

Article

# Direct, ECOC, ND and END Frameworks—Which One Is the Best? An Empirical Study of Sentinel-2A MSIL1C Image Classification for Arid-Land Vegetation Mapping in the Ili River Delta, Kazakhstan

Alim Samat <sup>1,2,3,\*</sup> , Naoto Yokoya <sup>4</sup> , Peijun Du <sup>5</sup>, Sicong Liu <sup>6</sup> , Long Ma <sup>1,2,3</sup> ,  
Yongxiao Ge <sup>1,2</sup>, Gulnura Issanova <sup>7,8</sup>, Abdula Saparov <sup>9</sup>, Jilili Abuduwaili <sup>1,2,3</sup> and Cong Lin <sup>5</sup>

<sup>1</sup> State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, CAS, Urumqi 830011, China

<sup>2</sup> Research Center for Ecology and Environment of Central Asia, CAS, Urumqi 830011, China

<sup>3</sup> University of Chinese Academy of Sciences, Beijing 100049, China

<sup>4</sup> RIKEN Center for Advanced Intelligence Project, Tokyo 103-0027, Japan

<sup>5</sup> Department of Geographical Information Science, Nanjing University, Nanjing 210093, China

<sup>6</sup> College of Surveying and Geoinformatics, Tongji University, Shanghai 200092, China

<sup>7</sup> Faculty of Geography and Environmental Sciences, Al-Farabi Kazakh National University, Almaty 050040, Kazakhstan

<sup>8</sup> Research Center for Ecology and Environment of Central Asia (Almaty), CAS, Almaty 050060, Kazakhstan

<sup>9</sup> U.U. Uspanov Kazakh Research Institute of Soil Science and Agrochemistry, Almaty 050060, Kazakhstan

\* Correspondence: alim\_smt@ms.xjb.ac.cn; Tel: +86-0991-782-7371

Received: 30 July 2019; Accepted: 16 August 2019; Published: 20 August 2019



**Abstract:** To facilitate the advances in Sentinel-2A products for land cover from Moderate Resolution Imaging Spectroradiometer (MODIS) and Landsat imagery, Sentinel-2A MultiSpectral Instrument Level-1C (MSIL1C) images are investigated for large-scale vegetation mapping in an arid land environment that is located in the Ili River delta, Kazakhstan. For accurate classification purposes, multi-resolution segmentation (MRS) based extended object-guided morphological profiles (EOMPs) are proposed and then compared with conventional morphological profiles (MPs), MPs with partial reconstruction (MPPR), object-guided MPs (OMPs), OMPs with mean values (OMPsm), and object-oriented (OO)-based image classification techniques. Popular classifiers, such as C4.5, an extremely randomized decision tree (ERDT), random forest (RaF), rotation forest (RoF), classification via random forest regression (CVRFR), ExtraTrees, and radial basis function (RBF) kernel-based support vector machines (SVMs) are adopted to answer the question of whether nested dichotomies (ND) and ensembles of ND (END) are truly superior to direct and error-correcting output code (ECOC) multiclass classification frameworks. Finally, based on the results, the following conclusions are drawn: 1) the superior performance of OO-based techniques over MPs, MPPR, OMPs, and OMPsm is clear for Sentinel-2A MSIL1C image classification, while the best results are achieved by the proposed EOMPs; 2) the superior performance of ND, ND with class balancing (NDCB), ND with data balancing (NDDB), ND with random-pair selection (NDRPS), and ND with further centroid (NDFC) over direct and ECOC frameworks is not confirmed, especially in the cases of using weak classifiers for low-dimensional datasets; 3) from computationally efficient, high accuracy, redundant to data dimensionality and easy of implementations points of view, END, ENDCB, ENDDDB, and ENDRPS are alternative choices to direct and ECOC frameworks; 4) surprisingly, because in the ensemble learning (EL) theorem, “weaker” classifiers (ERDT here) always have a better chance of reaching the trade-off between diversity and accuracy than “stronger” classifiers (RaF, ExtraTrees, and SVM here), END with ERDT (END-ERDT) achieves the best performance with less than a 0.5% difference in the overall accuracy (OA) values, but is 100 to 10000 times faster than END with RaF and ExtraTrees, and ECOC with SVM while using different datasets with various dimensions; and, 5) Sentinel-2A

MSIL1C is better choice than the land cover products from MODIS and Landsat imagery for vegetation species mapping in an arid land environment, where the vegetation species are critically important, but sparsely distributed.

**Keywords:** ND; END; ECOC; MRS; Extended object-guided morphological profiles; Multiclass classification; Arid-land vegetation mapping; Sentinel-2A MSIL1C; Central Asia

---

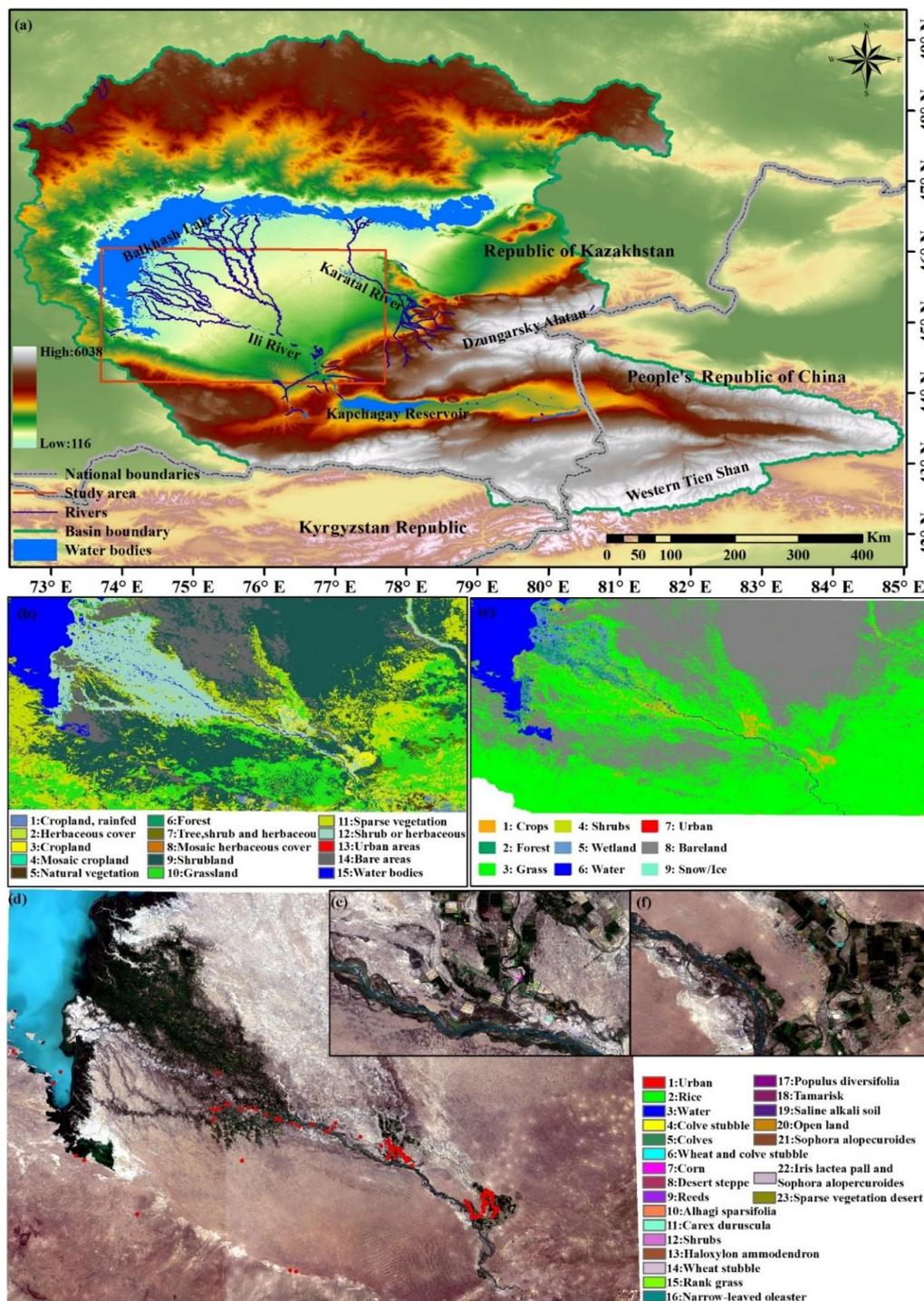
## 1. Introduction

Arid and semiarid lands encompass approximately 30–40% of the Earth's surface, and Central Asia contains one of the world's largest arid and semiarid areas. These areas have harsh climatic conditions and they are under high pressure to produce food and fibers for their rapidly increasing populations, which include a wide range of land utilization and management regimes, which results in a reduction in arid ecosystem quality. Understanding the effects and responses between landscapes and regional environments is fundamental to maintain their ecological and productive value in such circumstances. Hence, the effects and responses of landscape heterogeneity on the local and regional atmosphere, the surface energy balance, the carbon exchange, and climate changes are major topics that have attracted widespread interest [1–5]. Among these responses, the vegetation species, distribution, diversity, and biomass in these lands typically undergo wide seasonal and international fluctuations, which are largely regulated by water availability and impacted by both climatic shifts and human activities [6–8]. Thus, monitoring the vegetation status of these lands is an essential part of identifying problems, developing solutions, and assessing the effects of actions.

However, large-scale and long-term field sampling of vegetation information is challenging when considering the sampling efforts and costs. Moreover, the samples are often very sparsely distributed, and site revisits remain infrequent, while the success of any monitoring of vegetation dynamics depends on the availability of up-to-date and spatially detailed species richness and distribution at a regional scale [9–11]. Fortunately, satellite-based remote sensing (RS) data can address these challenges by identifying and detailing the biophysical characteristics of vegetation species' habitats, predicting the distribution and spatial variability in the richness, and detecting natural and human-caused changes at scales that range from individual landscapes to the whole world [1,9,12–18]. Therefore, an increasing number of geologists, ecologists, and biologists are turning to rapidly develop RS data sources for vegetation-based ecological and environmental research at local, regional, and global scales [19–24].

Regarding applications of RS data in vegetation studies, high- and moderate-resolution optical RS sensors, including IKONOS, Satellite for Observation of Earth (SPOT), Thematic Mapper™, Enhanced Thematic Mapper (ETM), ETM+, Operational Land Imager (OLI), Sentinel-2, and Moderate Resolution Imaging Spectroradiometer (MODIS), are widely accepted and are considered to be adequate for vegetation species, diversity, distribution, and biophysical information extraction in different settings [25–32]. Creating land cover maps that detail the biophysical cover of the Earth's surface is among the oldest and ongoing hot applications. Land cover maps have been continuously suggested proven especially valuable for predicting the distribution of both individual plant species and species assemblages across broad areas that could not otherwise be surveyed in more quantitative ways with respect to vegetation index (VI)-based approaches [9,33–35]. In particular, after various land cover products that are derived from RS data at the regional and global scales have been produced, and they are freely available at spatial resolutions from 30 m to 1 km. Solid proofs can be found for extensive applications of representative products, including the 1 km University of Maryland (UMD) land cover layer [36], the Global Land Cover 2000 (GLC2000) products [37], the MODIS products [38], the 500 m MODIS [39], the 300 m GlobCover [40], and the 30 m global land cover maps [41] for vegetation studies at the regional and global scales [42–49]. However, most of the existing land cover products are coarse, not only in the spatial resolution and land cover type details, but also in the

update frequency. For example, the 30 m global land cover maps are only available for 2010, 2015, and 2017 with a maximum of 10 land cover types (only eight types for our study area), while the MODIS products are only available for 2000, 2005, 2010, and 2015 with a maximum 300 m resolution (only 15 types for our study area). Furthermore, the differences between these products are very large, as shown in Figures 1b and 1c.



