remote sensing

MDPI

Article

Comparing Machine Learning Algorithms for Pixel/Object-Based Classifications of Semi-Arid Grassland in Northern China Using Multisource Medium Resolution Imageries

Nitu Wu [1], Luís Guilherme Teixeira Crusiol [2], Guixiang Liu [3], Deji Wuyun [4] and Guodong Han [1,*]

[1] Key Laboratory of Grassland Resources of the Ministry of Education, College of Grassland, Resources and Environment, Inner Mongolia Agricultural University, Hohhot 010011, China
[2] Embrapa Soja (National Soybean Research Center—Brazilian Agricultural Research Corporation), Londrina 86001-970, Brazil
[3] Institute of Grassland Research, Chinese Academy of Agricultural Sciences, Hohhot 010010, China
[4] Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China
* Correspondence: hanguodong@imau.edu.cn

Abstract: Knowledge of grassland classification in a timely and accurate manner is essential for grassland resource management and utilization. Although remote sensing imagery analysis technology is widely applied for land cover classification, few studies have systematically compared the performance of commonly used methods on semi-arid native grasslands in northern China. This renders the grassland classification work in this region devoid of applicable technical references. In this study, the central Xilingol (China) was selected as the study area, and the performances of four widely used machine learning algorithms for mapping semi-arid grassland under pixel-based and object-based classification methods were compared: random forest (RF), support vector machine (SVM), k-nearest neighbor (KNN), and naive Bayes (NB). The features were composed of the Landsat OLI multispectral data, spectral indices, Sentinel SAR C bands, topographic, position (coordinates), geometric, and grey-level co-occurrence matrix (GLCM) texture variables. The findings demonstrated that (1) the object-based methods depicted a more realistic land cover distribution and had greater accuracy than the pixel-based methods; (2) in the pixel-based classification, RF performed the best, with OA and Kappa values of 96.32% and 0.95, respectively. In object-based classification, RF and SVM presented no statistically different predictions, with OA and Kappa exceeding 97.5% and 0.97, respectively, and both performed significantly better than other algorithms. (3) In pixel-based classification, multispectral bands, spectral indices, and geographic features significantly distinguished grassland, whereas, in object-based classification, multispectral bands, spectral indices, elevation, and position features were more prominent. Despite the fact that Sentinel 1 SAR variables were chosen as an effective variable in object-based classification, they made no significant contribution to the grassland distinction.

Keywords: remote sensing; grassland classification; machine learning; random forest; support vector machine

check for updates

Citation: Wu, N.; Crusiol, L.G.T.; Liu, G.; Wuyun, D.; Han, G. Comparing Machine Learning Algorithms for Pixel/Object-Based Classifications of Semi-Arid Grassland in Northern China Using Multisource Medium Resolution Imageries. *Remote Sens.* **2023**, *15*, 750. https://doi.org/10.3390/rs15030750

Academic Editors: Carlos Antonio Da Silva Junior and Luciano Shozo Shiratsuchi

Received: 5 December 2022
Revised: 18 January 2023
Accepted: 20 January 2023
Published: 28 January 2023

1. Introduction

Grassland occupies approximately one-third of the Earth's terrestrial surface [1] and serves an irreplaceable role in livestock farming [2] and ecosystem services such as water conservation [3], climate regulation [4], and biodiversity protection [5]. In China, nearly 80% of the grassland is in arid and semi-arid regions, which are extremely sensitive to external disturbances. About 90% of native grassland in China has been degraded to various extents, which impedes the sustainable growth of the regional economy [6,7]. Grassland classification is preparatory work for further understanding grassland resources [8,9]. Detailed and timely knowledge of grassland classification is desperately needed for sustainable planning and management of grassland resources.

Remote sensing is the most available technical method for carrying out land cover classification at the macro scale. Particularly, recent advances in spectroscopy have improved the technical tools for land monitoring based on remote sensing [10]. Numerous studies emphasized the value of classifying land cover using data from multiple sources, including optical, microwave, and other auxiliary data [9,11–13]. However, such attempts mainly focus on integrated land cover classification. For the existing grassland studies, they focus more on extraction, and few involve further class distinction. According to the application case of remote sensing in other land-use types, such as forest and cropland, mapping detailed grassland types is apparently feasible. However, remote sensing classification is a complex process, and a good classification depends on the coordination of multi-factors, including data sources, classifier algorithms, classification approaches, etc.

Optical remote sensing (ORS) is frequently the preferred data source for land observation. Currently, a wide range of ORS data, at various spatial, spectral, and temporal scale resolutions, from local to global scales, is available for image classification [9,10]. For large-scale land use retrieval, the Landsat and Sentinel series imageries are especially popular. However, the similarity of spectral features among herbages makes it challenging to differentiate mixed grassland types using only a simple spectral image. In response, some studies suggested using hyperspectral remote sensing to obtain information available for grassland differentiation to the greatest extent in the narrow vegetation-sensitive band [14–16]. However, the limited observation range and high costs greatly discount the applicability of hyperspectral remote sensing in open grasslands. Others recommended integrating the use of multiple variables such as optical, geometry, texture, and environmental factors to improve the separability between vegetation from various dimensions [8,17–19]. Recent studies have also highlighted the significance of using time series remote sensing data in grassland classification [20,21]. Apart from the single ORS-based applications, some studies confirmed that the fusion of optical remote sensing and synthetic aperture radar (SAR) could significantly improve vegetation differentiation [12,22,23]. However, uncertainties and challenges of SAR application in low-vegetation areas still exist from soil moisture and texture interference [24,25].

Selecting an appropriate classifier is crucial for image classification. The classifiers can be divided into parametric and non-parametric depending on whether or not an additional assumption for data distribution is required [11]. Due to their simple training and operation, parametric classifiers such as logistic regression and naive Bayes (NB) are advantageous for classification issues involving straightforward logical relationships. In contrast, non-parametric classifiers can fit intricate non-linear relationships, making them better suited for monitoring complex land cover. Among them, random forests (RF), k-nearest neighbor (KNN), and support vector machines (SVM) are the most widely used in land use classification [26–30]. Studies have shown that these algorithms are effective at classifying grasslands as well. For instance, Zhao et al. [20] confirmed that combining RF and PROBA-V remote sensing images has great potential for central African grassland classification. Rapinel et al. [21] found that the SVM outperformed other classifiers for mapping grassland plant communities using Sentinel-2 time series data. Dusseux et al. [31] pointed out that the distribution of grassland management practices can be mapped very accurately (Kappa = 0.82) at a field scale over large agricultural areas using KNN and a series of satellite images. However, it is worth noting that these algorithms perform better only under suitable variables and sample conditions.

In addition, the basic analysis unit is an important factor in image classification. It could be a pixel, a group of neighboring pixels, or the whole image. Generally, the classification from remotely sensed imagery can be divided into two image analysis approaches: pixel based and object based [32]. The pixel-based analysis technique is conducted on a per-pixel level, using only the spectral information available for that individual pixel. In contrast, object-based analysis is carried out on a cluster group of pixels, considering the spatial properties of each pixel as they relate to each other [33]. Numerous studies have compared the applicability of these two analysis approaches under different conditions. For medium- or high-resolution satellite imagery, most claimed that the object-based image analysis technique has more potential than that of the pixel based [34–38]. However, some claim that, in a particular situation, there is no statistically significant difference between these two approaches [32,39]. Currently, limited studies have systematically compared the performance of these two approaches with popular machine learning algorithms for the semi-arid grassland classification in northern China, leaving this field lacking in methodologies and technical references.

To bridge the knowledge gap mentioned above, this study focused on central Xilingol, a representative semi-arid grassland of northern China, and carried out the following studies: (1) compared the performance of both pixel-based and object-based classification approaches with a selection of frequently used machine learning algorithms: RF, SVM, KNN, and NB on semi-arid grassland classification; (2) examined the effectiveness of Landsat OLI multispectral bands, spectral indices, geographic factors, Sentinel SAR C bands, geometric factors, and texture features for mapping semi-arid grassland.

2. Materials and Methods

2.1. Study Area

Xilingol (43°26′~45°52′N, 115°18′~118°06′E) is located in the south-central Mongolian Plateau and is a core region of native grassland in China (Figure 1). The terrain of this area is higher in the south and lowers in the north, with an average altitude of 1000 m (Figure 1). The climate belongs to a mid-temperate semi-arid continental climate, with total annual precipitation of 300 mm and an average annual temperature of 3 °C. According to the China grassland classification system (CGCS) released in the 1980s, the typical steppe is the main grassland type, and a few meadow steppes are in the eastern region. The dominant species include *Stipa grandis* (*S. grandis*), *Stipa krylovii* (*S. krylovii*), *Leymus chinensis* (*L. chinensis*), *Artemisia frigida* (*A. frigida*), *Caragana microphylla* (*C. microphylla*), and *Achnatherum splendens* (*A. splendens*).

2.2. Classification System and Sample Data

The predominant land types of the study area include grassland, shrubland, cropland, mining area, building, and waterbody. According to the CGCS, the grassland was further classified into hilly meadow steppe, hilly steppe, plain steppe, sandy steppe, saline meadow, and marshy meadow (Table 1). Compared with other classification standards, CGCS focuses more on integrating natural and social attributes of grassland and defines grassland types as a unit with the same natural and economic characteristics at a particular time and space [40–42]. The grassland types involved in this study are defined as follows:

(1) Hilly meadow steppe occurs mainly in the high relief sites of 900~1500 m with relatively moist and fertile soils, and the dominant species include *Stipa baicalensis*, *Filifolium sibiricum*, and *Leymus chinensis*.
(2) Hilly steppe is mainly formed at elevations between 600 and 1300 m and is dominated by xerophytic or semi-xeric bunchgrass.
(3) Plain steppe is the most widely distributed grassland type in this area, and occurs under a semi-arid climate with annual precipitation around 350 mm. The most common communities are dominated by *Stipa grandis*, *Stipa krylovii*, *Leymus chinensis*, *Cleistogenes squarrosa*, and *Artemisia frigida*.
(4) Sandy steppe has distinctive zonal characteristics and is mainly found in Hunshandak sandland.
(5) Saline meadow occurs mainly on salinized depression sites, broad valleys, fringes of lake basins, and river flats within steppe and desert regions. It is primarily composed of mesic perennial halophytes, such as *Achnatherum splendens* and *Leymus chinensis*.
(6) Marshy meadow is primarily composed of hygrophilous herbs, such as *Phragmites australis*, and has transitional characteristics between a meadow and a marsh.

