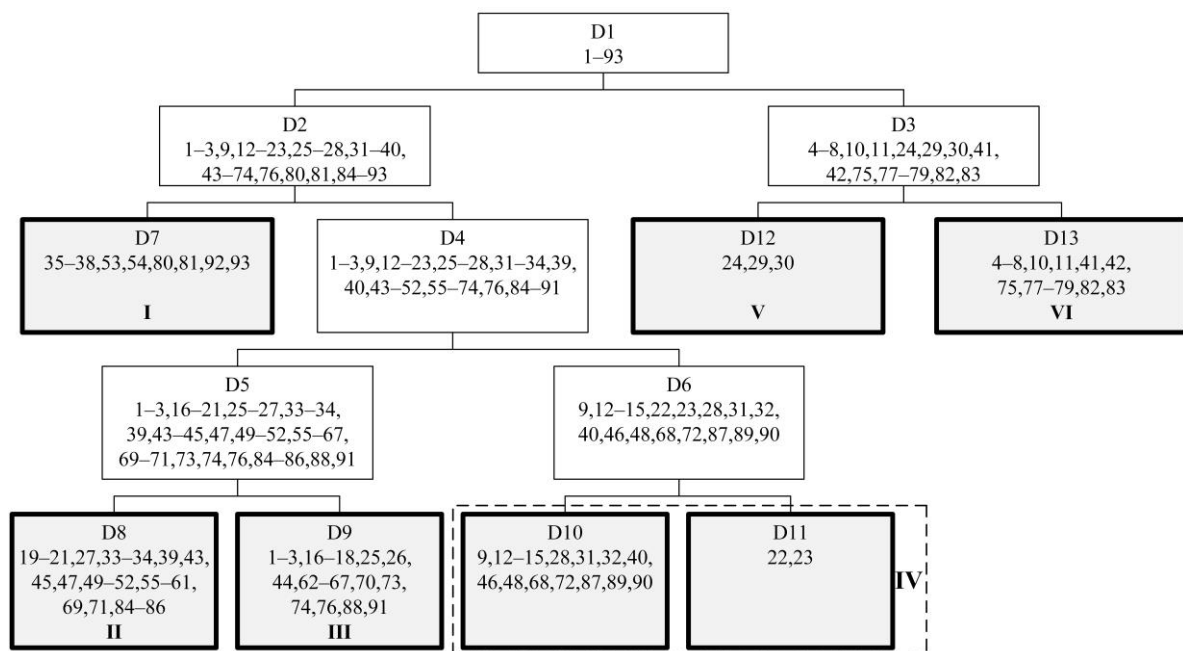


Figure 4. TWINSpan dendrogram.

Table 3. Ecological characteristics of TWINSpan classes.

Group	Description
Group I <i>P. annua</i> – <i>K. pygmaea</i> community	It is a typical native alpine meadow widely distributed on QTP, with the average community coverage of 41.7% and <i>P. annua</i> average coverage of 12.4%. Dominant species include: <i>P. annua</i> , <i>K. pygmaea</i> , and <i>T. lanceolata</i> . Accompanying species include: <i>G. straminea</i> , <i>Anaphalis lacteal</i> , and <i>Gentiana macrophylla</i> .

Table 3. Cont.

Group	Description
Group II <i>S. chamaejasme</i> – <i>P. annua</i> – <i>P. chinensis</i> community	It is a typical degraded meadow invaded by <i>S. chamaejasme</i> heavily. The average community coverage is 53.0%, and <i>S. chamaejasme</i> coverage is 29.5% with a range of 6.6–57.7%. Dominant species include: <i>S. chamaejasme</i> , <i>P. annua</i> , and <i>P. chinensis</i> . Accompanying species include: <i>M. ruthenica</i> , <i>A. rivularis</i> , and <i>T. mongolicum</i> .
Group III <i>S. chamaejasme</i> – <i>P. annua</i> – <i>G. straminea</i> community	It is a degraded meadow invaded by <i>S. chamaejasme</i> moderately. The average community coverage is 49.6%, and <i>S. chamaejasme</i> average coverage is 22.3% with a range of 6.7–46.7%. Dominant species include: <i>S. chamaejasme</i> , <i>P. annua</i> , and <i>G. straminea</i> . Accompanying species include: <i>A. rivularis</i> , <i>Morina kokonorica</i> , and <i>Bupleurum chinense</i> .
Group IV <i>P. annua</i> – <i>S. chamaejasme</i> community	It is an alpine meadow invaded by <i>S. chamaejasme</i> mildly. The average community coverage is 48.7%, and <i>S. chamaejasme</i> average coverage is 14.3% with a range of 4.0–46.1%. Dominant species include: <i>P. annua</i> , <i>S. chamaejasme</i> , and <i>A. rivularis</i> . Accompanying species include: <i>M. kokonorica</i> , <i>G. straminea</i> , and <i>T. mongolicum</i> .
Group V <i>E. nutans</i> – <i>M. ruthenica</i> community	It mainly contains native species and is distributed in flat valleys. The average community coverage is 54.8%, and <i>E. nutans</i> average coverage is 42.6%. Dominant species include <i>E. nutans</i> , <i>M. ruthenica</i> , and <i>T. mongolicum</i> . Accompanying species include <i>Ranunculus sceleratus</i> , <i>T. lanceolata</i> , and <i>Lancea tibetica</i> .
Group VI <i>P. fruticosa</i> – <i>P. annua</i> community	It is a typical native alpine meadow widely distributed on QTP. The average community coverage is 44.6%, and <i>P. fruticosa</i> average coverage is 20.9%. Dominant species include <i>P. fruticosa</i> , <i>P. asiatica</i> , and <i>Caragana jubata</i> . Accompanying species include <i>A. rivularis</i> , <i>P. chinensis</i> , and <i>M. ruthenica</i> .