**Figure 1.** (a) Geographic location of the study area, (b) 2015 Moderate Resolution Imaging Spectroradiometer land use and cover change (MODIS LUCC), (c) 2017 GLC30, (d) blue rectangle, (e) blue rectangle, and (f) green rectangle. Sentinel-2 RGB image of the study area with in situ points (red dots) and the corresponding land cover types.

The Sentinel-2 mission comprises a constellation of two polar-orbiting satellites placed in the same orbit with a five-day revisit time over land and coastal areas. Each satellite carries a MultiSpectral Instrument (MSI) with 13 bands spanning from the visible and the near-infrared (VNIR) portion of the electromagnetic spectrum to the short wave infrared (SWIR) portion of the spectrum and features four bands at a 10 m spatial resolution, six bands at a 20 m spatial resolution, and three bands at a 60 m spatial resolution with a 290 km field of view [50]. Since Sentinel-2A and Sentinel-2B were successfully launched on June 23, 2015, and March 7, 2017, respectively, their products (Level-2A, which covers the bottom-of-atmosphere reflectance in cartographic geometry; Level-1B, which covers the top-of-atmosphere (TOA) radiance in sensor geometry; and, Level-1C, which covers the TOA reflectance in fixed cartographic geometry) have been widely applied for monitoring land cover changes, agricultural applications, monitoring vegetation and retrieving biophysical parameters, observing coastal zones, monitoring inland water, monitoring glaciers and ice and mapping snow, mapping floods, and management [51–56]. However, these products have not been used for detailed vegetation mapping in arid land environments. Hence, the first objective of this paper is to investigate the performance of the Sentinel-2 MSIL1C product for vegetation mapping in an arid land environment.

Producing any substantial land cover/use maps always requires a specific classification method or ML algorithm. Although many methods and algorithms have been developed for satellite data classification applications, the search for advanced classification methods or algorithms is still a hot filed [57–60]. There are no supper classification methods or algorithms that could universally work at high performance, due to facts that classification performance not only controlled by robustness of adopted methods or algorithms, but also affected by discrimination and identification quality, size and distribution quantity of provided data [61,62]. According to the literatures from RS data classification community, the commonly used ML algorithms are artificial neural networks (ANNs) [63], support vector machine (SVM) [64,65], extreme learning machine (ELM) [66], decision trees (DTs) [67], ensemble methods, such as bagging, adaboost, and RaF [57,68,69], and deep neural networks (DNNs) [70,71]. In most scenarios, these algorithms involve a nominal class variable that has more than two values problem, because the real-world land surface usually recorded by EO sensors with simultaneous discrimination of numerous classes. In general, there are two approaches for addressing this type of problem: 1) adapting the binary algorithm to its multiclass counterpart to deal with multiclass problems directly (e.g., DTs, ANNs, ELMs, RaFs); and, 2) reduce the multiclass problem into several binary subproblems first, then form a multiclass prediction based on the results from several binary classifiers, such as AdaBoost, multiclass SVMs, and ND [72–74]. When compared with the direct approaches, the latter approach is appealing, because it does not involve any changes to the underlying binary algorithm [75]. In particular, a structural risk minimization (SRM) rule-based SVM can successfully work with limited quantities and quality of training samples and it often achieves a higher classification accuracy than linear discriminate analysis (LDA), DTs, ANNs, bagging, AdaBoost, and RaFs [64–66,76–78].

Well-known examples of the second approach are ECOC and pairwise classification, which often result in significant increases in the accuracy [72,75,79]. However, many studies have explicitly proven that ECOC works better than pairwise classification mainly due to its more advanced decoding strategies [80–82]. The ECOC framework consists of two steps: coding and decoding. Popular coding strategies include one-versus-all, one-versus-one, random sparse, binary complete, ternary complete, original and dense random coding, while the most frequently applied decoding designs are Hamming decoding, inverse Hamming decoding, Euclidean decoding, attenuated Euclidean decoding, loss-based decoding, probabilistic decoding,  $\beta$ -density-based distribution decoding, and loss-weighted decoding [72,82]. While the one-versus-one and one-versus-all strategies have been widely adopted in RS data classification, only a few works [83,84] have focused on applications of the ND and its ensemble variants, which have been proven to outperform the direct multiclass, ECOC, and pairwise classification methods while using C4.5 and logistic regression as the base learners [75,85,86]. Additionally, the most recent and more advanced direct multiclass classification methods may also see improved accuracy

by interacting with ND and END. Thus, the second objective of this paper is to investigate the performance of the popular ND algorithms, including NDCB, NDDB, NDRPS, NDFC, and their ensemble versions (i.e., ENDCB, ENDDDB, ENDRPS, and ENDBC, respectively) by setting C4.5 and bagging [87], AdaBoost [88], an RaF [89], a RoF [90], ExtraTrees [91], and an SVM [92] as the base learners.

The discrimination and identification quality of the provided data is another critical factor that controls the classification performance of the adopted classifier. Over the years, many approaches have been proposed to increase the discrimination and identification ability of the provided data. Among these approaches, structural filtering, MPs, random fields, object-based image analysis (OBIA) and geographic OBIA (GEOBIA), sparse representation (SR), and deep learning (DL)-based contextual information extraction are the most undertaken families of methods [77,93–96]. In the last ten years, mathematical morphology (MM)-based operators, such as MPs, EMPs, APs, and MPs with partial reconstruction (MPPR), have been the most widely accepted approaches in the RS image classification community, mainly due to their advanced and proven performances in contextual information extraction from HR/VHR RS imagery [68,77,93,96,97].

However, the SE sequences or attribute filters (AFs) that are necessarily adopted in the above operators always result in computationally inefficient and redundant high-dimensional features, which may become prohibitively large data processing cases. Additionally, the sequences of SE and AFs, with limited sizes and shapes, cannot match all of the sizes and shapes of the objects in an image; specifically, a single SE is not suitable for an entire image in each operation case [97,98]. MSER-MPs, SPMPs, and multi-resolution segmentation (MRS)-OMP have been proposed for the spectral-spatial classification of VHR multi/hyperspectral images with the ExtraTrees, ForestPA, and ensemble extremely randomized decision trees (EERTDs) ensemble classifiers in our previous works to overcome such challenges [98,99]. MSER\_MPs(M), SPMPs(M), and OMPs(M) were also proposed by considering the mean pixel values within regions, such as MSER objects, superpixels, and MRS objects, to foster effective and efficient spatial FE. The improvements from taking the mean values were clear. Specifically, the size of the regions generated was on a reasonable scale, which was mainly controlled by the spatial resolution and a readily available landscape image [99]. However, as hybrid methods of MPs and OBIA, comparison studies between SPMPs and OBIA, and between OMPs and OBIA were not carried out in our previous works. In addition, as generally known from the OBIA and GEOBIA communities, there are plenty of spectral, statistical, spatial, and geometrical measures of regions (i.e., objects) that can be adopted to further improve the classification accuracy [27,100–102]. Thus, extending the OMPs by considering more advanced object measures is interesting, especially when using Sentinel-2A MSIL1C data for vegetation mapping in large coverage areas in arid land environments, which is the last objective of this paper. In Table 1, we provide the acronym with corresponding full names that are used in this paper.

**Table 1.** Acronyms with corresponding full names used in this paper.

Acronyms	Full Name	Acronyms	Full Name
AA	Average accuracy	MSIL1C	MultiSpectral Instrument Level-1C
AFs	Attribute filters	MSER-MPs	Maximally stable extremal region-guided MPs
ANNs	Artificial neural networks	ND	Nested dichotomies
AVHRR	Advanced VHR Radiometer	NDBC	ND based on clustering
CBR	Closing by reconstruction	NDCB	ND with class balancing
CVRFR	Classification via RaF regression	NDDB	ND with data balancing
DL	Deep learning	NDFC	ND with further centroid
DNNs	Deep neural networks	NDRPS	ND with random-pair selection

Table 1. Cont.

Acronyms	Full Name	Acronyms	Full Name
DTs	Decision trees	OA	Overall accuracy
ECOC	Error-correcting output code	OBIA	Object-based image analysis
EERDTs	Ensemble of ERDTs	OBR	Opening by reconstruction
EL	Ensemble learning	OBPR	Opening by partial reconstruction
ELM	Extreme learning machine	OLI	Operational Land Imager
END	Ensembles of ND	OMPs	Object-guided MPs
ENDBC	Ensemble of NDBC	OMPsM	OMPs with mean values
ENDCB	Ensemble of NDCB	OO	Object-oriented
ENDDB	Ensemble of NDDB	OOBR	Object guided OBR
ENDRPS	Ensemble of NDRPS	PCA	Principal component analysis
END-ERDT	END with ERDT	RaF	Random forest
EOMPs	Extended object-guided MPs	RBF	Radial basis function
ERDT	Extremely randomized DT	ROI	Region of interest
ESA	European Space Agency	RoF	Rotation forest
ETM	Enhanced Thematic Mapper	SE	Structural element
ExtraTrees	Extremely randomized trees	SEOM	ESA's Scientific Exploration of Operational Missions
EVI	Enhanced vegetation index	SNAP	Sentinel Application Platform
GEOBIA	Geographic OBIA	SPOT	Satellite for Observation of Earth
GPS	Global positioning system	SR	Sparse representation
HR	High resolution	SRM	Structural risk minimization
LDA	Linear discriminate analysis	SVM	Support vector machine
LR	Logistic regression	SVM-B	SVM with Bayes optimization
ML	Machine learning	SVM-G	SVM with grid-search optimization
MM	Mathematical morphology	SWIR	Short wave infrared
MPs	Morphological profiles	UA	User accuracy
MPPR	MPs with partial reconstruction	UMD	University of Maryland
MRFs	Markov random fields	TOA	Top-of-atmosphere
MRS	Multi-resolution segmentation	VHR	Very high resolution
MODIS	Moderate Resolution Imaging Spectroradiometer	VI	Vegetation index
MSI	MultiSpectral Instrument	VNIR	Visible and the near-infrared