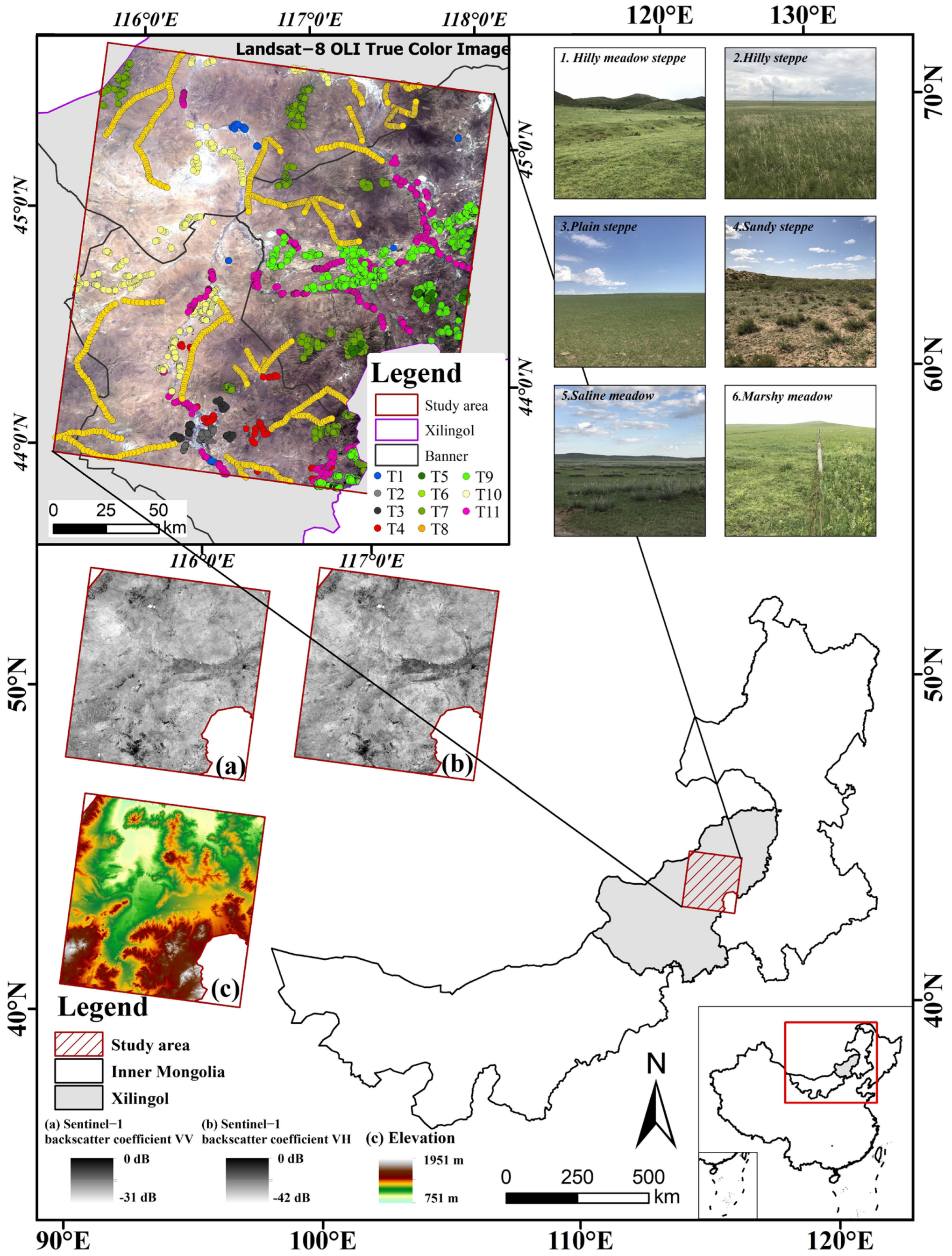

Figure 1. Location of the study area and distribution of the sample points.

A total of 3620 samples across the above-mentioned land covers were collected through field observation (taking photos and GPS positioning) combined with high-resolution imagery, Gaofen (2.5 m × 2.5 m), from July to August 2017 (Figure 1). After that, these samples were randomly divided into training and testing datasets in a ratio of 7:3. The distribution of samples for each land cover is shown in Table 1.

Table 1. Classification system and sample distributions.

Land Covers				No. of Samples
Non-vegetation		Waterbody	T1	47
		Building	T2	97
		Mining area	T3	50
Vegetation	Non-grassland	Cropland	T4	59
		Shrubland	T5	61
	Grassland (CGCS)	Hilly meadow steppe	T6	191
		Hilly steppe	T7	555
		Plain steppe	T8	826
		Sandy steppe	T9	870
		Saline meadow	T10	340
		Marshy meadow	T11	524

2.3. Remote Sensing Data and Preprocessing

2.3.1. Multispectral Imagery

Multispectral imagery (path/row: 124/29) of the Landsat-8 OLI/TIRS Level-1 product was downloaded from the Geospatial Data Cloud site, Chinese Academy of Sciences (http://www.gscloud.cn, accessed on 15 November 2022), and the imaging time was 17 July 2017. The Landsat is the longest-running Earth observation program, and Landsat 8 launched on 11 February 2013. Landsat-8 carries an optical sensor that provides a global coverage image every 16 days at a spatial resolution of 30 m (bands 1~7, 9) and 15 m (panchromatic). In this study, the visible light bands 2, 3, and 4, the near-infrared (NIR) band 5, and the short-wave infrared (SWIR) bands 6 and 7 were selected for training models. Moreover, three commonly used spectral indices, normalized difference vegetation index (NDVI) [43], normalized difference built-up index (NDBI) [44], and normalized difference water index (NDWI) [45], were calculated to enhance the separability between vegetation and non-vegetation. The formulas are as follows:

$$NDVI = \frac{NIR - R}{NIR + R} \tag{1}$$

$$NDBI = \frac{SWIR - NIR}{SWIR + NIR} \tag{2}$$

$$NDWI = \frac{G - NIR}{G + NIR} \tag{3}$$

where NIR, R, G, and $SWIR$ represent the reflectance of near-infrared, red, green, and short-wave infrared bands, respectively.

2.3.2. Synthetic Aperture Radar Data

SAR data of the Sentinel-1 SAR Ground Range Detected (GRD) product were downloaded from the Google Earth Engine (GEE) platform, and the imaging time was July 2017. The Sentinel-1 is equipped with twin polar-orbiting satellites and performs all-weather C-band SAR imaging. The GRD product consists of focused SAR data that have been detected, multi-looked, and projected to the ground range using an Earth ellipsoid model [46,47]. The GEE provides GRD data in three resolutions (10 m, 25 m, or 40 m), four band combinations, and three instrument modes. Each scene contains either 1 or 2 out of 4 possible polarization bands, depending on the instrument settings. In this study, except for VV and VH, the ratio form of these two bands, VV/VH and VH/VV, were also considered. All

SAR data were resampled to match a 30 m spatial resolution with Landsat imagery using cubic convolution.

2.3.3. Topographic Data

Topographic factors, including elevation, slope, and aspect of the ASTER Global Digital Elevation Model (GDEM) were obtained from the NASA Earth data platform (https://earthdata.nasa.gov, accessed on 15 November 2022). ASTER GDEM is a global-scale digital elevation model data with 30 m spatial resolution jointly released by NASA and Japan's Ministry of Economy, Trade, and Industry (METI) [48]. Compared with the previous version, version 2 mainly focuses on incorporating newer ASTER data and improving resolution by replacing the original 9 × 9 with a 5 × 5 window and partially removing biases and artifacts [48,49]. The data were resampled to 30 m of spatial resolution using cubic convolution.

2.4. Image Segmentation

Image segmentation is the first step for object-based image analysis. In this study, the multiresolution segmentation algorithm embedded in eCognition software version 9.0 was utilized. Based on the assumption of color and spatial continuity of individual image objects, this algorithm generates objects until a threshold representing the maximum object variation is reached [50]. The threshold parameters mainly consist of "scale", "shape", and "compactness". Among these, the "scale" value that defines the relative size of image objects is regarded as the most influential factor on the classification precision of the final map [32]. Thus, choosing an appropriate "scale" value is the most crucial step in image segmentation. In addition, the input layers are essential to the segmentation outcome. In this study, the selection of appropriate input layers and the above-mentioned tunable parameters was guided by prior knowledge and obtained through trial and error and visual inspection [8,32,39]. After loop testing various segmentation scales, the scale greater than 50 tended to under-segment the image with discernible mixtures of land cover, whereas the scale less than 50 tended to over-segment the image with adjacent objects of the same land cover observed. Moreover, when NDVI is used as the input layer, the differences between vegetation communities are more evident (Figure 2). Finally, blue, green, red, NIR, and NDVI were selected as input layers, and the "scale", "shape", and "compactness" were taken as 50, 0.1, and 0.5, respectively.

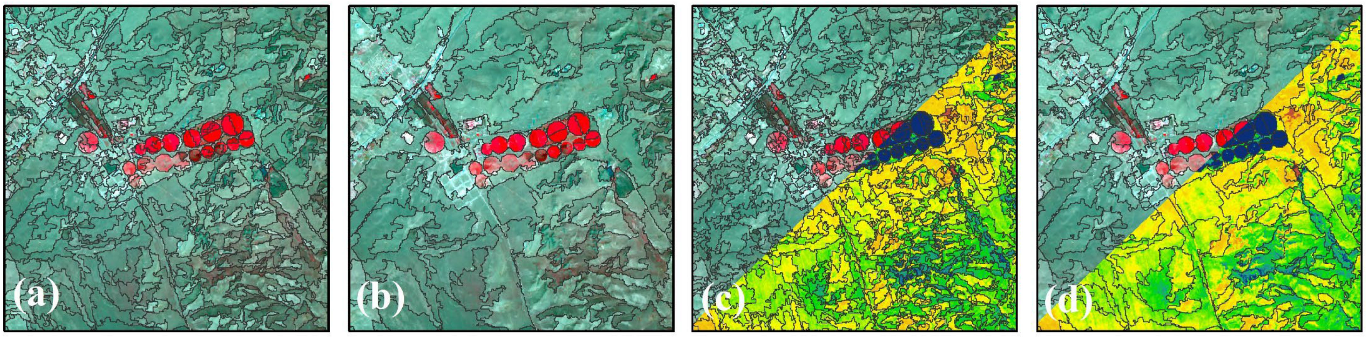

Figure 2. Image segmentation at different scales and input layers: (**a**) segmentation at a scale value of 30 with the multispectral band; (**b**) segmentation at a scale value of 50 with the multispectral band; (**c**) segmentation at a scale value of 30 in combination with the multispectral band and NDVI; (**d**) segmentation at scale value of 50 in combination with the multispectral band and NDVI.

2.5. Feature Selection

In this study, a total of 20 layers (features) were chosen as potential classification features for pixel-based image analysis from the multispectral bands (blue, green, red, NIR, SWIR1, and SWIR2), spectral indexes (NDVI, NDWI, and NDBI), topographic factors (elevation, aspect, and slope), position (pixel coordinates), and SAR variables (VH, VV, VH/VV, and VV/VH). Additionally, three different months of NDVI (April, July, and September) were employed to increase the separability between vegetation types (Table 2).

Table 2. Image layers used in pixel-based classifications.

Feature Types	Image Layers		No. of Features
S1	Multispectral band	Blue, green, red, NIR, SWIR1, SWIR2	6
S2	Spectral indices	NDVI-4, NDVI-7, NDVI-9, NDWI, NDBI	5
S3	Geography	Aspect, elevation, slope, X, Y	5
S4	SAR	VH, VV, VH/VV, VV/VH	4

Unlike pixels, object features permit contextual relationships between image objects to be incorporated into object-based image analysis. This means that the features are not limited to spectral, but can also include texture and geometry. In this study, a total of 158 object features belonging to spectral (mean, standard deviation), geometry, position, and texture were selected for object-based image analysis (Table 3). The object texture was analyzed using the gray-level co-occurrence matrix (GLCM) [51] and six commonly used statistical variables were calculated for each object: mean, homogeneity, dissimilarity, entropy, contrast, and correlation.

Table 3. Object features used in object-based classifications.