3.2. Optimal Feature Selection

The spectral discrimination ability of each conducted index was evaluated by the measure of TD Mean. Among all indices, the TD Mean values of MI_{13} , NDI_{23} , and DI_{23} (0.377, 0.363, and 0.341, respectively) were much higher than those of the other indices (Figure 5); the three indices were sensitive to the spectral variation in the quadrats and reflected the highest spectral separability among the quadrats. Furthermore, TD Mean values were calculated based on the integrating of original five bands and three sensitive indices. Compared to the use of the original bands, the spectral discrimination ability was improved in different degree by the use of the combining strategies (Figure 6). The result confirmed the contribution of sensitive indices to the accuracy improvement of plant community identification. Among all feature combining strategies, 5 bands+ MI_{13} rated at the top of the list with the TD Mean of 0.386, and was selected as the source of information in the classification of plant communities.

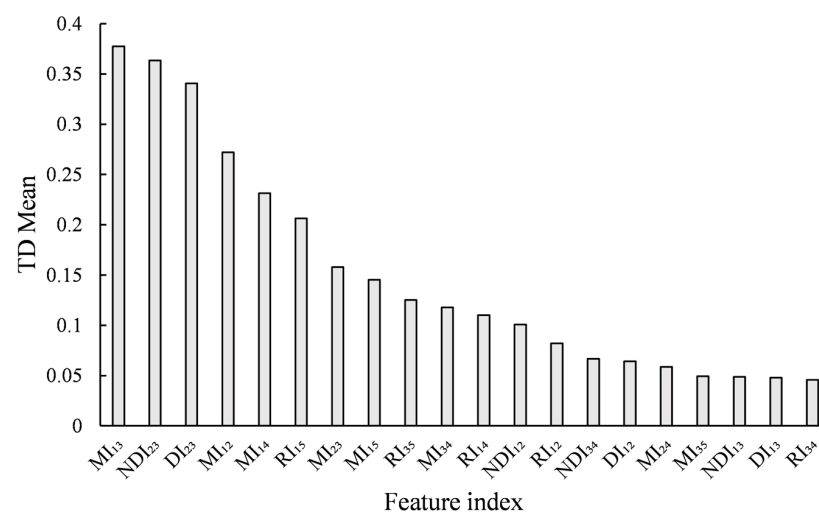


Figure 5. TD Mean values of the conducted spectral indices (the top 20 of 40 feature indices presented).

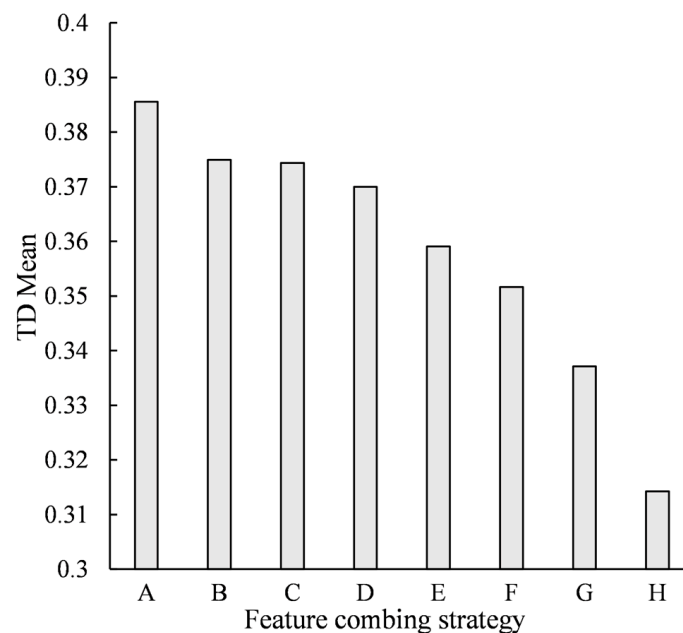


Figure 6. TD Mean values of the feature combining strategies. A: 5 bands+MI₁₃; B: 5 bands + MI₁₃ + NDI₂₃ + DI₂₃; C: 5 bands + MI₁₃ + DI₂₃; D: 5 bands+MI₁₃+NDI₂₃; E: 5 bands ++ NDI₂₃; F: 5 bands + DI₂₃; G: 5 bands + NDI₂₃ + DI₂₃; H: 5 bands.

3.3. Spectral Adjustment of TWINSpan Classes

The importance value is adopted in the TWINSpan vegetation classification in the study area, which is an integrative measure taking into account the coverage, height, and aboveground biomass of each species. However, the white-pink inflorescences make *S. chamaejasme* distinguishable from surrounding green vegetation in the flowering phase, *S. chamaejasme* coverage is the most important factor in spectral discriminating *S. chamaejasme* communities and the surrounding communities. TWINSpan classification indicated the spectral conflicts between and within the plant communities. *S. chamaejasme* quadrats and the native species quadrats were grouped together; *S. chamaejasme* quadrats with high-, moderate-, and low-coverage were involved simultaneously in each of the three *S. chamaejasme* communities. The result implied the unsuitability of original TWINSpan groups in further optimizing image classification. Therefore, TD is used to determine the spectral separability of the six TWINSpan communities and the quadrats are adjusted based on their spectral similarity, by using 5 bands+MVI₁₃ as data source. The resulted clusters were defined as follows:

- Four kinds of quadrats corresponding to the high-, moderate-, and low-coverage of *S. chamaejasme* (35–60%, 20–35%, and 0–20%, respectively), and green native species were picked out. TD values between each pair of the four groups were determined. Among them, the minimum value of 0.74 is taken as the TD threshold for the spectral adjustment of the TWINSpan communities;
- If the TD value between two quadrats is less than the threshold, they cannot be separated; if the TD value is greater than the threshold, the two can be separated;
- The TD values between each quadrat and all other quadrats within a class were less than the threshold;
- The TD values of the quadrats between different classes were greater than the threshold.