## 2. Materials and Methods

### 2.1. Materials

#### 2.1.1. Study Region

Our study area is located at the Ili River delta, in the central-western part of the Balkhash Lake basin, in the southeastern part of Kazakhstan (Figure 1a). The Balkhash Lake basin is one of the largest internal drainage areas in the arid and semiarid region in Central Asia; it is located between 72.44°–84.99°E and 42.24°–49.14°N, covering an area of approximately 500,000 km<sup>2</sup> and it is shared by

the Republic of Kazakhstan (approximately 60%) and the People's Republic of China (approximately 40%) [103]. Balkhash Lake is the world's fifth-largest inland water reservoir (605 km long and 4–74 km wide), with a volume of 87.7 km<sup>3</sup> and a catchment area of 15,730 km<sup>2</sup> [104]. All of the inflow to the Balkhash Lake is received from the western Tien-Shan and the Dzungarsky Alatau and the runoff from their ridges. The two largest rivers flowing into the lake are the Ili River and Karatal River, accounting for approximately 78% and 15% of the total inflow, respectively [105]. Balkhash Lake and several plentiful wetlands in its inflow deltas are considered to be very sensitive ecosystems, whose existence depends on variable climate conditions and extensive human activities, especially in the form of water abstractions from inflows. During the Soviet era, the inflow waters were largely used for irrigation (mainly for rice crops), industry, the water supply to populated areas, and the fishing industry, which resulted in a significant decrease in the water level and the degradation of the surrounding environments [103,105]. After the collapse of the Soviet Union, most of the social and economic activities in the Balkhash Lake basin rapidly diminished, which causes drastic changes in the land cover/use and broad rehabilitation of the ecosystem. Understanding the effects and responses between such drastic changes and the regional environment is crucial for the sustainable development of this basin, which can only be accomplished by the sustainable monitoring of entire environments. Many efforts have been made in recent decades; however, while most studies have focused on water, e.g., water resource management, water level and surface changes, chemical properties, regional-scale land cover/use changes, ecosystem services, and vegetation activity [106–111], only a few studies have focused on basin-level studies while using RS datasets [104]. In almost all of the above studies, datasets from Landsat, the Advanced Very-High-Resolution Radiometer (AVHRR), and MODIS were mainly used. Hence, it is of interest to use more advanced Sentinel-2A MSIL1C products with more advanced spatial FE and ML techniques for vegetation mapping in this area.

### 2.1.2. Datasets

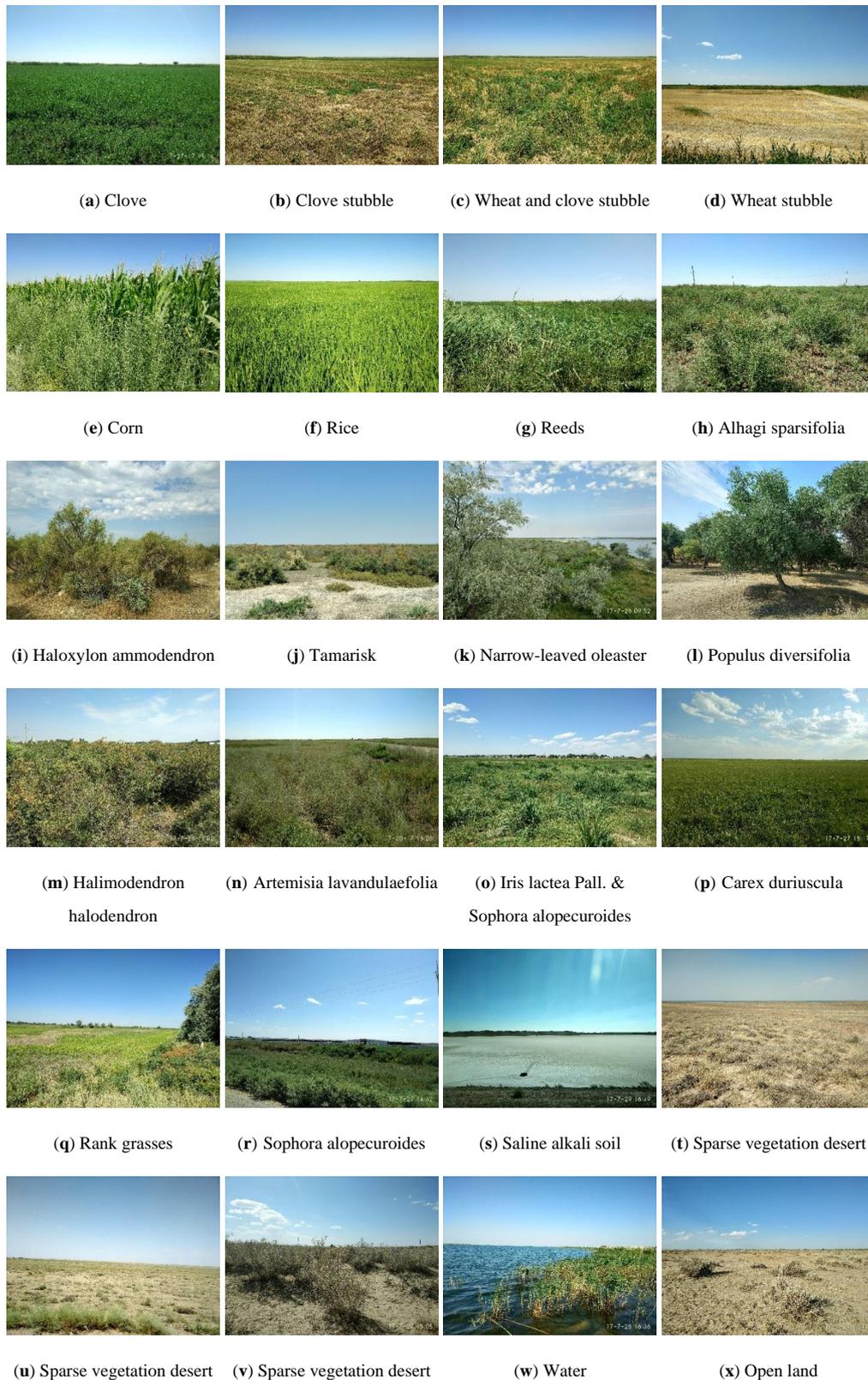
- Sentinel-2 data collection and preprocessing

In this study, Sentinel-2A geolocated TOA reflectance (L1C) products were acquired from the Copernicus Open Access Hub (<https://scihub.copernicus.eu>). We selected a total of six images with zero or near-zero (< 10%) cloud coverage, taken between 25 July and 8 August, 2017. Only the visible bands of blue (band 2), green (band 3), red (band 4), and the near-infrared (band 8) region with a 10 m spatial resolution were used. All of the images of the study region were projected to WGS 84/UTM zone 43N and then mosaicked while using the SNAP (v6.0), which is a free, open source software program that is distributed by the ESA under the GNU General Public License and was founded through the ESA's Scientific Exploration of Operational Missions (SEOM) Program. In Figure 1d, Figure 1e, and Figure 1f, true RGB color images were composited by setting band 4 to red, band 3 to green, and band 2 to blue, respectively, for real-world land surface illustration purposes.

- In situ data collection

In total, 120 valid in situ sites (red dots in Figure 1d) were visited on July 27, 28, 29, and 30, 2017. Specifically, 46 field sites were visited on July 27 in the Bakanas irrigation area (Figure 1e), 23 field sites were visited on July 28 in the Ili River delta region, six field sites were visited on July 29 on the way back from Balkhash city to Bakanas District, and 45 field sites were visited on July 30 in the Bakkakty irrigation area (Figure 1f). For all of the field sites, the coordinates were determined while using a differential global positioning system (GPS) and the Chinese BeiDou navigation system, which has a 2 m positioning accuracy. Additionally, the land cover type among 23 possibilities with 19 vegetation types was recorded (Figure 1), between 10 AM and 6 PM local time. Moreover, field sites are determined at locations with only large and uniform spatial coverage of the same land cover type for a more objective and representative in situ site selection. Figure 2 shows the ground photos of representative land cover types in our study area. According to the collected in situ information and

further referring to the high-resolution optical images in Google Earth, 582 regions of interest (ROIs) were selected for model training, validation, and data classification for vegetation mapping. Detailed ROI, training sample, and validation sample information are listed in Table 2.



**Figure 2.** Ground photos of the land cover types.

**Table 2.** Details of the land cover types in the training and validation samples in the test datasets.

LC Types	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	Total
ROIs	253	39	30	17	30	13	6	7	16	1	4	20	9	33	7	6	2	5	19	15	5	8	37	582
Train	147	230	295	144	325	225	128	19	304	12	92	164	42	277	117	45	27	75	976	171	30	238	475	4558
Test	2794	4366	5600	2726	6180	4269	2428	352	5768	218	1753	3113	794	5256	2226	852	516	1429	18542	3246	560	4523	9021	86532
Total	2941	4596	5895	2870	6505	4494	2556	371	6072	230	1845	3277	836	5533	2343	897	543	1504	19518	3417	590	4761	9496	91090

2.2. Methods

2.2.1. Related Methods

- Ensemble of nested dichotomies

In the area of statistics, ND are a standard technique for solving certain multiclass classification problems with logistic regression (LR). Generally, ND can be represented with a binary tree structure, where the set of classes is recursively split into two subsets until there is only one (Figure 3). In other words, the root node of the ND contains all of the classes that correspond to the multiclass classification problem, and each node contains a single class, which means that, for an  $n$ -class problem, there are  $n$  leaf nodes and  $n-1$  internal nodes. To build an ND approach based on such a tree structure, we perform the following steps: 1) at each internal node, store the instances pertaining to the classes associated with current node but no other instances; 2) group the classes pertaining to each node into two subsets to ensure that each subset holds the classes that are associated with exactly one of the node’s two successor nodes; and, 3) train the binary classifier at each node for the resulting binary class problem [75,85]. If the adopted binary classifier at each node can compute the class probability, the ND can compute class probability in a natural way, which is a convenient feature in real-world applications [112].