Feature Types		Object Features		No. of Features
C1	S1	Blue, green, red, NIR, SWIR1, SWIR2	Mean, standard deviation	12
	S2	NDVI-4, NDVI-7, NDVI-9, NDWI, NDBI		10
	S3	Aspect, elevation, slope		6
	S4	VH, VV, VH/VV, VV/VH		8
C2		Area, width, length, length/width, asymmetry, density, compactness, roundness	Geometry	8
C3		X max, X min, X center, Y max, Y min, Y center	Position	6
C4		Blue, green, red, NIR, SWIR1, SWIR2, NDVI-4, NDVI-7, NDVI-9, NDWI, NDBI, aspect, elevation, slope, VH, VV, VH/VV, VV/VH	Mean, homogeneity, dissimilarity, entropy, contrast, correlation	108

Recursive feature elimination (RFE), a wrapper-type feature selection algorithm, was applied to filter valuable features for both pixel-based and object-based classification. RFE operates by employing a base algorithm to identify a subset of features, then eliminating irrelevant features until only the useful component remains [52].

2.6. Classification Algorithms

RF is currently one of the most popular and widely used supervised classification algorithms on land use observation [26,27]. It is an extension of the bagging strategy as it utilizes both bagging and feature randomness to create a collection of uncorrelated trees [53]. Although RF has the least variability in its prediction accuracy among the popular machine learning algorithms [54], several tunable hyperparameters, such as the number of trees (ntree) and the number of features at each split point (mtry), remain to be considered. Among these, ntree controls the number of trees that stabilize the predicted error, and mtry determines the correlation between different trees. In addition, the tree complexity can be adjusted by the node size parameter.

SVMs are a set of supervised learning methods often used to solve land classification problems [29,30]. The core idea of SVM is to find a hyperplane that can classify data in the n-dimensional feature space by introducing the concept of the kernel function [55]. That is, the SVM aims to create the best line or decision boundary that can separate n-dimensional space into classes. As for the kernel function, it can either be linear or non-linear, depending on data complexity. In this study, the radial basis function (RBF), a commonly used non-linear kernel function, was applied. For the RBF, two tunable hyperparameters: penalty factor (c) and the kernel parameter (gamma) need to be adjusted. Among these, the penalty factor controls the trade-off between achieving a low error on the training data and minimizing the norm of the weights, and the gamma factor controls the distance of influence of a single training point [56].

KNN is a simple non-parametric supervised learning method that uses proximity to make classifications or predictions about the grouping of an individual data point [57]. That is, KNN only determines the group to which the data to be classified belongs based on the group of the nearest one or several samples. For classification problems, a sample is classified by a plurality vote of its neighbors, with it being assigned to the class most voted among its k-nearest neighbors. Obviously, the k value is the most critical tuning hyperparameter for the KNN algorithm. Generally, larger values of k reduce the effect of the noise on the classification but also make boundaries between classes less distinct [58].

NB is a set of supervised learning algorithms based on applying Bayes' theorem and the assumption of independent feature conditions, with a solid theoretical foundation and robust classification effect [59]. There is no single algorithm for training such classifiers except for a family of algorithms based on a shared principle. That is, all classifiers assume that the value of a particular feature is independent of the importance of any other part, given the class variable. In this study, Gaussian naive Bayes, which supports continuous-valued features and follows a Gaussian distribution, was applied as a classifier algorithm. This makes it possible to fit the model by simply computing the mean and standard deviation of the training data.

Model building and tuning of individual parameters were accomplished through repeated 5-fold cross-validation based on the training data. All the training procedures for models were carried out in R, and the involved main algorithm packages included "caret", "randomForest", "e1071", and "naivebayes".

2.7. Accuracy Assessment

All classifications were evaluated for accuracy using a confusion matrix [60], which is a standard format for accuracy evaluation consisting of a matrix with n rows and n columns. The confusion matrix contains two significant "single value" metrics for understanding classifier performance: the overall accuracy (OA), which represents the total number of accurate predictions across all classified classes, and the Kappa statistics, which measures the agreement between classification and truth values. For a specific category, it can be measured by the user accuracy (UA), which represents the probability that a value predicted to be in a certain class is that class, and the producer accuracy (PA), which is the probability that a value in a given class was classified correctly. In addition, the F1 score, the harmonic mean of the precision and recall, was used to evaluate the classification accuracy of each class.

McNemar's test [61] was used to determine whether or not there were prediction differences between each pair of classifiers in the same classification method and between the same classifier in two distinct classification methods. The McNemar's test is a non-parametric statistical test for paired nominal data, which checks the marginal homogeneity of two dichotomous variables. The null hypothesis is that there are no significant differences between selected classifiers. That is, if the calculated probability is lower than the selected significance level ($p < 0.05$), the null hypothesis is rejected, and it can be concluded that the classifiers' predictions are significantly different from each other.

3. Results

3.1. Pixel-Based Classifications

Table 4 lists the classification accuracy of the pixel-based classifiers under different feature combinations. Except for NB, the other three classifiers demonstrated significant prediction accuracy, with OA and Kappa exceeding 75% and 0.70, respectively, when only multispectral bands were employed as features. After spectral indices were added, the accuracy of all classifiers improved, particularly OA and Kappa of NB increased by 9.38% and 1.11. On this basis, accuracy improved further after adding geographic features, particularly for RF and SVM, with the OA reaching 95.58% and 95.31%, respectively, and the Kappa reaching 0.9467 and 0.9435. In contrast, adding SAR variables on the same basis did not improve accuracy significantly, and even KNN accuracy decreased. Although the classifiers had an acceptable prediction result in the condition of all features, it was not much different from the performance under the combination of the multispectral bands, spectral indices, and geographic features.

Table 4. Accuracy assessment for pixel-based classifiers under different feature combinations.

	RF		**SVM**		**KNN**		**NB**	
	OA (%)	**Kappa**	**OA (%)**	**Kappa**	**OA (%)**	**Kappa**	**OA (%)**	**Kappa**
S1	79.76	0.7553	82.98	0.7945	77.09	0.723	64.77	0.5813
S1 + S2	84.27	0.8096	85.92	0.8304	78.01	0.7338	74.15	0.6918
S1 + S2 + S3	95.58	0.9467	95.31	0.9435	80.40	0.7632	86.84	0.8416
S1 + S2 + S4	86.46	0.8365	86.48	0.837	76.63	0.7171	77.55	0.7323
S1 + S2 + S3 + S4	95.40	0.9445	95.77	0.949	79.21	0.7487	87.49	0.8495

The F1-score of each land cover under the five feature combinations is depicted in Figure 3. RF and SVM performed significantly better than the other two classifiers, particularly for grassland recognition. Under the combination of multispectral, spectral, and geographic features, the F1 score of each classifier was comparable to that under all features. This indicates that the SAR variables did not improve the performance of the model. In addition, geographic features significantly enhanced the F1 score of each land cover, especially grassland types. This demonstrates that topography and location played a crucial role in distinguishing land cover types. In contrast to other land types, SVM and NB performed poorly in identifying cropland and shrubland.

The relationship between classifier performance and the number of features is illustrated in Figure 4. It can be observed that classifiers have distinct feature requirements. After testing each feature, the best prediction was achieved when only 6 (NDVI-7, NDBI, NDWI, Elevation, X, and Y) and 15 (S1, S2, Elevation, X, VV, and VH) features were retained for RF and KNN, respectively. SVM and NB performed better when all features were used in the modeling process.

Figure 5 compares the classification confusion matrices of pixel-based methods utilizing the optimal features combination (The hyperparameter tuning process is shown in Appendix A.). RF presented the highest accuracy (OA: 96.32%, Kappa: 0.9556), followed by SVM (OA: 95.77%, Kappa: 0.949), NB (OA: 87.49%, Kappa: 0.8495), and KNN (OA: 80.86%, Kappa: 0.7692). In the figure, the orange upper and diagonal lines represent the PA and UA, respectively. From this, it can be observed that RF had a higher classification accuracy for each land cover. Except for cropland and shrubland, the PA and UA of other land covers were over 85%. SVM also performed well, except for cropland, where the UA was 69% and the rest of the land covers achieved an accuracy of over 80%. In the NB classification, 53% of cropland was mistakenly labeled as marshy meadow, which made the cropland UA only 47%. Other than these two types, the accuracy for all other land cover types exceeded 75%. In contrast, KNN returned a lower prediction for most land covers and a greater number of misclassifications and omissions.

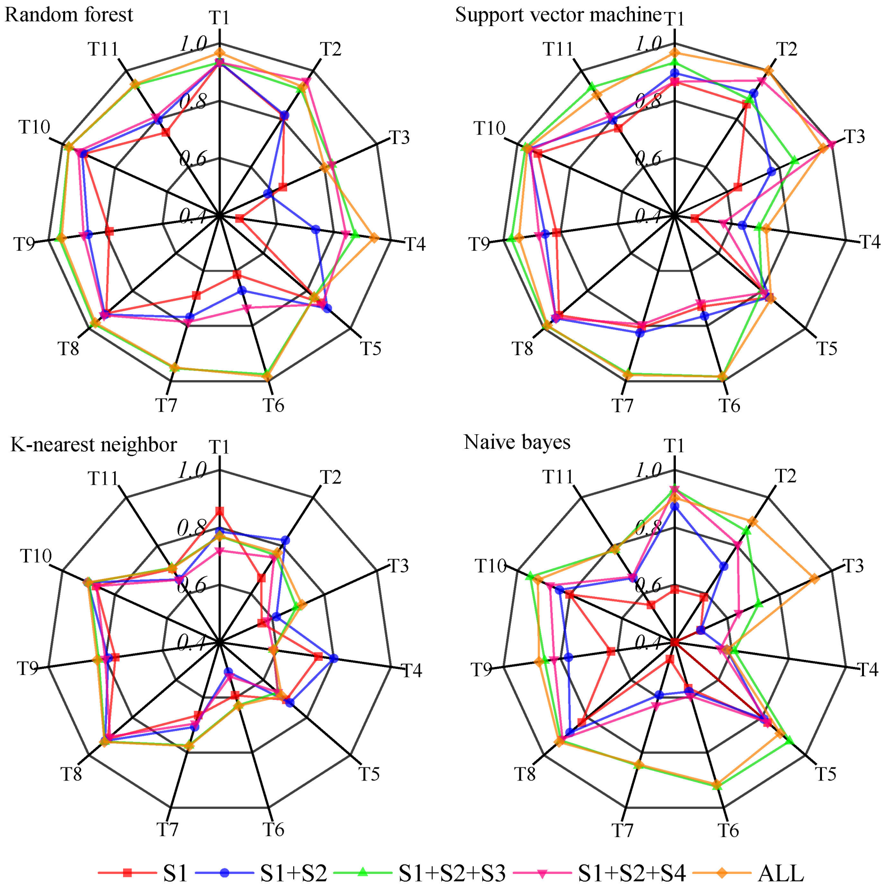

Figure 3. F1 score of pixel-based classifiers under different feature combinations.