Based on the above guidelines, six TWINSpan communities were regrouped into four classes (Table 4). Group 1, 2, and 3 are the *S. chamaejasme* communities in high, moderate, and low coverage, respectively, approximately corresponding to TWINSpan group II, III, and IV; Class 4 is mainly composed of TWINSpan group I, V, and VI, which were green native communities and can hardly be discriminated using reflectance data, with the

maximum TD value of 0.23 between the communities. As a result, a classification scheme for the spectrally separating of ecological community was proposed.

Table 4. Ecological characteristics of TWINSpan-adjusted classes.

Group	Description
Group 1 High-coverage <i>S. chamaejasme</i> community	The group mainly corresponds to TWINSpan Group II, contains 16 quadrats from Group II, also contains 2 quadrats from Groups III and 3 quadrats from Group IV. It is characterized by densely clumped patches of <i>S. chamaejasme</i> over large area. <i>S. chamaejasme</i> average coverage is 41.1%, ranging from 35.1% to 57.7%.
Group 2 Medium-coverage <i>S. chamaejasme</i> community	The group mainly corresponds to TWINSpan Group III and II, contains 9 quadrats from TWINSpan Group III and 5 quadrats from Group II, also contains 2 quadrats from Group IV. <i>S. chamaejasme</i> and <i>P. annua</i> are the dominant species. <i>S. chamaejasme</i> was distributed in the mosaic of dense and sparse patches. Its average coverage is 26.4%, ranging from 20.8% to 34.6%.
Group 3 Low-coverage <i>S. chamaejasme</i> community	The group is a compound of TWINSpan Group IV, III and II, containing 6 quadrats from TWINSpan Group IV, 5 quadrats from Group III and 4 quadrats from Group II. <i>P. annua</i> and <i>A. rivularis</i> are dominant species, followed by <i>S. chamaejasme</i> . It was distributed in sparse patches, its average coverage is 13.5%, with a range of 6–19.4%.
Group 4 Native species communities	The group mainly corresponds to the TWINSpan Groups I, V, and VI, contains 40 quadrats of native species, of which 12 quadrats is from TWINSpan Groups II, III and IV (<i>S. chamaejasme</i> communities). The dominant species include <i>P. annua</i> , <i>P. fruticosa</i> , <i>E. nutans</i> , and <i>T. lanceolata</i> , with an average community coverage of 46.3%.

3.4. Image Classification-TWINSpan-Adjusted Classes

Classification was performed based on the TWINSpan-adjusted scheme and 40 samples of other classes, by applying four machine learning algorithms. As a result, Groups 1, 2, and 3 were combined into *S. chamaejasme* communities; Group 4 included native species communities; the residential area, farmland, road, river, and the excluded area were combined into non-grassland area, with the results were shown in Table 5 and Figure 7. *S. chamaejasme* communities accounted for 16.13–31.21% of the total grassland area based on different classification models. Four cases indicated high-coverage *S. chamaejasme* community accounted for the largest proportion of total *S. chamaejasme* area with the range of 39.95–50.96%, followed by low- and moderate-coverage *S. chamaejasme* communities. Among them, *S. chamaejasme* area estimated using KNN was obviously larger than the other models. Spatial distributions of *S. chamaejasme* derived by RF and NBC were fairly similar, which reveal *S. chamaejasme* to be mainly distributed in gentle slopes of hills and flat river valleys, characterized by clumped patches in random dispersion. Image stitching influenced the classification using SVM and KNN, resulting in the apparent inconsistency of *S. chamaejasme* distribution crossing the area.

Table 5. *S. chamaejasme* distribution area estimated using the four algorithms (1×10^4 ha).

Categories		RF	SVM	NBC	KNN
<i>S. chamaejasme</i> communities	High-coverage community	2.69	1.98	1.75	4.50
	medium-coverage community	1.53	1.28	1.15	1.94
	Low-coverage community	1.72	1.46	1.48	2.39
	Total	5.94	4.72	4.38	8.83
Native species communities		21.60	23.59	22.77	19.46
Non-grassland area		11.65	10.88	12.04	10.90