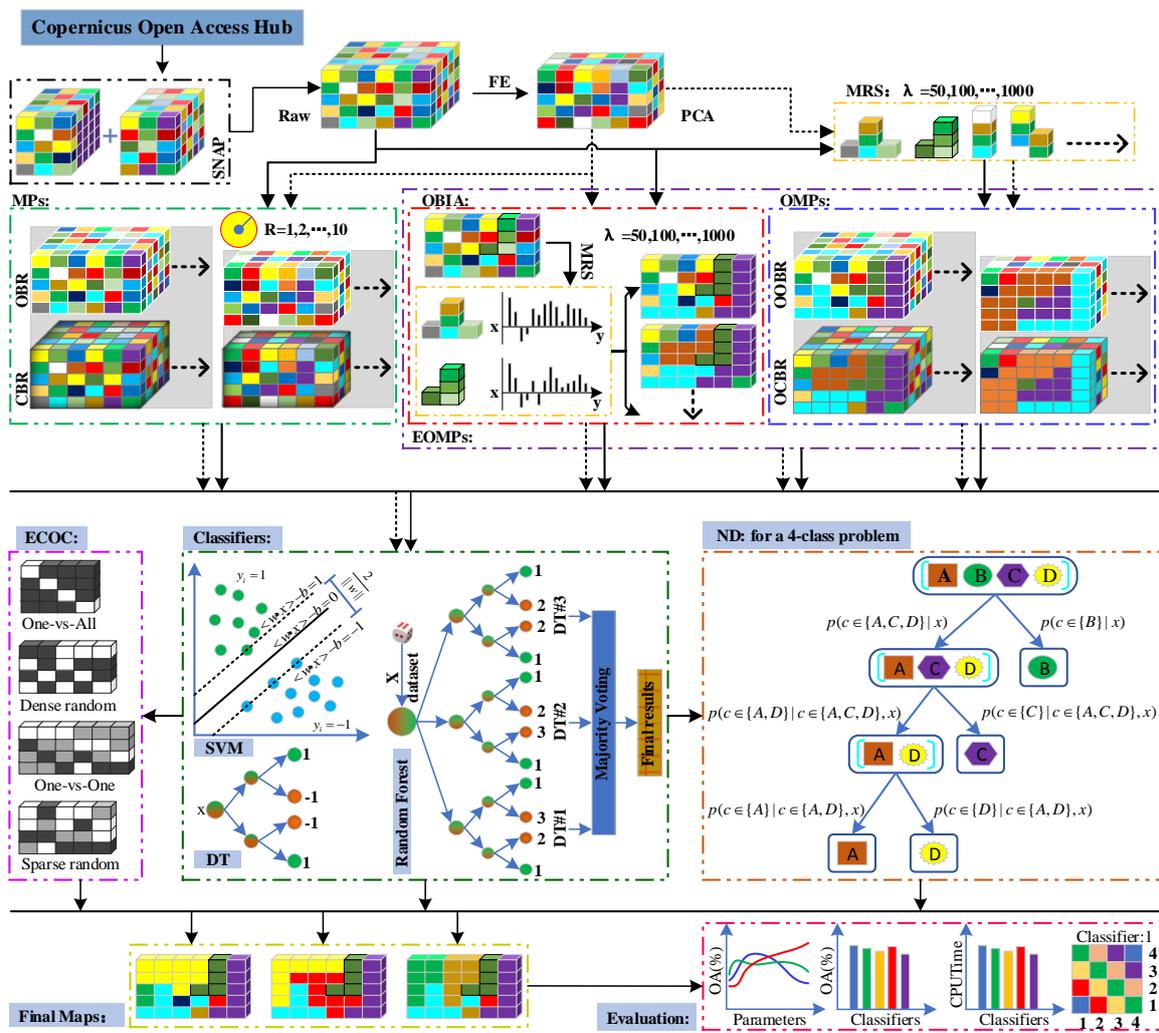


Figure 3. Overall technical flowchart for the methodology.

After an ND approach is built, one critical question is how to combine the probability estimates from individual binary problems to obtain class probability estimates for the original multiclass problem. Multiclass probability estimates can be obtained by simply multiplying the probability estimates returned from the individual binary learners because the individual dichotomies are statistically independent as they are nested. More specifically, let  $C_{i1}$  and  $C_{i2}$  be the two subsets of classes generated by a split of the set of classes  $C_i$  at internal node  $i$  of the tree structure, and let  $p(c \in C_{i1}|x, c \in C_i)$  and  $p(c \in C_{i2}|x, c \in C_i)$  be the conditional probability distributions that are estimated by the binary learner at node  $i$  for a given instance  $x$ . Subsequently, we can have the estimated class probably distribution for the original multiclass problem by [85]:

$$p(c = C|x) = \prod_{i=1}^{n-1} (I(c \in C_{i1})p(c \in C_{i1}|x, c \in C_i) + I(c \in C_{i2})p(c \in C_{i2}|x, c \in C_i)) \tag{1}$$

where  $I(\cdot)$  is the indicator function and the product is over all the internal nodes. Notably, not all of the nodes must be examined to compute the probability for a particular class value, which makes the evaluation of the path to the leaf associated with that class sufficient.

Ever since the basic form of the class subset split criterion was originally proposed by Frank and Kramer [85], many other sophisticated criteria, such as random selection, random-pair selection, clustering, multisubset evaluation, class-balanced-based optimization, data-balanced-based optimization, and genetic algorithm-based optimization, have been proposed and proven to have superior performance on the classification accuracy and model training efficiency, especially with END, which use common EL algorithms such as bagging, boosting, and RaFs [74,75,85,86,112–114].

According to the formation by Frank and Kramer [85], there are  $T(c) = (3^c - (2^{c+1} - 1))/2$  possible dichotomies for a  $c$ -class problem, which is very large and not ideal for efficient model training. Especially when large amounts of data are readily available, advanced, but computationally inefficient learners (e.g., ANNs and SVMs) are adopted in an ensemble scenario. One simple solution is using random selection dichotomies instead of complete selection dichotomies, which reduces the number of possible dichotomies to  $T(c) = (2c - 3) \times T(c - 1)$ , where  $T(1) = 1$ . Briefly, all the distinct dichotomies for a given  $n$ -class problem were uniformly sampled with replacement, and the class probability estimates for a given instance  $x$  were obtained by averaging the estimates from the individual END members. According to statistical theory regarding EL, reduced numbers of dichotomies are still large enough to ensure that there is a high level of diversity among END members to facilitate the improvement by the ensemble. One drawback of random selection is that it can produce very imbalanced tree structures, which results in a negative effect on the training time of the full model while the number of internal nodes remains the same in any ND for the same number of classes because an unbalanced tree often implies that the internal binary learners are trained on large datasets near the leaves. Dong et al. [75] proposed class-balanced and data-balanced versions of ND, namely, the NDCB and NDDB, respectively, to mitigate the effect of this issue. When compared with NDCB, NDDB can avoid the potential problem from multiclass problems with imbalanced samples. Empirical experiments have shown that NDCB and NDDB have little effect on the accuracy in most cases, but they have great benefits in reducing the time needed for model training, particularly for problems with many classes in ensemble NDCB and NDDB (ENDCB and ENDDB, respectively). The growth function for NDCB is [75]:

$$T(c) = \begin{cases} \frac{1}{2} \binom{c}{c/2} T(\frac{c}{2})^2, & \text{if } c \text{ is even} \\ \binom{c}{(c+1)/2} T(\frac{c+1}{2}) T(\frac{c-1}{2}), & \text{if } c \text{ is odd} \end{cases} \tag{2}$$

where  $T(2) = T(1) = 1$ .

When constructing an effective EL system, one is always faced with a dilemma between high classifier diversity and excellent performance, which is hard to satisfy in practice. Unfortunately, the above END methods are deterministic when generating subclass groups that cannot maintain the benefits of high diversity. Additionally, errors that are made by binary classifiers at earlier nodes can be inherently spread to lower nodes of the ND tree and they cannot be easily corrected. For these reasons, it is important to generate the dichotomies in a nondeterministic way to reach the high diversity requirement on the one hand and reduce the number of errors in and the size of the upper nodes of the ND tree on the other hand. NDBC, NDRPS, and their ensemble versions are good examples for this intension [74,112]. However, as compared with NDCB, NDRPS is more direct, easily discovers similar classes, and exhibits a degree of randomness, which leads to more diversity and a higher-performing ensemble. The growth function of the NDRPS was empirically estimated by [112]:

$$T(c) = p(c)T\left(\frac{c}{3}\right)T\left(\frac{2c}{3}\right) \quad (3)$$

where  $p(c)$  represents the size of the dichotomies from the base learner and  $T(2) = T(1) = 1$ .

- Multiresolution segmentation

OBIA is a classic technique in RS image interpretation that integrates the spatial and spectral features and it splits RS images into a set of nonoverlapping homogeneous regions or objects, depending on the segmentation method that was specified. During recent decades, OBIA has gained widespread attention in the RS community, mainly because it can overcome the limitations of pixelwise analysis, such as the neglect of geometric, contextual, and semantic information, particularly in the processing of HR/VHR RS imagery [100–102,115]. Over the years, many image segmentation methods have been proposed and extensively examined while using various RS imagery. Among these methods, MRS is one of the most frequently used methods, which is mainly due to its capability to produce high-quality segments at different scales [69,102,116,117].

MRS is a bottom-up region-merging-based segmentation technique that starts with one-pixel objects and it merges the most similar adjacent pixels or objects provided that the internal heterogeneity of the resulting object does not exceed a user-defined threshold [118]. The heterogeneity measure in eCognition considers the spectral heterogeneity, which allows for multivariant segmentation by adding weights to the image channels, and the shape heterogeneity, which describes the improvement in the shape with regard to the smoothness and compactness. In any OBIA, the segmentation scale determines the average size and number of segments that were produced. Defining an optimal scale segmentation to avoid oversegmentation and undersegmentation issues is always challenging because of the spectral similarity between different objects and landscape complexity in the real world [101,102,119]. For MRS, numerous studies demonstrate the importance of the scale parameter, because it controls the dimensions and the size of the segments, which may directly affect subsequent results [101,117,119]. A successful research result on scale optimization is to combine the local variance (LV) and the rates of change of the LV (ROC-LV) to determine the appropriate segmentation scales for a given resolution [120]. The automated selection of scale parameters is basically an automation of the ESP tool, where the production of a graph is replaced by an iterative procedure that segments an image at the first threshold that occurs in the LV graph. The readers are referred to the original works by [118] and [120] for more detailed information.

### 2.2.2. Proposed Method

MPs are composed of morphological opening and closing profiles, which consist of an ensemble of OBR and CBR operators. According to the definition of MPs, OBR and CBR operators are connected operators that satisfy the assertion of removing the structures that cannot contain the SE and preserving those structures that can contain the SE [121–123]. While applying such operators with a sequence of SEs of increasing size, one can extract information regarding the contrast and the size of the geometrical

structures that are present in the image. Originally, the formulation of the spatial information that was included in the MPs refers to a single-band image; therefore, the direct construction of MPs is not straightforward for multi/hyperspectral images. Several approaches have been considered to overcome this shortcoming [77,93,94,124,125]. Among these approaches, one simple, and yet efficient, approach is to use a few images that contain most of the spectral information that was obtained by some FE method, namely, the EMPs [121]. If we consider the first  $m$  principal components that were extracted from the multi/hyperspectral images with principal component analysis (PCA), the EMPs are obtained by stacking all of the MPs that are built on all  $m$  components.