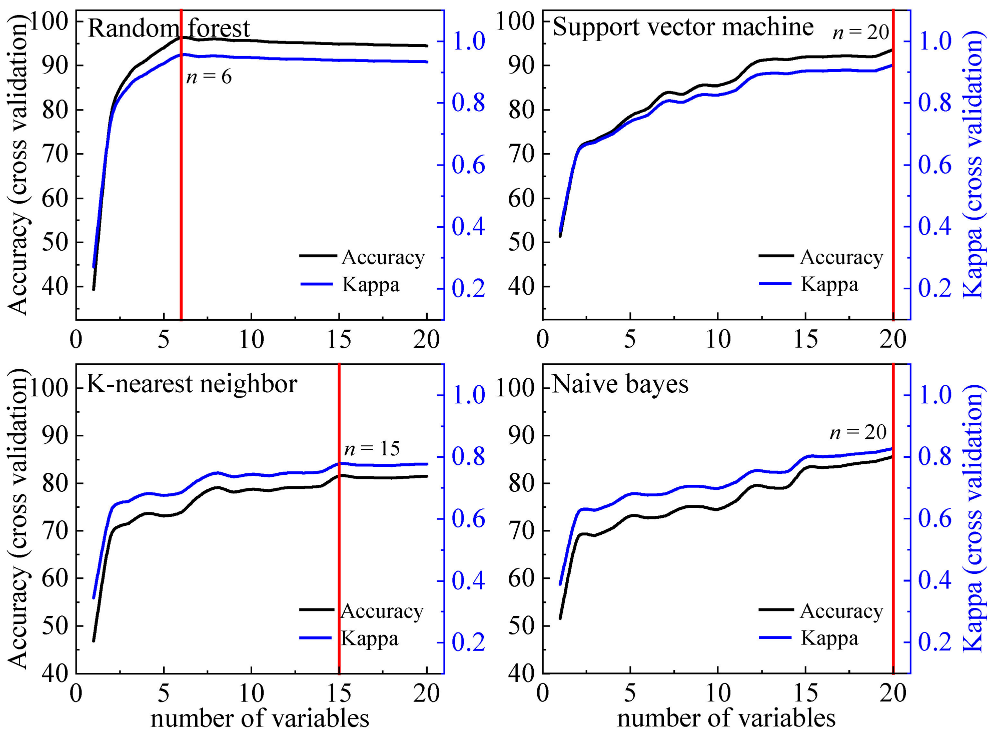

Figure 4. Relationship between the performance of pixel-based classifiers and the number of features (feature filtered by RFE).

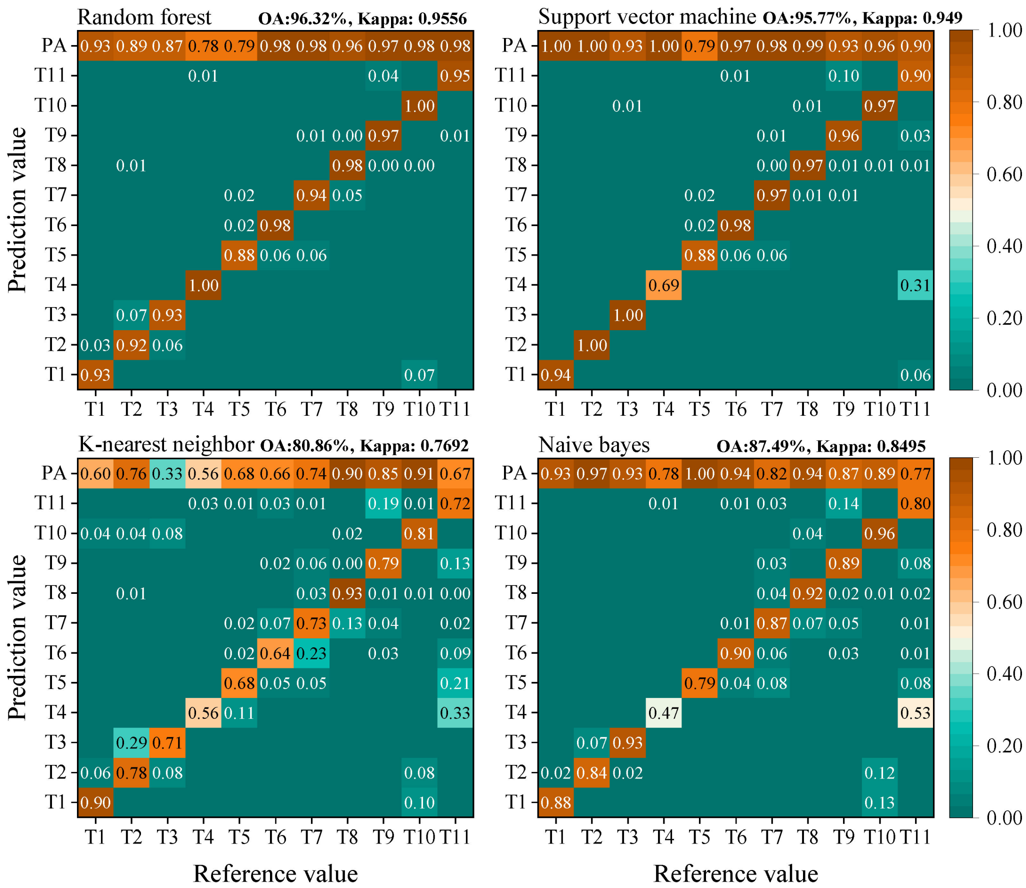

Figure 5. Confusion matrices for the optimized pixel-based classifications.

3.2. Object-Based Classifications

Table 5 lists the prediction accuracy of object-based classifiers under different feature combinations. The classifiers returned satisfactory results when only multispectral bands were used for modeling, particularly for RF and SVM, with OA and Kappa values of 0.9311 and 0.9281, respectively. This suggests that most land cover types can be distinguished using these spectral bands. After spectral indices were added, the accuracy of KNN and NB was significantly improved, with OA at 87.13% and 81.71% and Kappa at 0.8445 and 0.7804. The topographic variables also served to model improvement, particularly for NB, OA, and Kappa, which increased by 4.5% and 0.054, respectively. After adding a series of features from SAR, geometry, position, and texture, the performance of RF and SVM did not change much, while that of KNN and NB continuously increased. This implies that, in this study, excessive features were against model improvement and would hinder its performance.

Table 5. Accuracy assessment for object-based classifiers under different feature combinations.

	RF		**SVM**		**KNN**		**NB**	
	OA (%)	**Kappa**	**OA (%)**	**Kappa**	**OA (%)**	**Kappa**	**OA (%)**	**Kappa**
S1	94.30	0.9311	94.03	0.9279	85.48	0.8247	77.21	0.7269
S1 + S2	94.85	0.9379	93.84	0.9257	87.13	0.8445	81.71	0.7804
S1 + S2 + S3	96.88	0.9623	96.69	0.9601	88.33	0.8591	86.21	0.8344
C1	97.70	0.9722	97.43	0.9689	88.60	0.8625	88.79	0.8653
C1 + C2	97.43	0.9689	97.24	0.9667	87.50	0.8488	89.43	0.873
C1 + C2 + C3	**98.62**	**0.9833**	**97.59**	0.9754	94.68	0.9378	**92.12**	**0.9091**
C1 + C2 + C3 + C4	98.25	0.9789	97.43	**0.9689**	**94.94**	0.9389	91.36	0.8958

Figure 6 shows the F1-score of the object-based classifiers on each land cover. Under different feature conditions, RF and SVM outperformed the other two algorithms. It can be clearly seen, in both RF and SVM classifications, that the topographic features significantly improved the identifying capability of the model for the waterbody, and the position improved the accuracy of the mining area. Except for the hilly meadow steppe, the F1-score achieved over 0.9 for other grassland types when using only multispectral bands, indicating these two classifiers were sensitive to differences between grasslands. For all classifiers, most land covers presented the highest prediction accuracy when spectral, geometric, and position features were used.

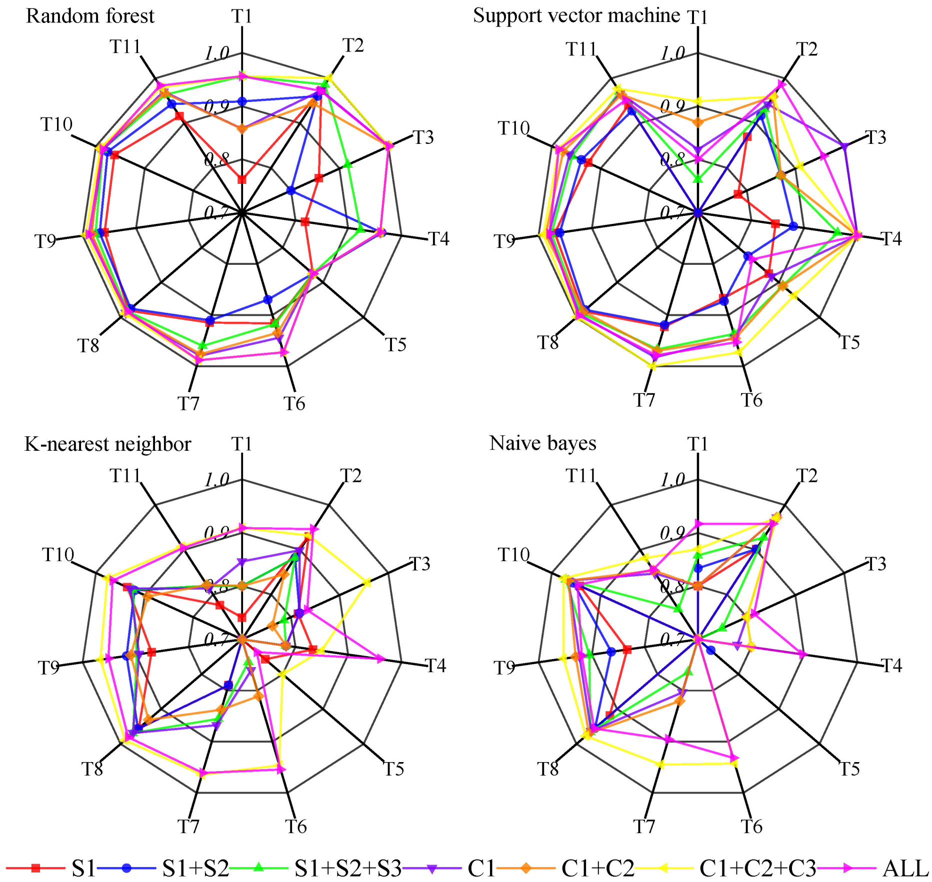

Figure 6. F1 score of object-based classifiers under different feature combinations.

As shown in Figure 7, selected classifiers have different requirements for features. RF had the most stable prediction accuracy after using only 10 features, and achieved the highest accuracy when retaining 32 useful features. SVM stabilized the accuracy after testing 60 features and performed best when selecting 95 valid features. The robustness of KNN and NB was inadequate, and the precision fluctuated as the number of features increased. The models were most accurate when all 158 features were retained for KNN and 40 features were retained for NB.

Table 6 lists the effective features of object-based classifiers. For RF, the effective features were multispectral bands, spectral indices, elevation, SAR, position, and texture statistics, whereas for SVM, some geometric features were chosen in addition to these. It can be noted that the selected texture features were more from homogeneity, entropy, and correlation statistics and presented irregularly. Among the selected features in NB, pixel mean and standard deviation accounted for the vast majority, and there were also some position and texture features. These indicate that multispectral bands, spectral indices, elevation, position, and SAR variables served critical roles in training these classifiers.