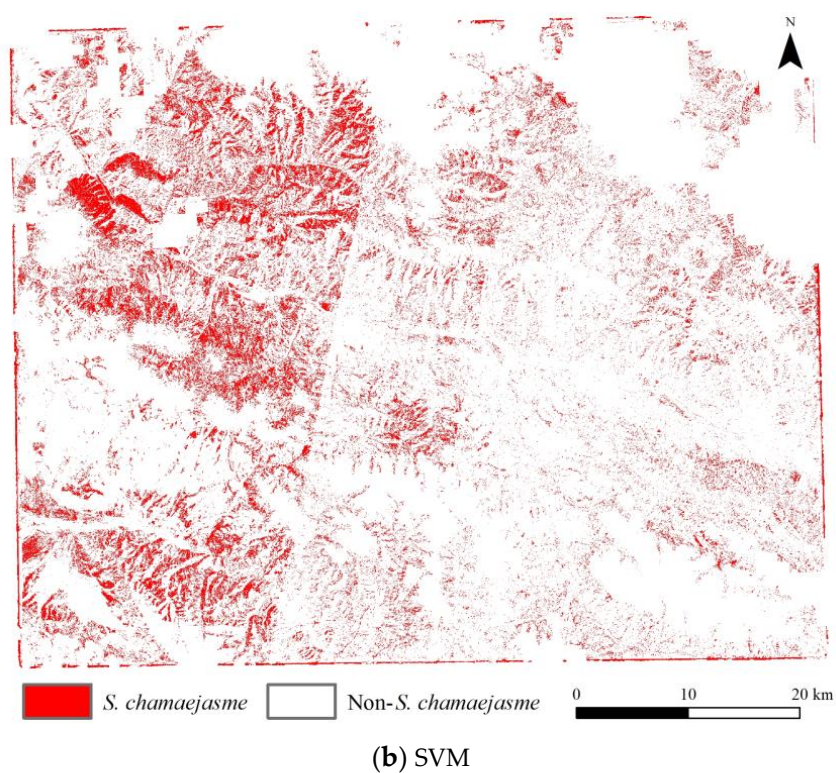
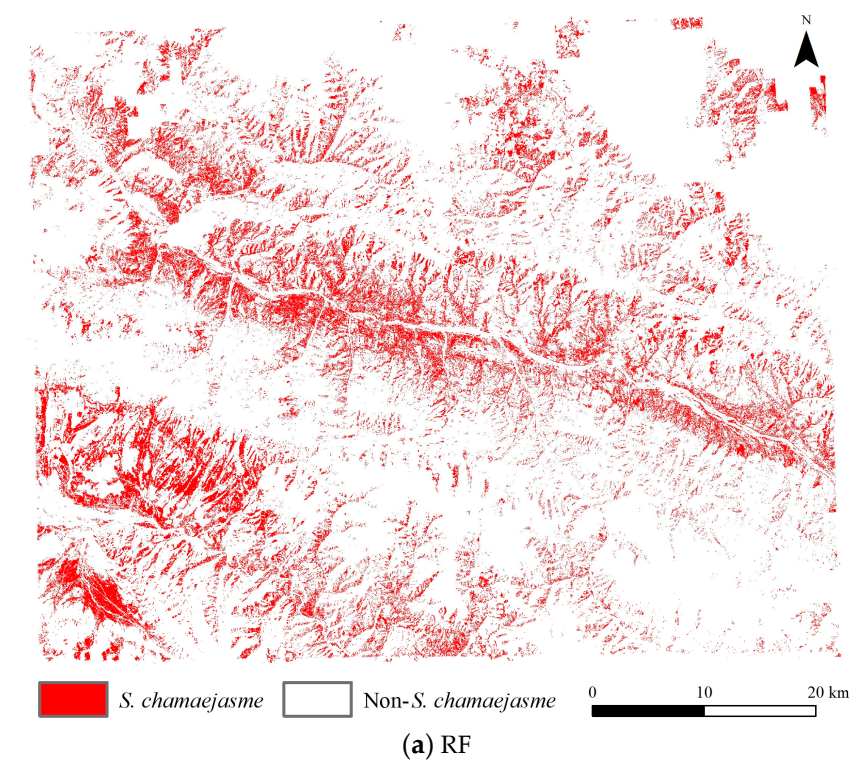


Figure 7. Cont.

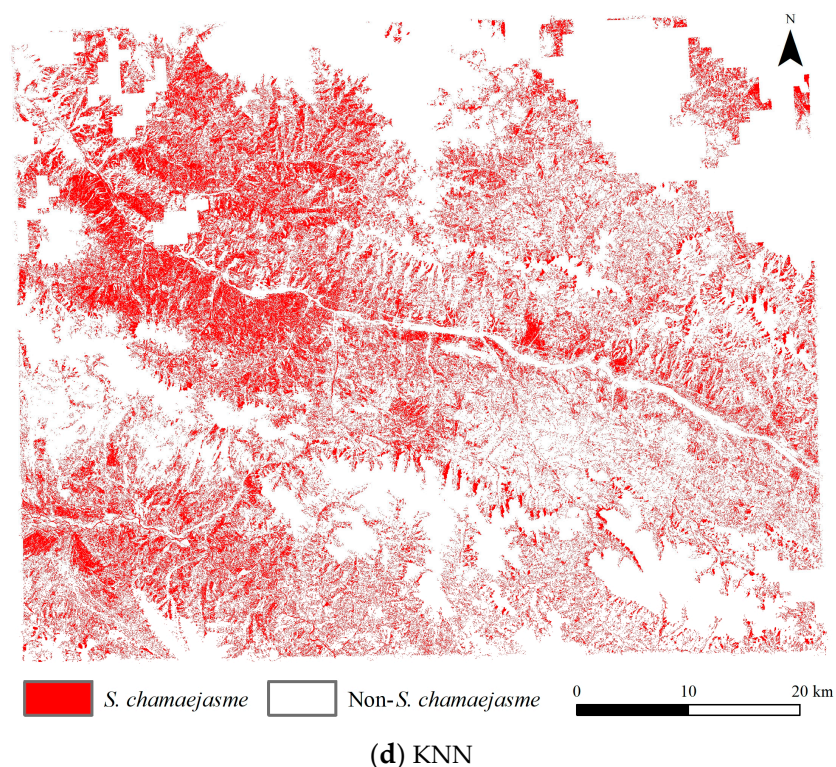
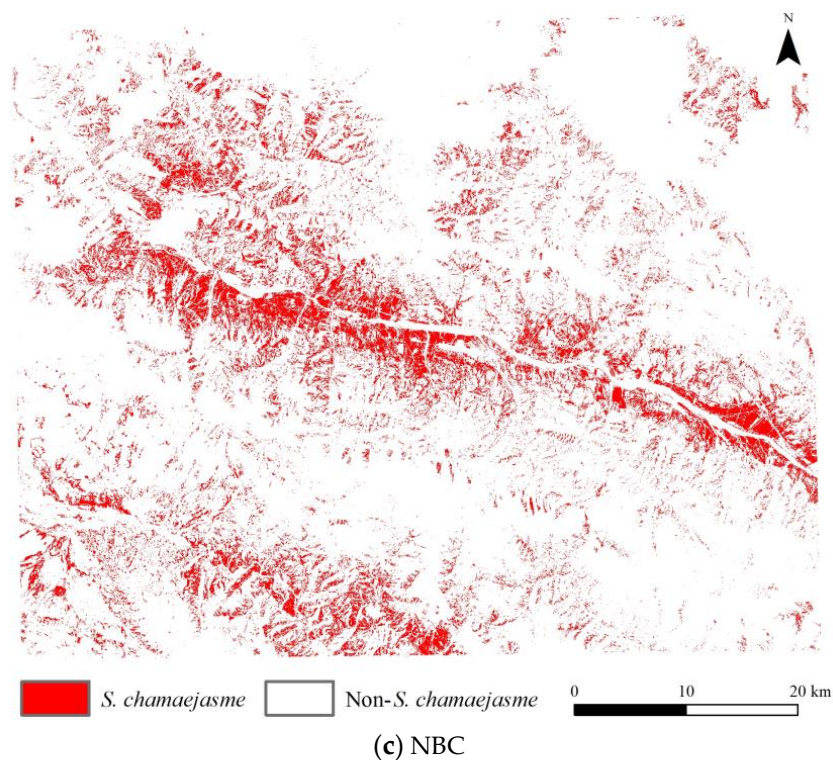