According to the definition from MM and our previous works [98,99], the MRS object-guided morphological OBR operators can be obtained by first eroding the input image while using segmented objects (where  $\Theta_S^\lambda$  represents the numbers ( $S$ ) of objects from MRS with scale  $\lambda$ ) in the SE approach and by using the result as a marker in geodesic reconstruction by a dilation phase:

$$OOBR(f) = R_f^D \left[ f \odot (\exists \Theta_{j,j \in S}^\lambda \in \Theta_S^\lambda) \right] \tag{4}$$

Similarly, we have

$$OCBR(f) = R_f^E \left[ f \oplus (\exists \Theta_{j,j \in S}^\lambda \in \Theta_S^\lambda) \right] \tag{5}$$

where the object-guided CBR (OCBR), which was obtained by complementing the image  $f^C$ , contains the object-guided OBR (OOBR) with SEs  $\exists \Theta_{j,j \in S}^\lambda$  and it complements the resulting procedure:

$$OCBR(f) = R_f^{DC} \left[ f^C \odot (\exists \Theta_{j,j \in S}^\lambda \in \Theta_S^\lambda) \right] \tag{6}$$

In MM, the erosion of  $f$  by  $b$  at any location  $(x, y)$  is defined as the minimum value of all the pixels in its neighborhood, denoted by  $b$ . In contrast, dilation returns the maximum value of the image in the window that was outlined by  $b$ . Subsequently, we can have the following new formations for the erosion and dilation operators:

$$\begin{aligned} \left[ f \odot (\exists \Theta_{j,j \in S}^\lambda \in \Theta_S^\lambda) \right] (x, y) &= \min_{(s,t) \in \Theta_{j,j \in S}^\lambda} \{f(x + s, y + t)\} \\ \left[ f \oplus (\exists \Theta_{j,j \in S}^\lambda \in \Theta_S^\lambda) \right] (x, y) &= \max_{(s,t) \in \Theta_{j,j \in S}^\lambda} \{f(x + s, y + t)\} \end{aligned} \tag{7}$$

By substituting Equation (13) into Equations (10) and (12), we have the formations of the OOBR and the OCBR as:

$$\begin{aligned} OOBR(f) &= R_f^D \left[ \min_{(s,t) \in \Theta_{j,j \in S}^\lambda} \{f(x + s, y + t)\} \right] \\ OCBR(f) &= R_f^E \left[ \max_{(s,t) \in \Theta_{j,j \in S}^\lambda} \{f(x + s, y + t)\} \right] = R_f^{DC} \left[ \min_{(s,t) \in \Theta_{j,j \in S}^\lambda} \{f^C(x + s, y + t)\} \right] \end{aligned} \tag{8}$$

If the SEs  $\exists \Theta_{j,j \in S}^\lambda$  are specified by MRS objects with a sequence of scale parameter  $\lambda$ , then the MRS object guided morphological profiles (OMPs) of an image  $f$  can be defined as:

$$OMP_S(f) = \left[ OOBR(f)^{(\exists \lambda \in \{\lambda_1^*, \lambda_2^*, \dots, \lambda_Q^*\})}, OCBR(f)^{(\exists \lambda \in \{\lambda_1^*, \lambda_2^*, \dots, \lambda_Q^*\})} \right] \tag{9}$$

where  $\{\lambda_1^*, \lambda_2^*, \dots, \lambda_Q^*\}$  represents the sets of  $Q$  numbers of the user-specified scale parameter  $\lambda$ .

By further considering the extensively proven performance from object profiles in OO-based image classification, the extended OMPs (EOMPs) can be calculated, as follows:

$$EOMPs(f) = \left[ OOBR(f)^{(\exists \lambda \in \{\lambda_1^*, \lambda_2^*, \dots, \lambda_Q^*\})}, OCBR(f)^{(\exists \lambda \in \{\lambda_1^*, \lambda_2^*, \dots, \lambda_Q^*\})}, (f)_{OO}^{(\exists \lambda \in \{\lambda_1^*, \lambda_2^*, \dots, \lambda_Q^*\})} \right] \quad (10)$$

where

$$(f)_{OO}^{(\exists \lambda \in \{\lambda_1^*, \lambda_2^*, \dots, \lambda_Q^*\})} = \left[ \left( O_{Min}^{\lambda_k^*}, O_{Mean}^{\lambda_k^*}, O_{Max}^{\lambda_k^*}, O_{Std}^{\lambda_k^*}, O_{Roun}^{\lambda_k^*}, O_{Comp}^{\lambda_k^*}, O_{Asym}^{\lambda_k^*}, O_{Rect}^{\lambda_k^*}, O_{MeanIn}^{\lambda_k^*}, O_{Density}^{\lambda_k^*}, O_{BorderI}^{\lambda_k^*}, O_{Shapel}^{\lambda_k^*}, O_{Elliptic}^{\lambda_k^*} \right) \Big|_{k=1, \dots, Q} \right] \quad (11)$$

represents the collections of 13 object features, including pixel value-based measures, such as the minimum, the maximum, the mean, the standard deviation, and the mean of the inner border, and geometrical measures, such as the roundness ( $O_{Roun}^{\lambda_k^*}$ ), the compactness ( $O_{Comp}^{\lambda_k^*}$ ), the asymmetry ( $O_{Asym}^{\lambda_k^*}$ ), the rectangular fit ( $O_{Rect}^{\lambda_k^*}$ ), the border index ( $O_{BorderI}^{\lambda_k^*}$ ), the shape index ( $O_{Shapel}^{\lambda_k^*}$ ), and the elliptic fit ( $O_{Elliptic}^{\lambda_k^*}$ ).

Finally, Figure 3 shows the overall technical flowchart for the proposed method.

### 2.2.3. Experimental Setup

To analyze the performance of the introduced the multiclass classification methods ND and END, state-of-the-art and classic ML algorithms, including C4.5 [87], END with ERDT (END-ERDT) [91], RaFs [89], ExtraTrees [91,98], classification via random forest (CVRaFs) [126], RoFs [90], and an SVM [64], were also applied in direct- or ECOC-based multiclass classification. The considered ECOC methods include one-versus-one (ECOC:1vs1), one-versus-all (ECOC:1vsAll), random correlation (ECOC:RC), dense random (ECOC:DR), sparse random (ECOC:SR), and ordinal (ECOC:Ordinal) methods. Critical tree parameters of C4.5, END-ERDT, RaF, RoF, CVRaF, and ExtraTrees classifiers are set by default, while the ensemble size is set to 100 by default for RaF, RoF, CVRaF, and ExtraTrees. The involved parameters of the radial basis function (RBF) kernel-based SVM were tuned by using Bayes optimization (SVM-B) and 10 by 10 grid-search optimization (SVM-G) [127].

We applied a disk-shaped SE with  $n = 10$  openings and closings by conventional and partial reconstructions to obtain the MPs and MPPR from the four raw bands of MSIL1C and the first three PCA-transformed components, ranging from one to ten with a step-size increment of one. These parameters mean that we obtain  $84 = 4 + 4 \times 10 \times 2$  dimensional datasets using four raw bands and  $63 = 3 + 3 \times 10 \times 2$  dimensional datasets using the first three PCA-transformed components, which are represented by Raw\_MPs, Raw\_MPPR, PCA\_MPs, and PCA\_MPPR in the graphs in the experimental parts. For fair evaluations from dimensionality, we set the MRS segregation scale parameter  $\lambda$  with 10 different values in the FE phase for OMPs and EOMPs. In other words, we obtained  $84 = 4 + 4 \times 10 \times 2$  and  $63 = 3 + 3 \times 10 \times 2$ -dimensional datasets for the raw and PCA-transferred data, respectively, while using OMPs and  $524 = 4 + 4 \times 13 \times 10$  and  $393 = 3 + 3 \times 13 \times 10$  dimensional OO feature datasets from the raw and PCA-transformed data, respectively. Naturally, there are  $604 = (524 - 4) + 84$  and  $453 = (393 - 3) + 63$  dimensional datasets for raw and PCA-transferred data, respectively, while using EOMPs.

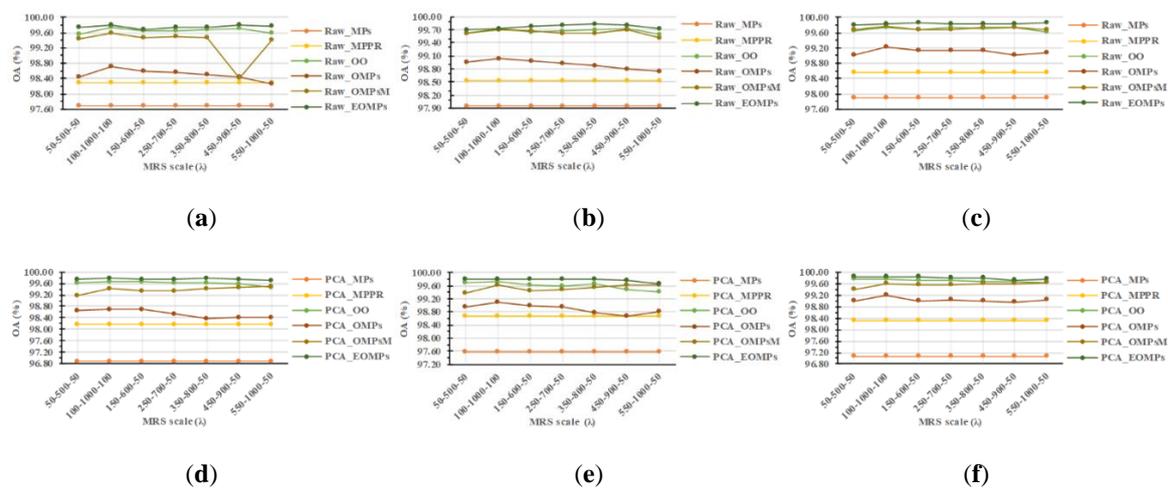
In the experiment, the average accuracy (AA), the overall accuracy (OA), the CPU running time (CPUTime), and the kappa statistic were used to evaluate the classification performance of all the considered methods. All of the experiments were conducted while using Oatave 5.1.0 on a Windows 10 64-bit system with an Intel Core i7-4790 3.60 GHz CPU and 64 GB of RAM.

### 3. Results

#### 3.1. Subsection Assessment of the Feature Extractors

##### 3.1.1. Accuracy Evaluation

Figure 4 illustrates the OA values from the ensemble methods, including RaF, ExtraTrees, and END-ERDT while using MPs, MPPR, and EOMP features that were extracted from the raw and PCA-transformed datasets. Each point on the  $x$ -axis represents the MRS scale sets for OO, OMPs, OMPsM, and EOMPs feature extractors (e.g., 50-500-50 means the scale parameter  $\lambda$  of MRS that starts with 50 and stops at 500 with total 10 steps by step 50), while the  $y$ -axis representation the OA values. First, the superiority of the proposed FE method EOMPs is obvious when compared with that of the MPs, MPPR, OMPs, and OMPsM, and the superiority of OO as compared with that of the MPs, MPPR, and OMPs. Specifically, the best improvements were achieved by EOMPs across all three classifiers with two datasets (see the dark green lines). Moreover, the superiority of MPPR compared to MPs and OMPs and the superiority of OMPsM when compared to MPPR is clear, which again supports the findings by Liao et al. [97] and Samat et al. [98]. Additionally, the performance of OO and OMPs could actually be limited by setting the segmentation scale parameter  $\lambda$  to very large values. For example, a decreasing trend in the OA values from OMPs can be observed after the starting scale is larger than 100 with 10 or 50 scale steps (see the brown lines).