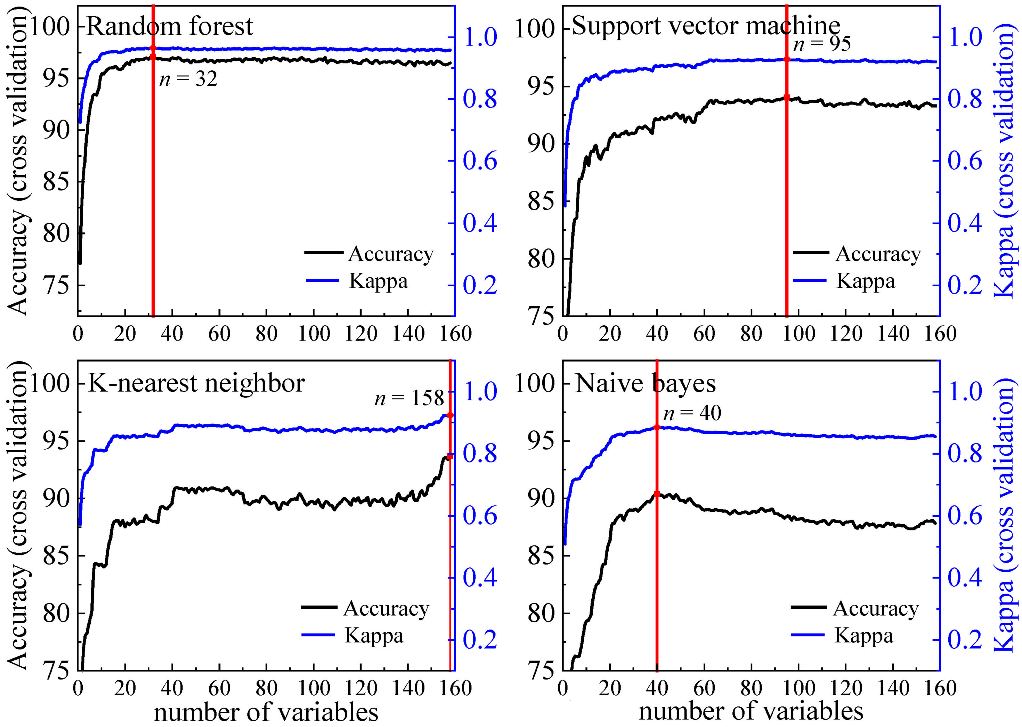

Figure 7. Relationship between the performance of object-based classifiers and the number of variables (feature filtered by RFE).

Table 6. Effective features of object-based classifiers.

Object-Based Algorithms	Selected Variables	No. of Variables
RF	Mean (S1, S2, elevation, VV, VH), standard deviation (blue, red, NDVI-7, NDWI, elevation), C3, homogeneity (blue, green, elevation), entropy (blue, elevation, slope), correlation (elevation)	32
SVM	Mean (S1, S2, S3, VV, VH), standard deviation (S1, S2, elevation), width, length, length/width, C3, GLCM mean (S1, S2), homogeneity (red, green, S2, S3), dissimilarity (S2), entropy (S1, elevation, S3, VV, VH), contrast (blue, green, NIR, S2, S3), correlation (red, S2, S3)	95
KNN	All	158
NB	Mean (red, NIR, SWIR1, SWIR2, S2, aspect, VV, VH), standard deviation (SWIR1, SWIR2, NDVI-4, NDVI-7, Aspect), X_{min}, X_{max}, X_{center}, GLCM mean (blue, green, red, NDVI-7, NDWI), homogeneity (red, SWIR1, SWIR2, NDVI-7, NDVI-9, NDBI, aspect, slope), entropy (SWIR1, SWIR2, aspect, slope, VV), dissimilarity (NDVI-7, NDVI-9)	40

Figure 8 compares the classification confusion matrices of object-based methods utilizing the optimal feature combinations (The hyperparameter tuning process is shown in Appendix A.). In general, the differences between the 4 classifications were not much; OA and Kappa were both higher than 90% and 0.9. The highest accuracy was achieved by RF (OA: 98.83%, Kappa: 0.9852), followed by SVM (OA: 97.63%, Kappa: 0.9733), KNN (OA: 94.94%, Kappa: 0.9389), and NB (OA: 92.56%, Kappa: 0.9101). Both RF and SVM had a considerable identification precision for each land cover, with few misclassifications and omissions. KNN achieved over 89% UA for all categories except shrubland and over 83% PA for all classes except cropland. NB performed poorly on each land cover classification, and the misclassification error was 33% and 42% for buildings and shrubland, respectively.

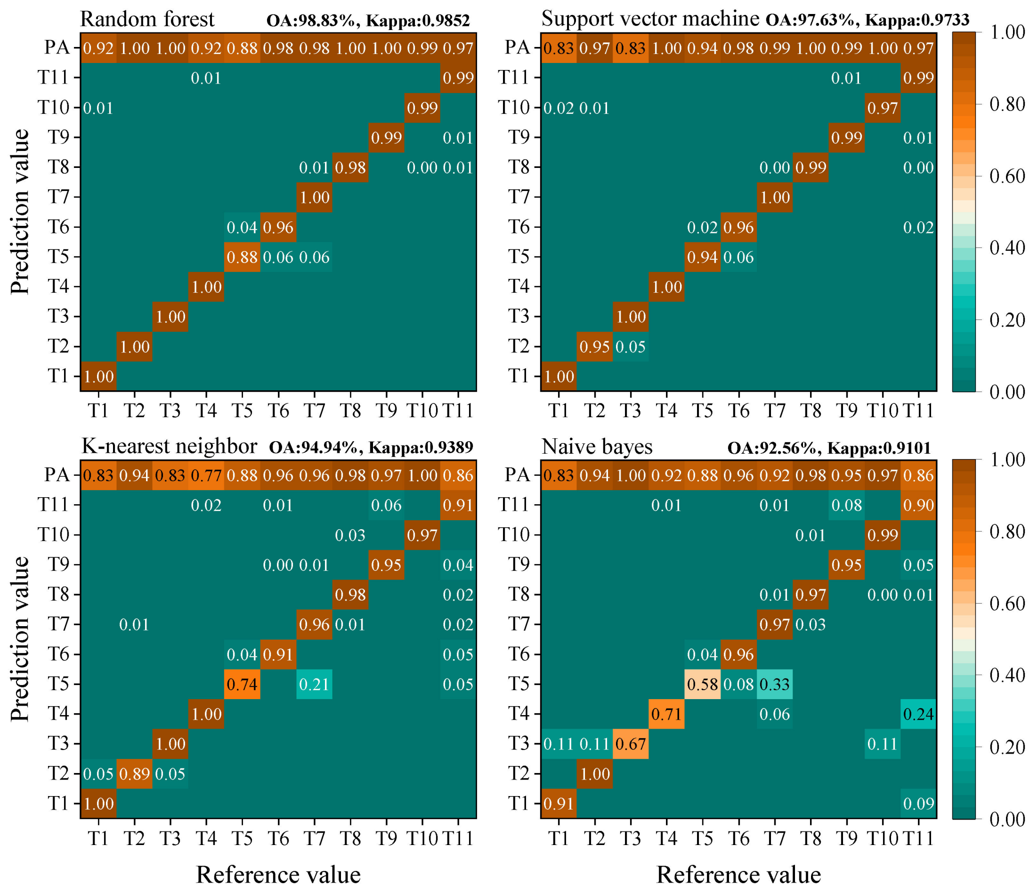

Figure 8. Confusion matrices for the optimized object-based classifications.

3.3. Comparison of Pixel-Based and Object-Based Classifications

The classification maps from each classifier, both pixel-based and object-based, are compared in Figure 9. In general, all classifications show a reasonably visual depiction of land covers in the study area. In pixel-based classifications, RF accurately described the actual grassland distribution, particularly in extracting sparsely distributed sandy steppe and hilly meadow steppe in complex terrain areas. It can be clearly seen that RF performed better in identifying some specific vegetation patches, such as marshy meadows in the river tidal flats and parks (Figure 9a). Although SVM also depicted a similar map, it underperformed in the extraction of marshy meadow, and there were omissions (Figure 9b). Both KNN and Bayes presented a relatively low performance in depicting some land types, such as hilly steppe and saline meadow, exaggerating their actual distribution. In addition, affected by the outlier pixel, the "salt-and-pepper noise" was also evident in these maps (Figure 9c,d). In object-based classifications, RF and SVM outperformed other classifiers for differentiating complex land covers, especially when the extracted water bodies, buildings, and mining areas were more consistent with the actual distributions (Figure 9). RF performed well in mapping marshy meadows and reflected the reed communities distributed along river banks and urban parks. Although NB described these land types, it was slightly inferior in identifying the specific boundaries (Figure 9h). When the same algorithm is compared, both pixel-based and object-based classifications showed similar patterns for most land covers, especially on the waterbody, building, mining area, and cropland areas.

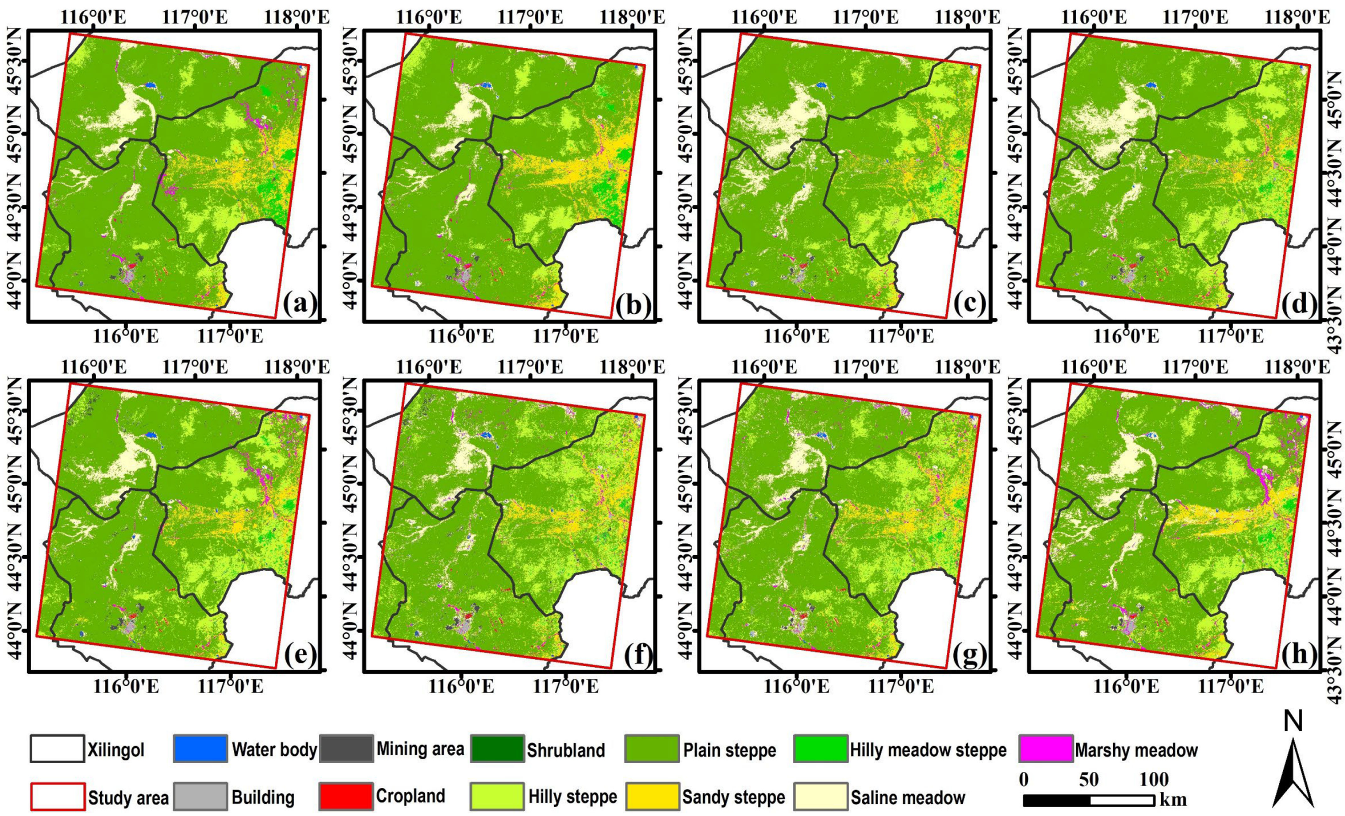

Figure 9. Comparison of pixel-based and object-based classification results: (**a**) pixel-based RF; (**b**) pixel-based SVM; (**c**) pixel-based KNN; (**d**) pixel-based NB; (**e**) object-based RF; (**f**) object-based SVM; (**g**) object-based KNN; (**h**) object-based NB.