Figure 7. Distribution map of *S. chamaejasme* based on the four algorithms.

3.5. Accuracy Assessment

In the study, 160 *S. chamaejasme* sites and 126 non-*S. chamaejasme* sites were used in the accuracy assessment. The error matrix and accuracy comparison revealed considerable variation in classification accuracy of the four algorithms (Table 6 and Figure 8). The OA varied in the range of 70.6–87.8%, corresponding to the Kappa coefficient of 0.41–0.75. The disagreement percentage for the four algorithms was mainly due to AD (11.2–25.2%)

rather than QD (1.0–4.9%). RF obtained the best classification result with the OA, PA, and UA of more than 85%, and the Kappa of 0.75, it also achieved the lowest disagreement percentage with the QD and AD of 1.0% and 11.2%, respectively. The result based on NBC is worse than RF; the metrics of OA, Kappa, QD, and AD were 76.9%, 0.53, 4.9%, and 18.2%, respectively. SVM and KNN showed relatively poor classification with the Kappa of 0.41 and 0.43, respectively, producing obvious higher spatial allocation disagreement, with the AD of 25.3% and 23.8%, respectively. In comparison, RF exhibits good potential in *S. chamaejasme* identification.

Table 6. Error matrixes of *S. chamaejasme* classification based on the four algorithms.

Reference Class		<i>S. chamaejasme</i>	Non- <i>S. chamae- jasme</i>	UA	OA	Kappa (%)	QD	AD
RF	<i>S. chamaejasme</i>	141	16	89.8				
	Non- <i>S. chamaejasme</i>	19	110	85.3				
	PA (%)	88.1	87.3		87.8	75.2	1.0	11.2
SVM	<i>S. chamaejasme</i>	112	36	75.7				
	Non- <i>S. chamaejasme</i>	48	90	65.2				
	PA (%)	70.0	71.4		70.6	41.0	4.2	25.2
NBC	<i>S. chamaejasme</i>	134	40	77.0				
	Non- <i>S. chamaejasme</i>	26	86	76.8				
	PA (%)	83.8	68.3		76.9	52.6	4.9	18.2
KNN	<i>S. chamaejasme</i>	113	34	76.9				
	Non- <i>S. chamaejasme</i>	47	92	66.2				
	PA (%)	70.6	73.0		71.7	43.2	4.5	23.8

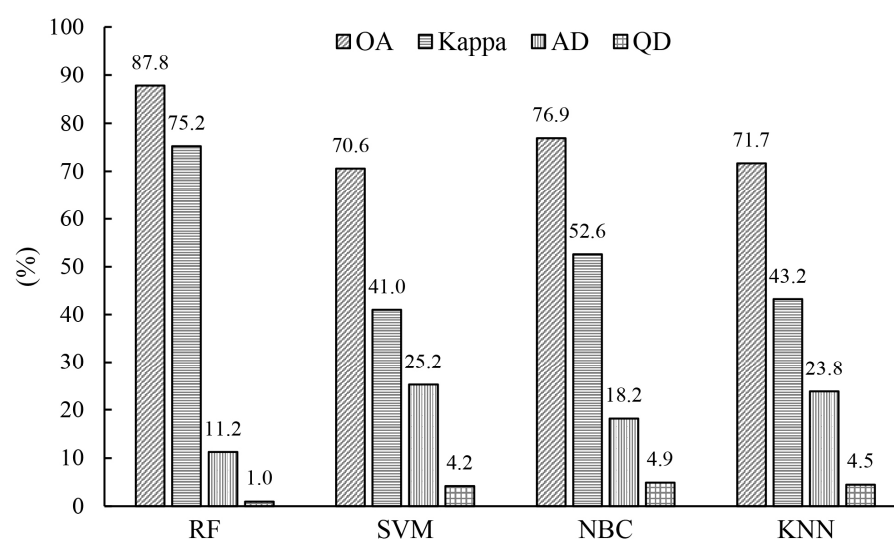


Figure 8. Accuracy comparison for the four algorithms.

4. Discussion

4.1. Spectral Confusion and Adjustment of Ecological Classes

Spectral responses of plant communities are influenced by factors of various kinds; ecologically distinct associations may have a similar spectral response, which accounts for

the difficulties of remote sensing mapping of vegetation [24]. Our study indicated there were obvious spectral conflicts between and within the TWINSpan ecological communities. In Group II (*S. chamaejasme* average coverage of 29.5%), 16 of the 26 quadrats were high-coverage *S. chamaejasme* quadrats, the spectral conflicts were relatively small. Group III and IV (*S. chamaejasme* average coverage of 22.3% and 14.3%) were the compounds of high-, moderate-, and low-coverage *S. chamaejasme* quadrats and native species quadrats, the spectral conflicts obviously increased. Groups I, V, and VI exhibited similar green spectral behaviors and could hardly be separated. The result revealed the different spectral confusions caused by the variation in coverage of invasive species, in agreement to the previously study [23].