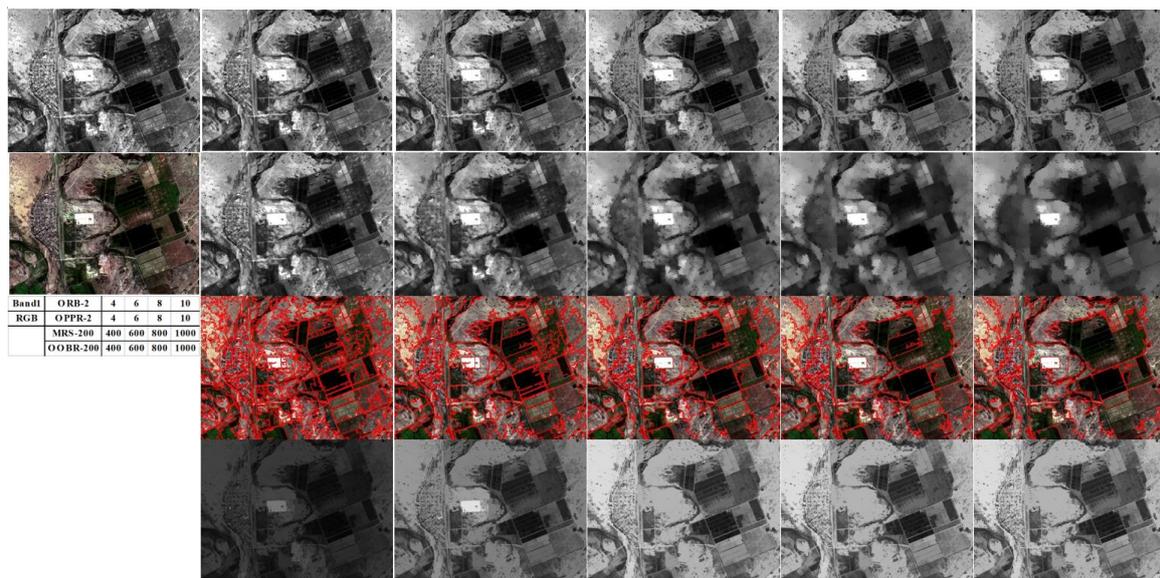


**Figure 4.** Overall accuracy (OA) values from RaF (a and d), ExtraTrees (b and e), and END-ERDT (c and f) using morphological profiles (MPs) (disk, 1-10-1), MPs with partial reconstruction (MPPR) (disk, 1-10-1), and extended object-guided MPs (EOMP) features extracted from the raw bands.

##### 3.1.2. Visual Evaluation

In Figure 5, a  $600 \times 800$  image patch was selected from the south-central area of the Bakbakty irrigation area (Figure 1f) to show the differences between the OBR, opening by partial reconstruction (OBPR) and object guided OBR (OOBR) operators with different scale parameter settings while using the first raw band. According to the graphs in the first row of Figure 4, the image becomes slightly grayer as the size of the SEs increases in OBR, with most of the small details, such as boundaries between objects, still remain. In contrast, boundaries between different objects become too blurred and indistinguishable as the size of the SEs increases in OPBR; many large objects, such as the urban area in the central-western area that should appear at a certain scale of the area attribute, remain at a low scale, and disk shapes, such as new objects, are created with large SE size after OPBR (see the last image in row 2 of Figure 4). For the proposed EOMPs, the target image becomes slightly grayer as the scale parameter  $\lambda$ , which controls the total number and individual scales of segments, increases. Furthermore, most of the boundaries between the different land cover types remain exactly as in the original, which is mainly

due to the EOMP's only filtering the areas within the corresponding boundaries. However, when the segments are too large and are composed of many different objects, the performance of OOB-OB could be limited by returning profiles of only one object. In addition, different segments could have the same minimum and/or maximum pixel values; as a result, OOB-OB could return similar, or even the same, profiles for different objects. For example, a very bright and rectangular building in the center of the target image cannot be distinguished from its surroundings when the value of  $\lambda$  is greater than 600 (see the last row of Figure 4).



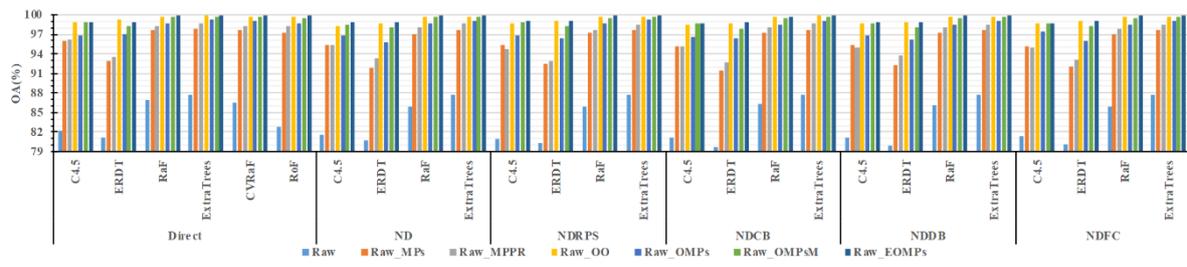
**Figure 5.** Examples of OBR (row 1), opening by partial reconstruction (OBPR) (row 2), multi-resolution segmentation (MRS) segments (row 3), object-guided OBR (OOBR) (row 4) computed from band 1 (first image at row 1) at the center-bottom of Figure 1f (the numbers in the table in row 3 show the disk sizes in OBR and OBPR and the segmentation scale  $\lambda$  in MRS).

### 3.2. Evaluation of ND and END

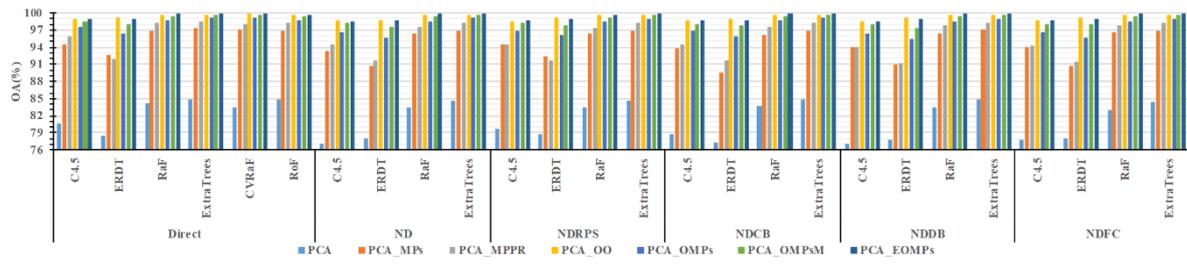
#### 3.2.1. Classification Accuracy

As mentioned in Part 1, the second objective of this paper is to investigate the performance of popular ND algorithms and their ensemble versions. Hence, Figure 6 presents the OA values from various classifiers that were adopted in direct, ND, ECOC, and END multiclass classification frameworks by using all of the considered features.

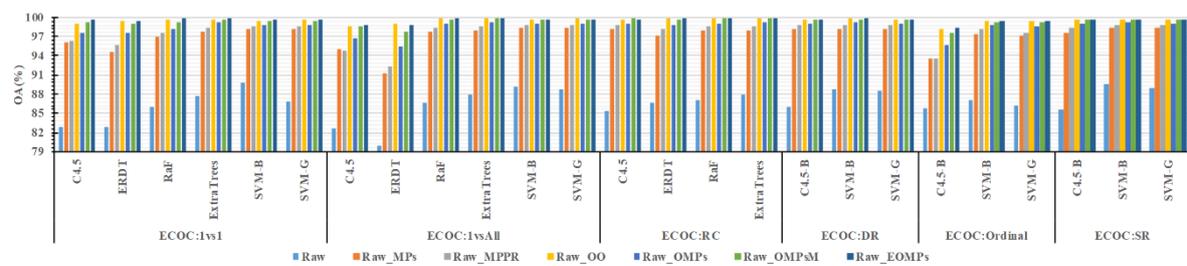
If we simply compare the OA bars from all of the adopted classification algorithms while using various features in all three multiclass classification framework scenarios, the results of MPPR are superior to those of MPs and OMPs, and the results of OMPs are superior to those of MPPR, OO is superior to MPs and MPPR, EOMPs is superior to all others, and are uniformly shown in almost all of the classification scenarios, which confirms the superiority of our proposed method.



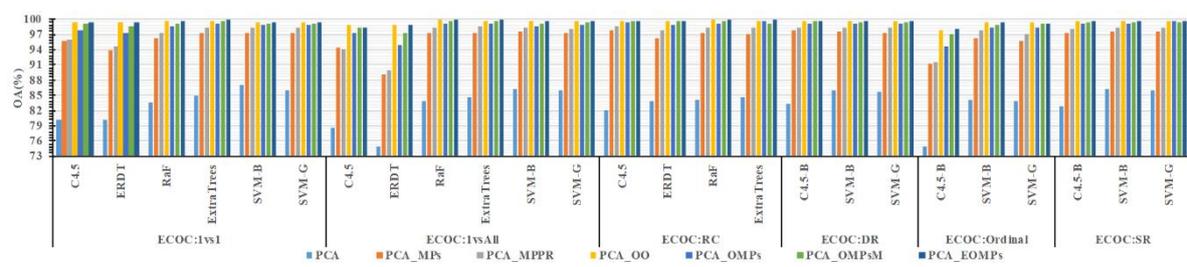
(a)



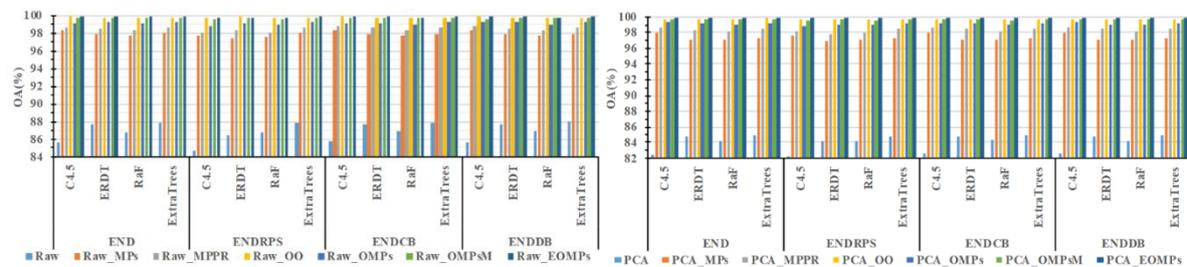
(b)



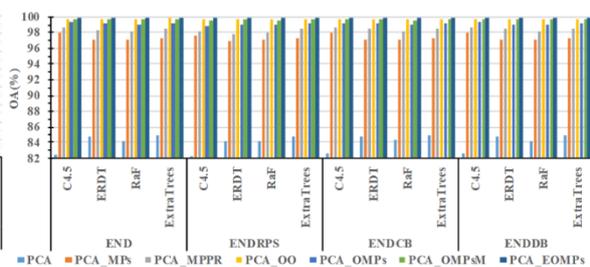
(c)



(d)



(e)



(f)

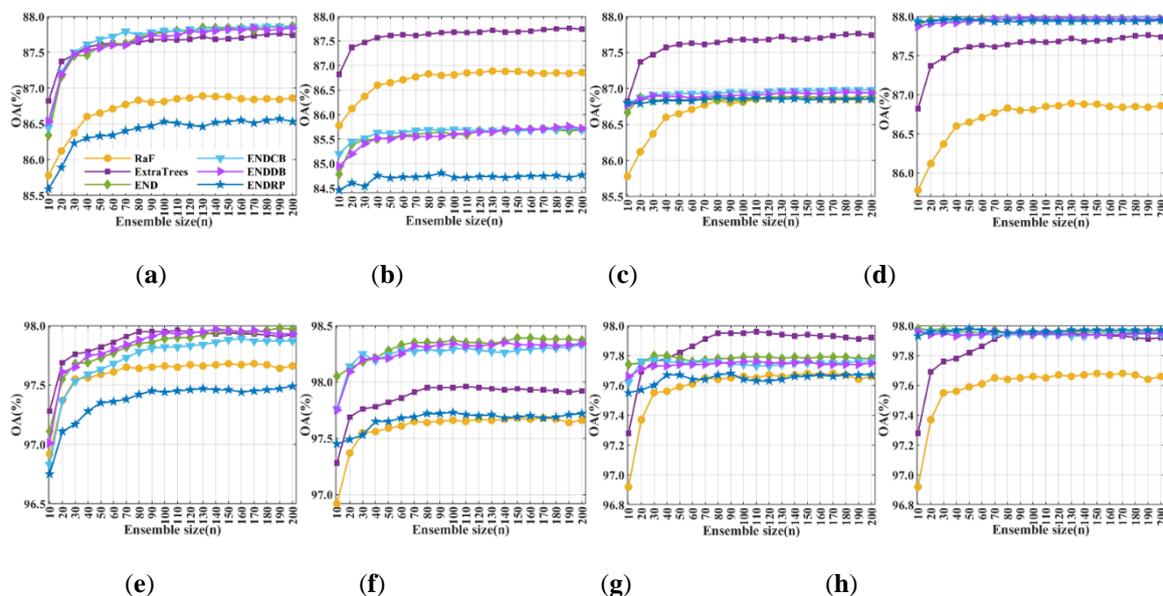
**Figure 6.** OA values from various classifiers in different multiclass classification frameworks (Direct and ND: a, b; ECOC: c, d; END: e, f) using all the considered features.