The pixel-based algorithm pairings exhibited a statistically significant difference ($p < 0.05$) according to McNemar's test (Table 7). That is, the predictions of these algorithms for the study area were significantly different from each other. Most of the non-significant pairings were for the waterbody, shrubland, and hilly meadow steppe. There was no real difference between RF and SVM for predicting the waterbody ($p = 0.500$), cropland ($p = 1.000$), shrubland ($p = 0.250$), sandy steppe ($p = 0.092$), and saline meadow ($p = 0.302$), and no real difference between RF and KNN for predicting the waterbody ($p = 0.500$), hilly meadow steppe ($p = 0.125$), and hilly steppe ($p = 0.375$). RF and NB had the same predictions for the waterbody ($p = 0.500$), shrubland ($p = 0.063$), hilly meadow steppe ($p = 0.219$), and hilly steppe ($p = 0.375$). Apart from cropland, observations for all land covers were statistically consistent between KNN and NB.

Table 7. McNemar's test for each of the two classifiers in the pixel-based method.

	All	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11
RF vs. SVM	0.000	**0.500**	0.001	0.000	**1.000**	**0.250**	0.039	0.021	0.001	**0.092**	**0.302**	0.000
RF vs. KNN	0.000	**0.500**	0.000	0.000	0.000	0.002	**0.125**	**0.375**	0.000	0.000	0.000	0.000
RF vs. NB	0.000	**0.500**	0.000	0.000	0.006	**0.063**	**0.219**	**0.375**	0.000	0.000	0.000	0.000
SVM vs. KNN	0.000	**1.000**	0.038	0.000	0.000	0.016	**0.625**	0.001	0.035	0.000	0.000	0.000
SVM vs. NB	0.000	**1.000**	0.038	0.000	0.006	**0.500**	**0.250**	0.001	0.008	0.000	0.000	0.000
KNN vs. NB	0.002	**1.000**	**1.000**	**0.286**	0.000	**0.063**	**1.000**	**1.000**	**0.375**	**1.000**	**0.388**	**0.222**

McNemar's test (Table 8) determined that only the prediction of RF and SVM were statistically the same ($p = 0.618$) in the object-based pairing algorithms. Selected algorithms had the same predictions for the waterbody, shrubland, and hilly meadow steppe. This means these land covers are more prone to be identified. There were no significant differences between RF and SVM for predicting the waterbody (p = 1.000), shrubland (p = 0.063), hilly meadow steppe (p = 0.125), plain steppe (p = 0.089), sandy steppe (p = 0.519), and marshy meadow (p = 0.066), and no real difference between RF and KNN in the waterbody (p = 1.000), shrubland (p = 1.000), hilly meadow steppe (p = 1.000), and hilly steppe (p = 0.238). There was a statistical difference in observation of the building and mining area between RF and NB.

Table 8. McNemar's test for each of the two classifiers in the object-based method.

	All	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11
RF vs. SVM	**0.618**	**1.000**	0.000	0.000	0.016	**0.063**	**0.125**	0.004	**0.089**	**0.519**	0.000	**0.066**
RF vs. KNN	0.000	**1.000**	0.000	0.038	0.031	**0.250**	**1.000**	0.008	**0.238**	0.001	0.000	0.000
RF vs. NB	0.004	**0.500**	0.000	0.000	**0.500**	**1.000**	**0.500**	**0.063**	**0.388**	**0.222**	**0.710**	**0.085**
SVM vs. KNN	0.000	**1.000**	**1.000**	0.004	**1.000**	**0.500**	**0.250**	**1.000**	0.000	0.000	0.016	0.007
SVM vs. NB	0.002	**1.000**	**0.125**	**0.675**	**0.125**	**0.125**	**0.625**	**0.125**	0.011	**0.101**	0.000	0.000
KNN vs. NB	0.000	**0.500**	**0.219**	0.029	**0.219**	**0.500**	**1.000**	**0.250**	**0.804**	0.000	0.000	0.000

McNemar's test (Table 9) showed that only the paring classifications of pixel-based and object-based SVM were statistically consistent (p = 0.318) and the other three algorithms had significantly different predictions. These algorithms had the same predicting accuracies for the waterbody, shrubland, and hilly meadow steppe. There was no statistical difference in the waterbody (p = 1.000), cropland (p = 1.000), shrubland (p = 1.000), hilly meadow steppe (p = 1.000), and hilly steppe (p = 0.125) between two RF predictions. SVM had the same prediction on most of the land covers, except for the building and sandy steppe. The paring KNN had the same predictions on the waterbody (p = 0.500), shrubland (p = 0.065), and hilly meadow steppe (p = 0.063), and presented different observations on other classes. There was no real difference between the two NB predictions on the waterbody (p = 1.000), shrubland (p = 0.219), hilly meadow steppe (p = 0.250), and hilly steppe (p = 0.219).

Table 9. McNemar's test for the same classifier based on different classification methods.

	All	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11
RF vs. RF	0.000	**1.000**	0.000	0.000	**1.000**	**1.000**	**1.000**	**0.125**	0.000	0.000	0.000	0.000
SVM vs. SVM	**0.318**	**1.000**	0.000	**0.403**	**0.070**	**0.727**	**0.289**	**0.629**	**0.054**	0.007	**0.424**	**1.000**
KNN vs. KNN	0.000	**0.500**	0.000	0.000	0.000	**0.065**	**0.063**	0.039	0.007	0.000	0.000	0.000
Bayes vs. NB	0.000	**1.000**	0.000	0.000	0.012	**0.219**	**0.250**	**0.219**	0.003	0.000	0.000	0.000

4. Discussion

4.1. Comparison of Classification Methods for Semi-Arid Grassland

In contrast, object-based classifications presented a more realistic land cover, even though both pixel-based and object-based methods offered similar land cover depictions of the study area. According to McNemar's test, object-based classification outperformed pixel-based classification even when using the same algorithm, indicating that this technique is more appropriate for grassland classification tasks. This is consistent with most previous studies [62–66]. Object-based classification is more closely aligned with the visual interpretation process than pixel-based classification, allowing classifiers to identify target objects by combining the use of the objects themselves and contextual information. Although some studies suggested that the object-based analysis is better suited to high spatial resolution imagery [38,62,67], this study considers that it is also ideal for classifying open grassland with coarser resolution imagery. This is because a grassland type perhaps contains many different communities (or herbages), and pixel-based statistics are easier to be affected by outliers, thus exaggerating or reducing the potential differences of these grasslands. While some object characteristics, such as mean and standard deviation, can avoid this to some extent, RF and SVM are consistently the most widely used algorithms due to their few tunable hyperparameters and high performance [68]. Numerous studies have shown that both algorithms are effective at classifying grasslands [8,20–22]. The results show that RF had the highest accuracy in the pixel-based classification, while in the object-based classification, both RF and SVM significantly outperformed the other two algorithms and were not statistically different from one another. However, RF was superior in terms of feature selection and model stability, offering significant prediction under a limited number of variables. This suggests that RF is better suited for semi-arid area grassland classification tasks. KNN demonstrated high prediction accuracy in both pixel-based and object-based classifications despite being a straightforward nonparametric algorithm. However, these algorithms, including NB, are prone to sample imbalance and are erroneous in their assessments of rare categories. The results show that there was significant misclassification and omission in the classification of some land covers for both algorithms.

4.2. Feature Selection for Semi-Arid Grassland Classifications

A growing number of studies have concentrated on using high spatial resolution and hyperspectral imagery to produce a high-quality map of land cover classification as remote sensing techniques have advanced. However, most places lack access to these data sources, particularly in open grassland. Due to this, medium-resolution multispectral imagery from series such as the Landsat or Sentinel continues to be the primary source of information for grassland observation in these regions. Meanwhile, the grassland in northern China is characterized by low-vegetated, complex community structures, all of which make grassland classification more challenging. Most studies suggested improving the separability of land covers through the enrichment of classification features [12,19,69]. This study investigated the potential impact of multisource features on the classification of grasslands. Although selected methods performed different requirements for features, some similarities could still be seen. In the case of land covers, particularly grassland types, geographical factors, such as topography and position, had a significant impact on classification accuracy. This is because the CGCS is a standard for classifying grasslands based on the habitat of the vegetation, and each class is determined by the topographic and climatic characteristics of the vegetation. Geographical conditions still play a role in determining vegetation distribution, even for other standards. Furthermore, in this study, these methods demonstrate that too many features hinder prediction. In other words, removing redundant features is crucial for enhancing model performance. Especially for object-based classification, since features are not limited to pixel values but can be geometry, texture, and context, it is challenging to reasonably filter influential variables. The results show that only a few GLCM texture statistics helped with classification and that geometric features were inappropriate for classifying land cover. This is because native grassland lacks obvious geometric features, making it difficult to distinguish it from other regular-shaped land types such as buildings and cropland. SAR has been widely used for land classification in recent years and has the advantages of strong penetration and all-weather observation [70,71]. According to this study, SAR had no substantial role in pixel-based classification. Although it was selected as an effective variable in object-based classification, no significant contribution to the grassland distinction was seen. These show that the Sentinel 1 SAR C band is inappropriate for classifying semi-arid grasslands. Since effective scattering only happens when the particles are on the scale of the radar wavelength [72], this may have something to do with why most grassland types did not meet the SAR's scattering requirement.

4.3. Limitations and Uncertainties

Although it has been established that remote-sensing-based grassland classification is somewhat feasible, there are still some limitations and uncertainties. First, unlike the land covers with apparent perceptual characteristics, such as buildings, cropland, and forest, grassland classification is a highly abstract idea, and a defined class may include a variety of herbages and communities. This makes it challenging to find their similarities to create a classification identity. For instance, even though the CGCS specifies the habitat traits for each type, it is more of an empirical range and falls short of including all potential grasslands. Second, grasslands are susceptible to external disturbances such as climate changes and human activities, and community structures are prone to changes. Because of this, it is difficult to accurately represent a natural grassland using the few remote sensing data sources available. Additionally, the spectral and temporal resolution of the available data makes it challenging to accurately reflect the complex phenology of grassland vegetation. In these regards, we consider that different classification systems should be formulated according to the classification purpose and scale and that a more flexible standard should be used for grassland classification accuracy assessment. In addition, we recommend using multisource data, including multi-temporal data, to try to improve the separability between vegetation communities from multiple perspectives. The transferability aspect is also important to highlight the domain adaptation of the models proposed in this study. Understanding how classifiers learn from a source domain and generalize to a target domain is critical for applying the models in other areas. In light of this, we proposed that the sample-based and feature-based methods could be used [73]. They are appropriate to transfer the model to the target region because they have been widely used in land use and land cover classification for aligning disparities between domains so that the trained model can be generalized into the domain of interest [74]. Density ratio estimation is the typical solution for minimizing the discrepancy between the re-weighted source sample distribution and the actual target sample distribution [75]. A deep residual network is also being developed as a feature-based method to address the problem of step dispersion during model training for the target domain [76].