Establishment of communities using the rule of spectral similarity provides a new way for the improvement of image classification. After the spectral adjustment of TWINSpan classification, the derived Groups 3, 2, and 1 were spectral separated in terms of consistency in *S. chamaejasme* coverage. Group 4 contained all native species quadrats, which eliminated the compounding of native species quadrats in *S. chamaejasme* communities and enhanced the spectral difference between *S. chamaejasme* communities and the native species communities. Meanwhile, the adjusted communities maintained their ecological significance. Coverage gradient of *S. chamaejasme* in Group 3, 2, and 1 markedly increased (13.5%, 26.4%, and 41.1%, respectively), and the three *S. chamaejasme* communities displayed clear and different patch patterns. Thus, the invasive process of *S. chamaejasme* was depicted more precisely. Group 4 represented the typical native plant species in the region, characterized by *P. annua*, *E. nutans*, and *P. fruticose*. Our study implied the adjustment of the ecological communities based on spectral criterion forms the *S. chamaejasme* identification scheme which are ecologically significant and spectrally delineated.

4.2. Spectral Separability between *S. chamaejasme* Communities and the Native Species Communities

There are more or less spectral differences between different plant species or communities, even if a number of vegetation types of similar structure are compared [37]. Spectral separability is commonly measured using the metrics of TD, BH (Bhattacharyya distance), and JM (Jeffries–Matusita distance) [38]. The greater the spectral distance between species or communities, the higher the separability there will be. Spectral difference are obvious between forest types or major physiognomic types (forest, shrub, and grassland); JM distance > 1.9 was used as the threshold for the spectral separability of boreal forest types in northern Canada [24]. Spectral distances between grassland types significantly decreased due to higher similarity in community structure; JM Distance > 1.4 was used as the threshold to spectrally delineate the eight grassland species in Africa, by using leaf spectra measured in the laboratory [39]. Spectral differences among 27 saltmarsh vegetation types were examined based on canopy spectra of densely homogeneous plots; the lowest JM Distances were 0.81–0.96 [37]. Spectral separability among the alpine grassland types on the QTP were relatively lower, the pair-wise JM distances between the different communities ranged in 0.44–0.93 [40]. The results revealed that the spectral distances of <1 were observed among grassland species due to the effects of soil background and coverage variation.

Spectral separability was evaluated between the four TWINSpan-adjusted communities using the TD measurement. With the decrease in *S. chamaejasme* coverage, the white-pink color weakened, and the spectral differences between *S. chamaejasme* communities and the native species community attenuated; TD > 0.74 was used as the threshold for distinguishing the four communities. Studies have indicated that there is a nonlinear relationship between spectral separability and classification accuracy, an increasing spectral distance does not necessarily indicate successful discrimination of classes [37,38]. Therefore, the small spectral distance does not mean poor classification result; it is still possible to obtain the classification accuracy of 70–90% with the TD value of large than 0.70 [24,37].

In the study, the definite separability threshold was determined between the four plant communities, which provides the basis for *S. chamaejasme* discrimination.

4.3. Potential of Planet Image and Machine Learning Algorithms in *S. chamaejasme* Discrimination

Planet imagery combined with machine learning algorithms obtained moderate or good accuracy for *S. chamaejasme* identification. The combination of five bands+MVI₁₃ showed the highest spectral separability between the quadrats with the TD Mean of 0.39, compared to the original five bands with the TD Mean of 0.31. Due to the subtle spectral difference across all quadrats, the improvement of TD Mean implies that the sensitive index spectrally characterizing *S. chamaejasme* can remarkably enhance the discrimination ability of *S. chamaejasme* from surrounding species. The result also confirms the potential of high-resolution multispectral image in the detection of toxic weed invasion, which is consistent with previous research [15,22,23]. Classification accuracy of the four algorithms varied greatly, due to the relatively small spectral distance between *S. chamaejasme* communities and the native species communities. High AD percentage is an important factor in affecting the classification result of SVM, KNN, and NBC. Meanwhile, SVM and KNN have low compatibility with the source of information, resulting in the varying degrees of mosaic effects [41,42]. RF achieved the lowest disagreement percentage and highest Kappa, owing to the strong anti-interference and stable performance of this method [43,44]. In view of the OA, Kappa, AD, and QD, the RF model outperformed SVM, NBC, and KNN.

5. Conclusions

In this study, an integrative approach of the ecological classification and image classification were proposed for *S. chamaejasme* discrimination over a large scale. The feature of white-pink inflorescences in the flowering phase makes *S. chamaejasme* distinguishable from the surrounding species. Determination of the definite spectral separability provides a basis for adjusting TWINSpan ecological classes into spectrally separated *S. chamaejasme* communities and the native species communities, which are amenable to image classification. The study confirms the potential of high-resolution Planet imagery combined with machine learning algorithms in the monitoring of toxic weed invasion, which was proved to obtain good accuracy for *S. chamaejasme* discrimination in alpine grasslands. Environmental variables have important influences on the distribution of plant species and communities, and the optimization of *S. chamaejasme* discrimination needs to be further carried out by involving ecological ordination techniques and environmental data.