When comparing the OA bars in Figures 6a and 6b from C4.5, ERDT, RaF, and ExtraTrees in direct and ND multiclass classification, improvements from ND, NDCB, NDDB, NDRPS, and NDFC over the direct framework is not clear. Interestingly, the performance of the weak, direct multiclass classification algorithm can be reduced in the case of readily available low-dimensional data with low discrimination capability. For instance, the C4.5 classifier reached OA values that were greater than 82% and 80% individually while using the original raw bands and the PCA-transformed datasets, respectively, in the direct multiclass classification framework; moreover, C4.5 in ND, NDCB, NDDB, NDRPS, and NDFC multiclass classification frameworks uniformly reached OA values that were less than 82% and 79% while using the original raw bands and the PCA-transformed datasets, respectively. Similar results can also be found for ERDT not only using the original raw and PCA-transformed datasets, but also using MPs and MPPR features from the original raw and PCA-transformed datasets, whereas the ERDT has proven much weaker than C4.5 [98]. When comparing the OA bars of ensemble classifiers, such as RaF and ExtraTrees, there are no obviously increased or decreased OA values observed for ExtraTrees in the direct and ND, NDCB, NDDB, NDRPS, and NDFC multiclass classification frameworks, but a slightly decreasing trend is shown by RaF in the ND, NDCB, NDDB, NDRPS, and NDFC frameworks while using the original raw and PCA-transformed datasets.

According to the results that are shown in Figures 6c and 6d, there are no obvious increases or decreases in the OA for the same classifiers with different ECOC techniques, except for ERDT and C4.5 in the one vs. all (1 vs. all) and C4.5 in ordinal multiclass classification cases while using original raw and PCA transformed datasets. Additionally, differences in OA values from ECOC techniques using OO and spatial features are smaller and more stable than those from the ND frameworks. Take the C4.5 classifier as an example, 95%–99% and 93%–98% OA value ranges for ND multiclass classification framework becomes into 98%–99.80% and 97.5%–99.80% OA values for ECOC RC. Additionally, more interestingly in comparing with direct and ND frameworks, better OA results can always be reached for weak classifiers (e.g., C4.5, ERDT) in ECOC one vs. one, random correlation, dense random, and sparse random multiclass classification techniques. For instance, a minimum larger than 82% (ECOC 1vs all) and maximum around 86% (ECOC RC, DR and SR) OA values are shown by C4.5 in ECOC frameworks while using original raw bands, while minimum larger than 81% (NDRP) and maximumly larger than 82% (ND) OA values are shown in ND, NDCB, NDDB, NDRP, and NDFC frameworks. On the contrary, when the stronger classifiers, such as RaF, ExtraTree, and SVM are adopted, differences between them in direct, ND, and ECOC frameworks are much smaller, especially from those using high dimensional datasets with high discrimination capabilities. In contrast with RaF and ExtraTrees, better OA values could be reached by SVM in ECOC frameworks while using low dimensional datasets with low discrimination capabilities in the original raw bands and PCA-transformed datasets.

By comparing the results in Figures 6e and 6f with the results in Figures 6a and 6b, we can clearly observe the superiority in the OA values of END, ENDRPS, ENDCB, and ENDDDB over ND, NDRPS, NDCB, and NDDB, respectively, which is in accordance with the findings from Frank and Kramer [93], Dong et al. [83], and Rodríguez et al. [94]. Interestingly, the OA values of ERDT in the END, ENDRPS, ENDCB and ENDDDB frameworks always reached better OA values than C4.5 and RaF (except for ENDRPS) with the same multiclass classification sets, even when using the original raw bands and PCA-transformed datasets with low discrimination capabilities (see the bars in light blue in Figures 6e and 6f). When better data with high discrimination capabilities are available, the END, ENDRPS, ENDCB, and ENDDDB multiclass classification frameworks are capable of reaching better OA values while using weak but simple classifiers (e.g., C4.5 and ERDT) than direct and ECOC when using stronger but more complex classifiers (e.g., RaF, ExtraTrees, and SVM). For example, the OA values for C4.5 and ERDT are approximately 98% larger in the END framework, while the OA values for SVM-B and SVM-G are approximately 97% larger in the ECOC:Ordinal framework while using various spatial features that were extracted from the original raw bands.

In Figure 7, we present the OA curves from the direct and END-based multiclass classification frameworks with incrementally increased ensemble size. The conventional C4.5 and ERDT classifiers are adopted in direct multiclass classification approaches RaF and ExtraTrees, respectively. C4.5, ERDT, RaF, and ExtraTrees are adopted as the base learners in the END, ENDCB, ENDDB, and ENDRPS frameworks. Note that the size of RaF and ExtraTrees are set to 100 in the END, ENDCB, ENDDB, and ENDRPS frameworks. Based on the results, the superiority of ENDCB and ENDDB over END is not obvious in the context of the OA values, as shown in a study by Dong et al. [75]. In contrast, ENDRPS showed the worst results while using the C4.5 and ERDT classifiers. Additionally, the END, ENDCB, and ENDDB frameworks with the ERDT classifier can achieve classification accuracy results that are better than those attained by RaF, by using both the original raw bands and the MPs features that were extracted from raw bands (see the results in Figures 7a and 7e). However, optimum results can be reached by feeding the ExtraTrees to the END, ENDCB, ENDDB, and ENDRPS frameworks. For effects from the ensemble size, increasing the ensemble size beyond 80 does not yield obvious improvements in the OA values for the END, ENDCB, ENDDB, and ENDRPS frameworks with C4.5 and ERDT while using the considered features, while increasing the ensemble size beyond 30 does not yield obvious improvements in the OA values for the END, ENDCB, ENDDB, and ENDRPS frameworks with RaF and ExtraTrees.



**Figure 7.** OA values versus the ensemble size of the END, ENDCB, ENDDB, and ENDRPS frameworks with ERDT (a, e), C4.5 (b, f), RaF (c, g), and ExtraTrees (d, h) classifiers while using the original raw bands (a, b, c, d) and the MPs (e, f, g, h).

### 3.2.2. Computational Efficiency

Computational efficiency is always considered to be another key factor after the classification accuracy when evaluating a classifier's performance. In accordance with Figures 6 and 7, Figure 8 shows the CPUTime (in seconds) in the training phase for various classifiers in different multiclass classification frameworks and using all of the considered features, while Figure 9 shows the results for the END, ENDCB, ENDDB, and ENDRPS frameworks with different ensemble sizes.

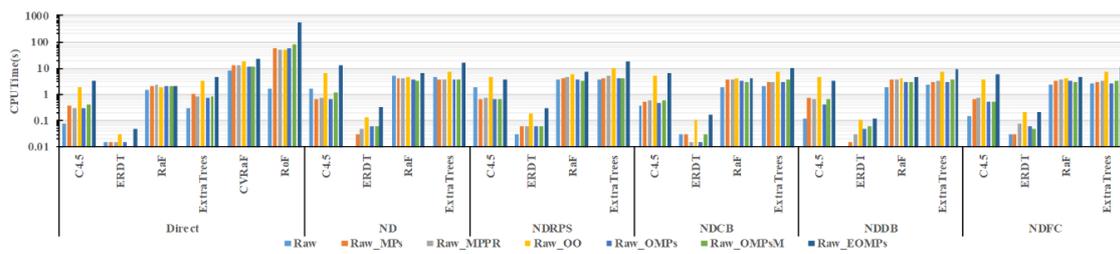
When comparing the charts in Figure 8, direct ERDT is at least 10 to 1000 times faster than the C4.5, RaF, ExtraTrees, CVRaF, and RoF classifiers, ExtraTrees is faster than RaF, CVRaF, and RoF, which is in accordance with our previous findings [98]. The extremely fast operability of ERDT is inherently available in the ND, NDCB, NDDB, NDFC multiclass classification frameworks, and in their ensemble versions, as shown in Figures 8e and 8f. Specifically, using ERDT in the ND, NDCB, NDDB, NDFC multiclass classification frameworks is at least 10 times faster than using C4.5 and at

least 100 times faster than using RaF and ExtraTrees. It is reasonable that the ensemble size of RaF and ExtraTrees is set to 100 as the default in those frameworks.

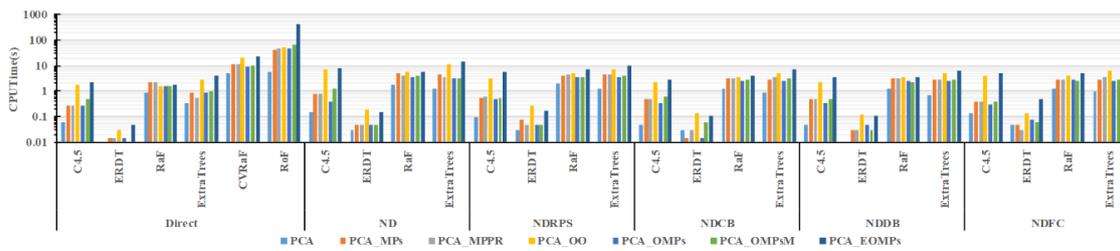
In contrast with results from the ECOC frameworks, as shown in Figures 8c and 8d, C4.5, ERDT, RaF, and ExtraTrees in the ND, NDCB, NDDB, NDRPS, and NDFC frameworks are slightly faster than their corresponding frameworks in the ECOC:1vs1, ECOC:1vsAll, and ECOC:RC frameworks. As expected, the worst computational efficiency is shown by the ECOC frameworks with SVM-B and SVM-G parameter optimization techniques. Specifically, SVM-B is 10 times faster than SVM-G, whereas the former is at least 1000 times slower than ERDT in the ND frameworks and at least 100 times slower than ERDT in the END frameworks.