5. Conclusions

In this study, the central Xilingol (China) was selected as the study area, and the performance of four commonly used machine learning algorithms (RF, SVM, KNN, and NB) in classifying semi-arid grassland under pixel-based and object-based classification methods was compared. Meanwhile, the capabilities of Sentinel C-band SAR, Landsat OLI multispectral bands, texture features, and geographical factors for grassland classification were evaluated. Overall, the object-based methods were more accurate and depicted a more realistic land cover distribution than the pixel-based methods. In the pixel-based classification, RF performed best, and OA and Kappa achieved 96.32% and 0.9556, respectively, when only 6 valid features were used. In the object-based classification, the prediction of RF and SVM had no statistical difference, but both were significantly higher than other algorithms. Given feature selection, robustness, and prediction accuracy, the object-based RF method is more suitable for semi-arid grassland classification tasks in northern China. In pixel-based classification, multispectral bands, spectral indices, and geographic features contribute significantly to distinguishing grassland, whereas, in object-based classification, predominant features are more derived from multispectral bands, spectral indices, elevation, and position features. Although Sentinel 1 SAR variables were selected as an effective variable in object-based classification, there was no significant contribution to the grassland distinction.

Author Contributions: Conceptualization, N.W. and G.H.; methodology, N.W.; software, N.W.; validation, G.L. and D.W.; formal analysis, G.L.; writing—original draft preparation, N.W.; writing—review and editing, L.G.T.C.; project administration, G.H.; funding acquisition, G.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Innovative Research Team of Ministry of Education of China (No. IRT_17R59) and the Inner Mongolia Key Project (No. ZDZX2018020).

Data Availability Statement: Not applicable.

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

Appendix A

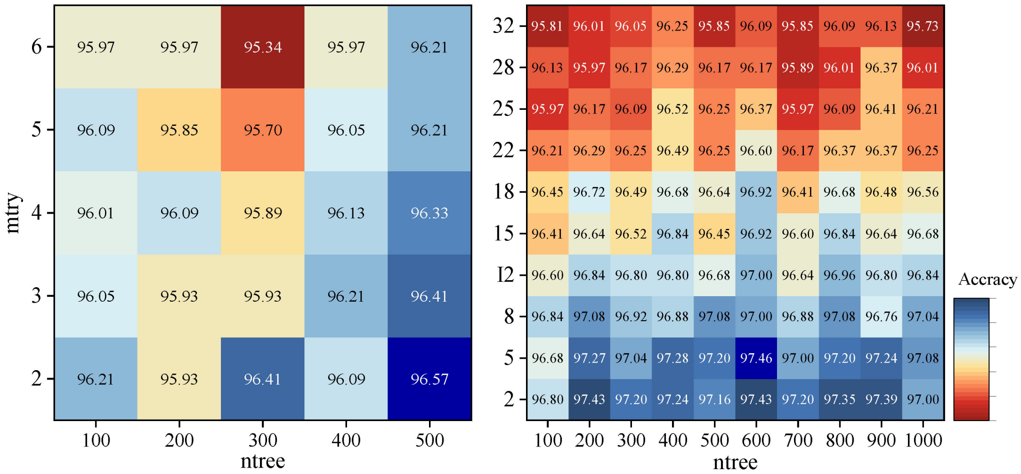

Figure A1. Hyperparameter grid search of RF in pixel-based (**right**) and object-based (**left**) classification methods under optimal feature combination.

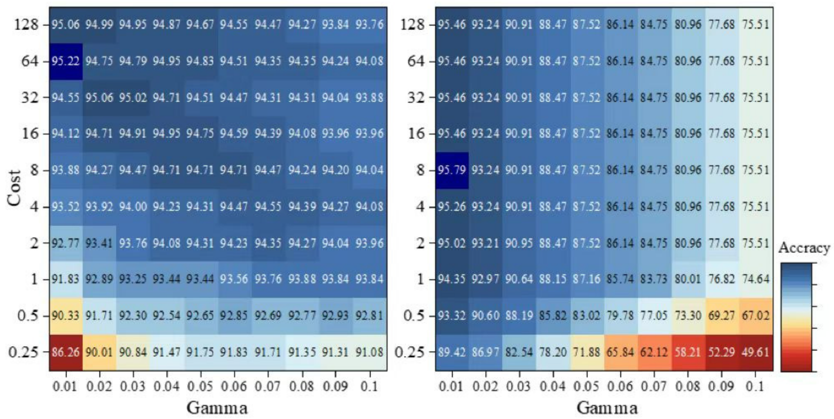

Figure A2. Hyperparameter grid search of SVM in pixel-based (**right**) and object-based (**left**) classification methods under optimal feature combination.

References

1. Bengtsson, J.; Bullock, J.M.; Egoh, B.; Everson, C.; Everson, T.; O'Connor, T.; O'Farrell, P.J.; Smith, H.G.; Lindborg, R. Grasslands—More important for ecosystem services than you might think. *Ecosphere* **2019**, *10*, e02582. [CrossRef]
2. O'Mara, F.P. The role of grasslands in food security and climate change. *Ann. Bot.* **2012**, *110*, 1263–1270. [CrossRef] [PubMed]
3. Richter, F.; Jan, P.; El Benni, N.; Lüscher, A.; Buchmann, N.; Klaus, V.H. A guide to assess and value ecosystem services of grasslands. *Ecosyst. Serv.* **2021**, *52*, 101376. [CrossRef]
4. Zhao, Y.; Liu, Z.; Wu, J. Grassland ecosystem services: A systematic review of research advances and future directions. *Landsc. Ecol.* **2020**, *35*, 793–814. [CrossRef]
5. Tilman, D.; Reich, P.B.; Knops, J.M. Biodiversity and ecosystem stability in a decade-long grassland experiment. *Nature* **2006**, *441*, 629–632. [CrossRef]
6. Nan, Z. The grassland farming system and sustainable agricultural development in China. *Grassl. Sci.* **2005**, *51*, 15–19. [CrossRef]
7. Wu, N.; Liu, A.; Ye, R.; Yu, D.; Du, W.; Chaolumeng, Q.; Liu, G.; Yu, S. Quantitative analysis of relative impacts of climate change and human activities on Xilingol grassland in recent 40 years. *Glob. Ecol. Conserv.* **2021**, *32*, e01884. [CrossRef]
8. Xu, D.; Chen, B.; Shen, B.; Wang, X.; Yan, Y.; Xu, L.; Xin, X. The Classification of Grassland Types Based on Object-Based Image Analysis with Multisource Data. *Rangel. Ecol. Manag.* **2019**, *72*, 318–326. [CrossRef]
9. Xie, Y.; Sha, Z.; Yu, M. Remote sensing imagery in vegetation mapping: A review. *J. Plant Ecol.* **2008**, *1*, 9–23. [CrossRef]
10. Wang, Z.; Ma, Y.; Zhang, Y.; Shang, J. Review of Remote Sensing Applications in Grassland Monitoring. *Remote Sens.* **2022**, *14*, 2903. [CrossRef]
11. Lu, D.; Weng, Q. A survey of image classification methods and techniques for improving classification performance. *Int. J. Remote Sens.* **2007**, *28*, 823–870. [CrossRef]

12. Barrett, B.; Nitze, I.; Green, S.; Cawkwell, F. Assessment of multi-temporal, multi-sensor radar and ancillary spatial data for grasslands monitoring in Ireland using machine learning approaches. *Remote Sens. Environ.* **2014**, *152*, 109–124. [CrossRef]
13. Corbane, C.; Lang, S.; Pipkins, K.; Alleaume, S.; Deshayes, M.; García Millán, V.E.; Strasser, T.; Vanden Borre, J.; Toon, S.; Michael, F. Remote sensing for mapping natural habitats and their conservation status—New opportunities and challenges. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *37*, 7–16. [CrossRef]
14. Lopatin, J.; Fassnacht, F.E.; Kattenborn, T.; Schmidtlein, S. Mapping plant species in mixed grassland communities using close range imaging spectroscopy. *Remote Sens. Environ.* **2017**, *201*, 12–23. [CrossRef]
15. Marcinkowska-Ochtyra, A.; Gryguc, K.; Ochtyra, A.; Kopeć, D.; Jarocińska, A.; Sławik, Ł. Multitemporal Hyperspectral Data Fusion with Topographic Indices—Improving Classification of Natura 2000 Grassland Habitats. *Remote Sens.* **2019**, *11*, 2264. [CrossRef]
16. Mansour, K.; Mutanga, O.; Everson, T.; Adam, E. Discriminating indicator grass species for rangeland degradation assessment using hyperspectral data resampled to AISA Eagle resolution. *ISPRS J. Photogramm. Remote Sens.* **2012**, *70*, 56–65. [CrossRef]
17. Meng, B.; Yang, Z.; Yu, H.; Qin, Y.; Sun, Y.; Zhang, J.; Chen, J.; Wang, Z.; Zhang, W.; Li, M.; et al. Mapping of Kobresia pygmaea Community Based on Umanned Aerial Vehicle Technology and Gaofen Remote Sensing Data in Alpine Meadow Grassland: A Case Study in Eastern of Qinghai–Tibetan Plateau. *Remote Sens.* **2021**, *13*, 2483. [CrossRef]
18. Corbane, C.; Alleaume, S.; Deshayes, M. Mapping natural habitats using remote sensing and sparse partial least square discriminant analysis. *Int. J. Remote Sens.* **2013**, *34*, 7625–7647. [CrossRef]
19. Buck, O.; Millán, V.E.G.; Klink, A.; Pakzad, K. Using information layers for mapping grassland habitat distribution at local to regional scales. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *37*, 83–89. [CrossRef]
20. Zhao, Y.; Zhu, W.; Wei, P.; Fang, P.; Zhang, X.; Yan, N.; Liu, W.; Zhao, H.; Wu, Q. Classification of Zambian grasslands using random forest feature importance selection during the optimal phenological period. *Ecol. Indic.* **2022**, *135*, 108529. [CrossRef]
21. Rapinel, S.; Mony, C.; Lecoq, L.; Clément, B.; Thomas, A.; Hubert-Moy, L. Evaluation of Sentinel-2 time-series for mapping floodplain grassland plant communities. *Remote Sens. Environ.* **2019**, *223*, 115–129. [CrossRef]
22. Badreldin, N.; Prieto, B.; Fisher, R. Mapping Grasslands in Mixed Grassland Ecoregion of Saskatchewan Using Big Remote Sensing Data and Machine Learning. *Remote Sens.* **2021**, *13*, 4972. [CrossRef]
23. Hong, G.; Zhang, A.; Zhou, F.; Brisco, B. Geoinformation. Integration of optical and synthetic aperture radar (SAR) images to differentiate grassland and alfalfa in Prairie area. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *28*, 12–19.
24. Pulliainen, J.T.; Kurvonen, L.; Hallikainen, M.T. Multitemporal behavior of L- and C-band SAR observations of boreal forests. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 927–937. [CrossRef]
25. Forkuor, G.; Benewinde Zoungrana, J.-B.; Dimobe, K.; Ouattara, B.; Vadrevu, K.P.; Tondoh, J.E. Above-ground biomass mapping in West African dryland forest using Sentinel-1 and 2 datasets—A case study. *Remote Sens. Environ.* **2020**, *236*, 111496. [CrossRef]
26. Chang, S.; Wang, Z.; Mao, D.; Guan, K.; Jia, M.; Chen, C. Mapping the Essential Urban Land Use in Changchun by Applying Random Forest and Multi-Source Geospatial Data. *Remote Sens.* **2020**, *12*, 2488. [CrossRef]
27. Rodriguez-Galiano, V.F.; Ghimire, B.; Rogan, J.; Chica-Olmo, M.; Rigol-Sanchez, J.P. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J. Photogramm. Remote Sens.* **2012**, *67*, 93–104. [CrossRef]
28. Pacheco, A.D.P.; Junior, J.A.d.S.; Ruiz-Armenteros, A.M.; Henriques, R.F.F. Assessment of k-Nearest Neighbor and Random Forest Classifiers for Mapping Forest Fire Areas in Central Portugal Using Landsat-8, Sentinel-2, and Terra Imagery. *Remote Sens.* **2021**, *13*, 1345. [CrossRef]
29. Jamsran, B.-E.; Lin, C.; Byambakhuu, I.; Raash, J.; Akhmadi, K. Applying a support vector model to assess land cover changes in the Uvs Lake Basin ecoregion in Mongolia. *Inf. Process. Agric.* **2019**, *6*, 158–169. [CrossRef]
30. Ul Din, S.; Mak, H.W.L. Retrieval of Land-Use/Land Cover Change (LUCC) Maps and Urban Expansion Dynamics of Hyderabad, Pakistan via Landsat Datasets and Support Vector Machine Framework. *Remote Sens.* **2021**, *13*, 3337. [CrossRef]
31. Dusseux, P.; Vertès, F.; Corpetti, T.; Corgne, S.; Hubert-Moy, L. Agricultural practices in grasslands detected by spatial remote sensing. *Environ. Monit. Assess.* **2014**, *186*, 8249–8265. [CrossRef] [PubMed]
32. Duro, D.C.; Franklin, S.E.; Dubé, M.G. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. *Remote Sens. Environ.* **2012**, *118*, 259–272. [CrossRef]
33. Liu, D.; Xia, F. Assessing object-based classification: Advantages and limitations. *Remote Sens. Lett.* **2010**, *1*, 187–194. [CrossRef]
34. Flanders, D.; Hall-Bayer, M.; Pereverzoff, J. Preliminary evaluation of eCognition object-based software for cut block delineation and feature extraction. *Can. J. Remote Sens.* **2003**, *29*, 441–452. [CrossRef]
35. Li, C.; Wang, J.; Wang, L.; Hu, L.; Gong, P. Comparison of Classification Algorithms and Training Sample Sizes in Urban Land Classification with Landsat Thematic Mapper Imagery. *Remote Sens.* **2014**, *6*, 964–983. [CrossRef]
36. Autrey, B.; Reif, M. Segmentation and object-oriented classification of wetlands in a karst Florida landscape using multi-season Landsat7 ETM+ imagery. *Int. J. Remote Sens.* **2011**, *32*, 1471–1489. [CrossRef]
37. Zerrouki, N.; Bouchaffra, D. Pixel-Based or Object-Based: Which Approach is More Appropriate for Remote Sensing Image Classification? In Proceedings of the International Conference on Systems, Man, and Cybernetics (SMC), Banff, AB, Canada, 5–8 October 2014.