Author Contributions: Conceptualization and methodology, Y.L. (Yongmei Liu) and N.H.; software and formal analysis, N.H., X.G. and X.D.; investigation, Y.L. (Yongmei Liu), N.H., X.D., H.W. and X.G.; resources, Y.L. (Yongmei Liu); writing—original draft preparation, N.H. and Y.L. (Yongmei Liu); writing—review and editing, Y.L. (Yongmei Liu) and N.H.; visualization, N.H.; supervision, L.W. and Y.L. (Yongqing Long); funding acquisition, Y.L. (Yongmei Liu). All authors have read and agreed to the published version of the manuscript.

Funding: The study was supported by the National Natural Science Foundation of China (Grant No. 41871335).

Data Availability Statement: The data that support the findings of this study are available on request from the corresponding author.

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

References

1. Dai, L.; Fu, R.; Guo, X.; Ke, X.; Du, Y.; Zhang, F.; Cao, G. Effect of grazing management strategies on alpine grassland on the northeastern Qinghai-Tibet Plateau. *Sci. Total Environ.* **2021**, *173*, 106418. [\[CrossRef\]](#)
2. Zhang, J.; Fang, S.; Liu, H. Estimation of alpine grassland above-ground biomass and its response to climate on the Qinghai-Tibet Plateau during 2001 to 2019. *Glob. Ecol. Conserv.* **2022**, *35*, e02065. [\[CrossRef\]](#)
3. Harris, R.B. Rangeland degradation on the Qinghai-Tibetan Plateau: A review of the evidence of its magnitude and causes. *J. Arid. Environ.* **2010**, *74*, 1–12. [\[CrossRef\]](#)

4. Wen, L.; Dong, S.; Li, Y. The effects of biotic and abiotic factors on the spatial heterogeneity of alpine grassland vegetation at a small scale on the Qinghai-Tibet Plateau, China. *Environ. Monit. Assess.* **2013**, *10*, 8051–8064. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Zhao, M.; Gao, X.; Wang, J.; He, X.; Han, B. A review of the most economically important poisonous plants to the livestock industry on temperate grasslands of China. *J. Appl. Toxicol.* **2013**, *33*, 9–17. [\[CrossRef\]](#)
6. Sun, G.; Luo, P.; Qiu, P.; Gao, Y.; Chen, H.; Shi, F.S. *S. chamaejasme* L. increases soil N availability, turnover rates and microbial biomass in an alpine meadow ecosystem on the eastern Tibetan Plateau of China. *Soil Biol. Biochem.* **2009**, *41*, 86–91. [\[CrossRef\]](#)
7. Shi, Z.C. *The Poisonous Plants of Chinese Grassland*; China Agriculture Press: Beijing, China, 1997; pp. 140–150.
8. Zhang, Z.; Zhang, Y.; Sun, H. The reproductive biology of *Stellera chamaejasme* (Thymelaeaceae): A self-incompatible weed with specialized flowers. *Flora-Morphol. Distrib. Funct. Ecol. Plants* **2011**, *206*, 567–574. [\[CrossRef\]](#)
9. 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**, *112*, 237–245. [\[CrossRef\]](#)
10. Taylor, S.; Kumar, L.; Reid, N.; Lewis, C.R. Optimal band selection from hyperspectral data for *Lantana camara* discrimination. *Int. J. Remote Sens.* **2012**, *33*, 5418–5437. [\[CrossRef\]](#)
11. 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. [\[CrossRef\]](#)
12. Rai, P.K.; Singh, J. Invasive alien plant species: Their impact on environment, ecosystem services and human health. *Ecol. Indic.* **2020**, *111*, 106020.
13. 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. [\[CrossRef\]](#)
14. Dorigo, W.; Lucieer, A.; Podobnikar, T.; Čarni, A. Mapping invasive *Fallopia japonica* by combined spectral, spatial, and temporal analysis of digital orthophotos. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *19*, 185–195. [\[CrossRef\]](#)
15. Kganyago, M.; Odindi, J.; Adjorlolo, C.; Mhangara, P. Evaluating the capability of Landsat 8 OLI and SPOT 6 for discriminating invasive alien species in the African Savanna landscape. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *67*, 10–19. [\[CrossRef\]](#)
16. Ji, W.J.; Wang, L.; Knutson, A.E. Detection of the spatiotemporal patterns of beetle-induced tamarisk (*Tamarix* spp.) defoliation along the Lower Rio Grande using Landsat TM images. *Remote Sens. Environ.* **2017**, *193*, 76–85. [\[CrossRef\]](#)
17. Mirik, M.; Ansley, R.J.; Steddom, K.; Jones, D.C.; Rush, C.M., Jr.; 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. [\[CrossRef\]](#)
18. Noujdina, N.V.; Ustin, S.L. Mapping Downy Brome (*Bromus tectorum*) Using Multidate AVIRIS Data. *Weed Sci.* **2008**, *56*, 173–179. [\[CrossRef\]](#)
19. Dehaan, R.; Louis, J.; Wilson, A.; Hall, A.; Rumbachs, R. Discrimination of blackberry (*Rubus fruticosus* sp. agg.) using hyperspectral imagery in Kosciuszko National Park, NSW, Australia. *J. Photogramm. Remote Sens.* **2007**, *62*, 13–24. [\[CrossRef\]](#)
20. Gallego-Fernández, J.B.; Martínez, M.L.; García-Franco, J.G.; Zunzunegui, M. The impact on plant communities of an invasive alien herb, *Oenothera drummondii*, varies along the beach-coastal dune gradient. *Flora* **2019**, *260*, 151466. [\[CrossRef\]](#)
21. Qian, W.; Huang, Y.; Liu, Q.; Fan, W.; Sun, Z.; Dong, H.; Wan, F.; Qiao, X. UAV and a deep convolutional neural network for monitoring invasive alien plants in the wild. *Comput. Electron. Agric.* **2020**, *174*, 105519. [\[CrossRef\]](#)
22. Martín, M.; Barreto, L.; Fernández-Quintanilla, C. Discrimination of sterile oat (*Avena sterilis*) in winter barley (*Hordeum vulgare*) using QuickBird satellite images. *Crop Prot.* **2011**, *30*, 1363–1369. [\[CrossRef\]](#)
23. Casady, G.M.; Hanley, R.S.; Seelan, S.K. Detection of Leafy Spurge (*Euphorbia esula*) Using Multidate High-Resolution Satellite Imagery. *Weed Technol.* **2005**, *19*, 462–467. [\[CrossRef\]](#)
24. Thomas, V.; Treitz, P.; Jelinski, D.; Miller, J.; Lafleur, P.; McCaughey, J.H. Image classification of a northern peatland complex using spectral and plant community data. *Remote Sens. Environ.* **2003**, *84*, 83–99. [\[CrossRef\]](#)
25. Brook, R.; Kenkel, N. A multivariate approach to vegetation mapping of Manitoba's Hudson Bay Lowlands. *Int. J. Remote Sens.* **2002**, *23*, 4761–4776. [\[CrossRef\]](#)
26. Malik, R.N.; Husain, S.Z. Linking remote sensing and ecological vegetation communities: A multivariate approach. *Pak. J. Bot.* **2008**, *40*, 337–349.
27. Schmidtlein, S.; Zimmermann, P.; Schüpferling, R.; Weiß, C. Mapping the floristic continuum: Ordination space position estimated from imaging spectroscopy. *J. Veg. Sci.* **2007**, *18*, 131–140. [\[CrossRef\]](#)
28. Li, J.; Liu, Y.; Mo, C.; Wang, L.; Pang, G.; Cao, M. IKONOS image-based extraction of the distribution area of *Stellera chamaejasme* L. in Qilian County of Qinghai Province, China. *Remote Sens.* **2016**, *8*, 148. [\[CrossRef\]](#)
29. Bao, G.; Wang, Y.; Song, M.; Wang, H.; Yin, Y.; Liu, S.; Yang, Y.; Yang, M. Effects of *Stellera chamaejasme* patches on the surrounding grassland community and on soil physical-chemical properties in degraded grasslands susceptible to *S. chamaejasme* invasion. *Acta Prataculturae Sin.* **2019**, *28*, 51–61.
30. Zhao, B.; Liu, Z.; Lu, H. Damage and control of poisonous weeds in western grassland of China. *Agric. Sci. China* **2010**, *10*, 1512–1521. [\[CrossRef\]](#)
31. Hill, M.O. *TWINSPAN-A FORTRAN Program for Arranging Multivariate Data in an Ordered Two-way Table by Classification of the Individuals and Attributes*; Cornell University Press: New York, NY, USA, 1979.
32. Zhang, J.T. *Quantitative Ecology*; Science Press: Beijing, China, 2004; pp. 50–60.