38. Oruc, M.; Marangoz, A.; Buyuksalih, G. Comparison of Pixel-Based and Object-Oriented Classification Approaches Using Landsat-7 ETM Spectral Bands. In Proceedings of the 20th ISPRS Congress on Technical Commission VII, Istanbul, Turkey, 12–23 July 2004; pp. 1118–1123.
39. Dingle Robertson, L.; King, D. Comparison of pixel- and object-based classification in land cover change mapping. *Int. J. Remote Sens.* **2011**, *32*, 1505–1529. [CrossRef]
40. Li, F.Y.; Jäschke, Y.; Guo, K.; Wesche, K. Grasslands of China. In *Encyclopedia of the World's Biomes*; Goldstein, M.I., DellaSala, D.A., Eds.; Elsevier: Oxford, UK, 2020; pp. 773–784.
41. Ministry of Agriculture of the People's Republic of China. *Rangeland Resources of China*; China Science and Technology Press: Beijing, China, 1996.
42. Xilingol League Grassland Station. *Xilingol Grassland Resources*; Inner Mongolia Daily Youth Press: Hohhot, China, 1988.
43. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [CrossRef]
44. Zha, Y.; Gao, J.; Ni, S. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *Int. J. Remote Sens.* **2003**, *24*, 583–594. [CrossRef]
45. McFeeters, S.K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* **1996**, *17*, 1425–1432. [CrossRef]
46. Soudani, K.; Delpierre, N.; Berveiller, D.; Hmimina, G.; Vincent, G.; Morfin, A.; Dufrêne, É. Potential of C-band Synthetic Aperture Radar Sentinel-1 time-series for the monitoring of phenological cycles in a deciduous forest. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *104*, 102505. [CrossRef]
47. Mullissa, A.; Vollrath, A.; Odongo-Braun, C.; Slager, B.; Balling, J.; Gou, Y.; Gorelick, N.; Reiche, J. Sentinel-1 SAR Backscatter Analysis Ready Data Preparation in Google Earth Engine. *Remote Sens.* **2021**, *13*, 1954. [CrossRef]
48. Abrams, M.; Crippen, R.; Fujisada, H. ASTER Global Digital Elevation Model (GDEM) and ASTER Global Water Body Dataset (ASTWBD). *Remote Sens.* **2020**, *12*, 1156. [CrossRef]
49. Tachikawa, T.; Kaku, M.; Iwasaki, A.; Gesch, D.; Oimoen, M.; Zhang, Z.; Danielson, J.; Krieger, T.; Curtis, B.; Haase, J. ASTER Global Digital Elevation Model Version 2—Summary of Validation Results. NASA: Washington, DC, USA, 2011.
50. Nussbaum, S.; Menz, G. *eCognition Image Analysis Software*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 29–39.
51. Haralick, R.M.; Shanmugam, K.; Dinstein, I. Textural Features for Image Classification. *IEEE Trans. Syst. Man Cybern.* **1973**, *SMC-3*, 610–621. [CrossRef]
52. Guyon, I.; Weston, J.; Barnhill, S.; Vapnik, V. Gene Selection for Cancer Classification using Support Vector Machines. *Mach. Learn.* **2002**, *46*, 389–422. [CrossRef]
53. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
54. Probst, P.; Bischl, B.; Boulesteix, A.-L. Tunability: Importance of hyperparameters of machine learning algorithms. *J. Mach. Learn. Res.* **2018**, *20*, 1934–1965.
55. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]
56. Awad, M.; Khanna, R. Support Vector Regression. In *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*; Awad, M., Khanna, R., Eds.; Apress: Berkeley, CA, USA, 2015; pp. 67–80.
57. Cover, T.; Hart, P. Nearest neighbor pattern classification. *IEEE Trans. Inform. Theory* **1967**, *13*, 21–27. [CrossRef]
58. Landau, S.; Leese, M.; Stahl, D.; Everitt, B.S. *Cluster Analysis*; John Wiley & Sons: Hoboken, NJ, USA, 2011.
59. Russell, S.J. *Artificial Intelligence: A Modern Approach*, 3rd ed.; Prentice Hall: Upper Saddle River, NJ, USA, 2010.
60. Townsend, J.T. Theoretical analysis of an alphabetic confusion matrix. *Percept. Psychophys.* **1971**, *9*, 40–50. [CrossRef]
61. McNemar, Q. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* **1947**, *12*, 153–157. [CrossRef]
62. Whiteside, T.G.; Boggs, G.S.; Maier, S.W. Comparing object-based and pixel-based classifications for mapping savannas. *Int. J. Appl. Earth Observ. Geoinform.* **2011**, *13*, 884–893. [CrossRef]
63. Fu, B.; Wang, Y.; Campbell, A.; Li, Y.; Zhang, B.; Yin, S.; Xing, Z.; Jin, X. Comparison of object-based and pixel-based Random Forest algorithm for wetland vegetation mapping using high spatial resolution GF-1 and SAR data. *Ecol. Indic.* **2017**, *73*, 105–117. [CrossRef]
64. Du, B.; Mao, D.; Wang, Z.; Qiu, Z.; Yan, H.; Feng, K.; Zhang, Z. Mapping Wetland Plant Communities Using Unmanned Aerial Vehicle Hyperspectral Imagery by Comparing Object/Pixel-Based Classifications Combining Multiple Machine-Learning Algorithms. *IEEE J. Select. Top. Appl. Earth Observ. Remote Sens.* **2021**, *14*, 8249–8258. [CrossRef]
65. Qu, L.a.; Chen, Z.; Li, M.; Zhi, J.; Wang, H. Accuracy Improvements to Pixel-Based and Object-Based LULC Classification with Auxiliary Datasets from Google Earth Engine. *Remote Sens.* **2021**, *13*, 453. [CrossRef]
66. Tassi, A.; Gigante, D.; Modica, G.; Di Martino, L.; Vizzari, M. Pixel- vs. Object-Based Landsat 8 Data Classification in Google Earth Engine Using Random Forest: The Case Study of Maiella National Park. *Remote Sens.* **2021**, *13*, 2299. [CrossRef]
67. Qian, Y.; Zhou, W.; Yan, J.; Li, W.; Han, L. Comparing Machine Learning Classifiers for Object-Based Land Cover Classification Using Very High Resolution Imagery. *Remote Sens.* **2015**, *7*, 153–168. [CrossRef]
68. Sheykhmousa, M.; Mahdianpari, M.; Ghanbari, H.; Mohammadimanesh, F.; Ghamisi, P.; Homayouni, S. Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 6308–6325. [CrossRef]

69. Abdi, A.M. Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data. *GIScience Remote Sens.* **2020**, *57*, 1–20. [CrossRef]
70. Parikh, H.; Patel, S.; Patel, V. Classification of SAR and PolSAR images using deep learning: A review. *Int. J. Image Data Fusion* **2020**, *11*, 1–32. [CrossRef]
71. Joshi, N.; Baumann, M.; Ehammer, A.; Fensholt, R.; Grogan, K.; Hostert, P.; Jepsen, M.R.; Kuemmerle, T.; Meyfroidt, P.; Mitchard, E.T.A.; et al. A Review of the Application of Optical and Radar Remote Sensing Data Fusion to Land Use Mapping and Monitoring. *Remote Sens.* **2016**, *8*, 70. [CrossRef]
72. Flores-Anderson, A.I.; Herndon, K.E.; Thapa, R.B.; Cherrington, E. *The SAR Handbook: Comprehensive Methodologies for Forest Monitoring and Biomass Estimation*; NASA: Washington, DC, USA, 2019.
73. Kouw, W.M.; Loog, M. A Review of Domain Adaptation without Target Labels. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *43*, 766–785. [CrossRef] [PubMed]
74. Farahani, A.; Voghoei, S.; Rasheed, K.; Arabnia, H.R. A Brief Review of Domain Adaptation. *Proc. Adv. Data Sci. Inf. Eng.* **2021**, *2021*, 877–894.
75. Zadrozny, B. Learning and evaluating classifiers under sample selection bias. In Proceedings of the Twenty-First International Conference on Machine Learning ICML, Banff, AB, Canada, 4–8 July 2004; Volume 2004. [CrossRef]
76. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2016; pp. 770–778.