33. Landgrebe, D.A. *Signal Theory Methods in Multispectral Remote Sensing*; John Wiley & Sons: New York, NY, USA, 2003; Volume 24.
34. Mausel, P.; Wu, Y.; Li, Y.; Moran, E.F.; Brondizio, E.S. Spectral identification of successional stages following deforestation in the Amazon. *Geocarto Int.* **1993**, *8*, 61–71. [[CrossRef](#)]
35. Zhang, X.S. *Vegetation Map of the People's Republic of China (1:1,000,000,000)*; Geology Press: Beijing, China, 2007; pp. 60–70. Normative references: Vegetation Map of the People's Republic of China (1:1000000), 2021, Plant Data Center of Chinese Academy of Sciences, CSTR:34735.11.PLANTDATA.0155. [[CrossRef](#)]
36. Pontius, R.G., Jr.; Millones, M. Death to Kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment. *Int. J. Remote Sens.* **2011**, *32*, 4407–4429. [[CrossRef](#)]
37. Schmidt, K.; Skidmore, A. Spectral discrimination of vegetation types in a coastal wetland. *Remote Sens. Environ.* **2003**, *85*, 92–108. [[CrossRef](#)]
38. Swain, P.; King, R. Two effective feature selection criteria for multispectral remote sensing. *LARS Tech. Rep.* **1973**, *39*, 1–6.
39. Schmidt, K.; Skidmore, A. Exploring spectral discrimination of grass species in African rangelands. *Int. J. Remote Sens.* **2001**, *22*, 3421–3434. [[CrossRef](#)]
40. Zhang, X.; Nan, Z.; Sheng, Y.; Zhao, L.; Zhou, G.; Yue, G.; Wu, J. Analysis of time-series Modis 250M vegetation index data for vegetation classification in the Wenquan area over the Qinghai-Tibet Plateau. In Proceedings of the 2010 IEEE International Geoscience and Remote Sensing Symposium, Honolulu, HI, USA, 25–30 July 2010.
41. Chen, B.; Fan, Y.; Lan, W.; Liu, J.; Cao, C.; Gao, Y. Fuzzy support vector machine with graph for classifying imbalanced datasets. *Neurocomputing* **2022**, *514*, 296–312. [[CrossRef](#)]
42. Wang, Y.; Pan, Z.; Dong, J. A new two-layer nearest neighbor selection method for kNN classifier. *Knowl-Based. Syst.* **2022**, *235*, 107604. [[CrossRef](#)]
43. Triscowati, D.W.; Sartono, B.; Kurnia, A.; Dirgahayu, D.; Wijayanto, A.W. Classification of rice-plant growth phase using supervised random forest method based on Landsat-8 multitemporal data. *Int. J. Remote Sens.* **2020**, *16*, 187–196. [[CrossRef](#)]
44. Belgiu, M.; Drăguț, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm Remote Sens.* **2016**, *114*, 24–31. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.