Critical tree parameters, including the minimum leaf size and the maximum depth, are also tuned using Bayes optimization in the ECOC:DR, ECOC:Ordinal, and ECOC:SR frameworks to identify the computational effects from parameter optimization. As shown in Figures 8c and 8d, the computational burden from the parameter tuning process is also severe for C4.5. For instance, the ECOC:SR framework with C4.5 took approximately 1000 seconds of CPUTime on the four original raw bands, while less than 5, 10, and 100 seconds are usual in the direct, ND, NDCB, NDRPS, NDDB frameworks, and their ensemble version frameworks. If we correspondingly look back at the OA results that are shown in Figure 6, obvious improvements in the OA values are not indicated. In other words, the computational complexity that was brought by parameter optimization could be further eliminated in more sophisticated ECOC multiclass classification frameworks without an obvious reduction in the accuracy.

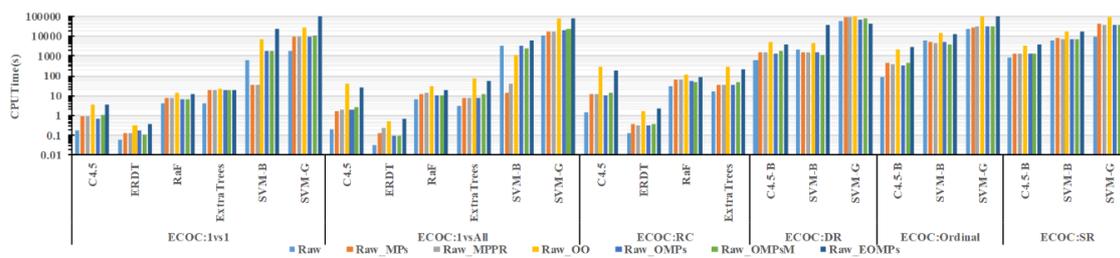
According to the results that are shown in Figure 9, it is clear that the direct classifier ExtraTrees is faster than RaF, and RaF is faster than the END, ENDCB, ENDRPS, ENDDDB multiclass classification frameworks while using C4.5, ERDT, RaF, and ExtraTrees as the base learners. Moreover, the computational efficiency of ENDCB and ENDDDB over END is also clear, while all of the computational costs of END, ENDCB, ENDRPS, and ENDDDB frameworks linearly increase as the ensemble size increases. Interestingly, both the adopted classifier and the ND frameworks can influence the computational efficiency. For instance, the worst computational efficiency is shown by ENDRPS with ERDT while using both the regional raw and MPs datasets (see Figures 9a and 9e), while END with C4.5, RaF and ExtraTrees showed the worst computational efficiency. According to the results that are shown in Figure 7, ENDRPS with ERDT might not be the optimal choice for both accurate and efficient classification with respect to the performance of the END, ENDCB, and ENDDDB frameworks.



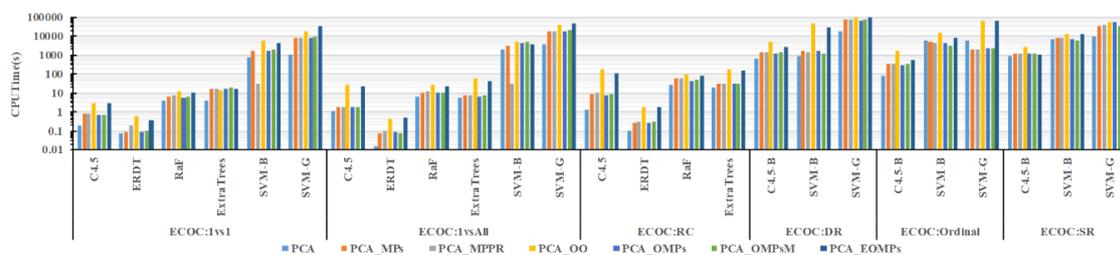
(a)



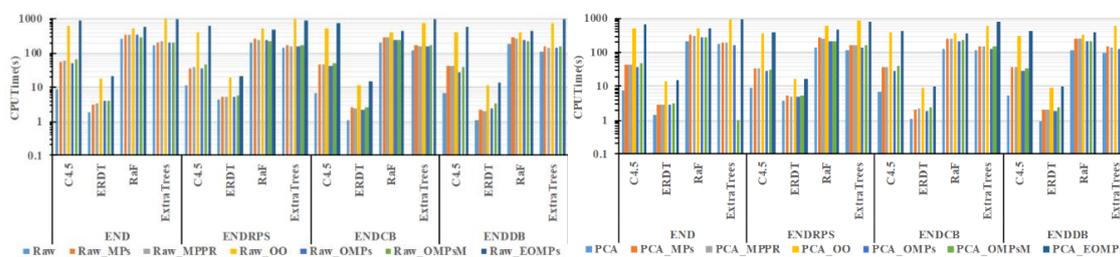
(b)



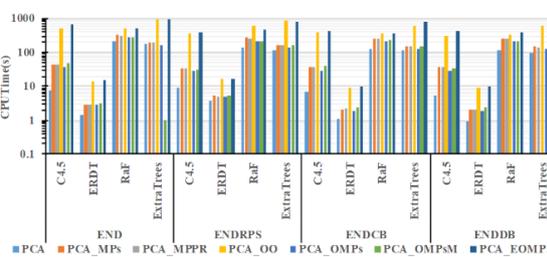
(c)



(d)

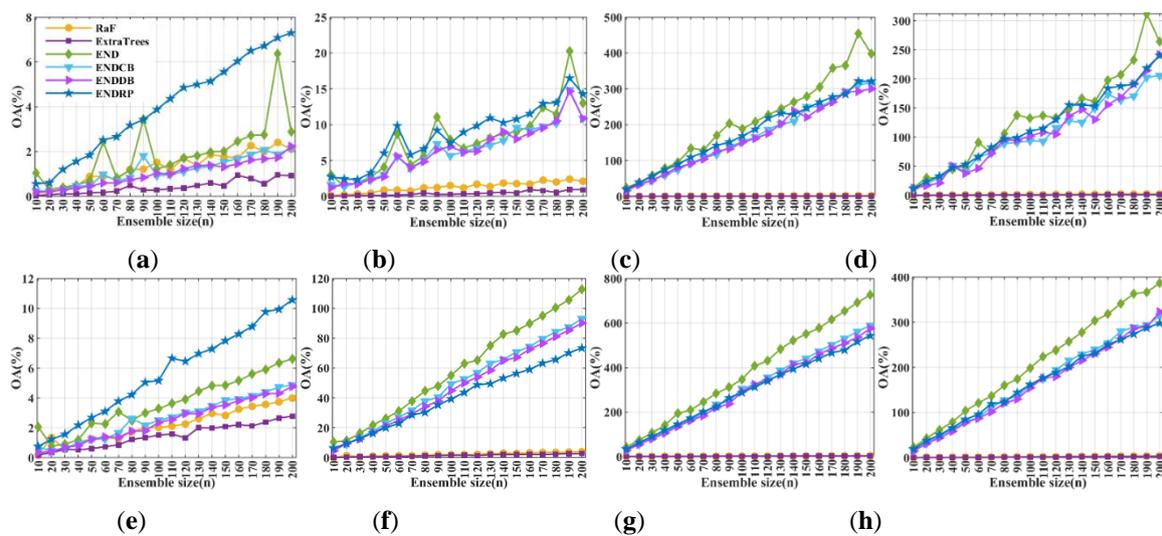


(e)



(f)

**Figure 8.** CPU running in seconds for various classifiers in different multiclass classification frameworks (Direct and ND: a, b; ECOC: c, d; END: e, f) using considered features.



**Figure 9.** CPU runtime in seconds versus the ensemble sizes of the END, ENDCB, ENDDDB, and ENDRPS frameworks with the ERDT (a, e), C4.5 (b, f), RaF (c, g), and ExtraTrees (d, h) classifiers using the original raw bands (a, b, c, d) and features from the MPs (e, f, g, h).

### 3.2.3. Robustness to the Data Dimensionality

Data quality is also a critical factor that controls the classification performance of adopted classifiers, and many approaches can be used to increase the discrimination and identification quality of the provided data by introducing new features. However, increasing the number of data dimensions by introducing new features could limit the training samples large enough to mitigate the Hughes phenomenon on the one hand and increase the computational complexity of feature space splitting-based classifiers (e.g., C4.5, RaF, and RoF) on the other hand. Hence, it is of interest to comparatively investigate the robustness of ND and END to the data dimensionality.

According to the results in Figures 6 and 7, the improved data quality by introducing new features is clear. For various single and ensemble methods, direct and ND-based classifiers, C4.5 is more robust than ERDT to the data dimensionality in the direct, ND, NDCB, NDDDB, NDRPS, and NDFC frameworks. For example, ND with ERDT achieves OA values between 92% and 99% after features from MPs, MPPR, OMPs, OMPsM, and EOMPs are introduced, while ND with C4.5 achieves OA values that are between 95% and 99% (see Figure 6a). The ensemble versions of C4.5 and ERDT are less robust than the RaF, RoF, CVRaF, and ExtraTrees, both in direct and various ND. When compared with the results from direct and various ND frameworks, uniformly better robustness to data dimensionality is shown by all of the ECOC frameworks, especially with the RaF, ExtraTrees, and SVM classifiers. Taking the ECOC:RC framework with C4.5 as an example, the OA values range between 95% and 99% for ND with C4.5 and they shrink to a range between 98% and 99% after features from MPs, MPPR, OMPs, OMPsM and EOMPs are introduced.

As expected, the ECOC frameworks with SVM show better robustness to data dimensionality than the ECOC frameworks with C4.5, ERDT, RaF, and ExtraTrees, whereas the SVM is capable of overcoming the Hughes phenomenon that is caused by the data dimensionality with kernel trick [64,65]. When comparing the OA values from various END-based multiclass classification frameworks, it is clear that 1) various END frameworks have better robustness to the data dimensionality than various ND frameworks; 2) differences in the robustness to the data dimensionality between C4.5 and ERDT, C4.5, and RaF, and ERDT and ExtraTrees in various END frameworks are much smaller than those from various ND frameworks; and, 3) similar and even better than ECOC frameworks on the robustness to the data dimensionality can be reached by the END frameworks. For instance, END with C4.5 showed an OA ranging between 98% and approximately 99.8% after various considered features are

introduced, while most of the ECOC frameworks with SVM show an OA that ranges between 97% and approximately 99.8%.

As shown in Figure 8, the computational complexity that was brought by the data dimensionality is clear for all classifiers in all of the multiclass classification frameworks. Especially for the C4.5, ERDT, ExtraTrees, RoF, and CVRaF classifiers that adopt feature splits or selection criteria in feature spaces that control the complexity of adopted DTs. For example, a higher computational cost is always shown for ERDT and ExtraTrees in the END frameworks by using Raw\_OO (with 524 dimensions) and Raw\_EOMPs (with 604 dimensions) features, while a similar and lower computational cost is shown by using Raw\_MPs, Raw\_MPPR, Raw\_OMP, and Raw\_OMP\_M features (see Figure 8e). RaF is more robust than the C4.5, ERDT, ExtraTrees, RoF, and CVRaF classifiers to the data dimensionality. From a computational efficiency point of view, the best robustness to the data dimensionality is always shown by ERDT in the direct, ND, ECOC and END frameworks. Additionally, because of the kernel trick, differences in the robustness to the data dimensionality from SVM in ECOC frameworks are smaller than those from the DT-based classifiers that were adopted in the END frameworks.

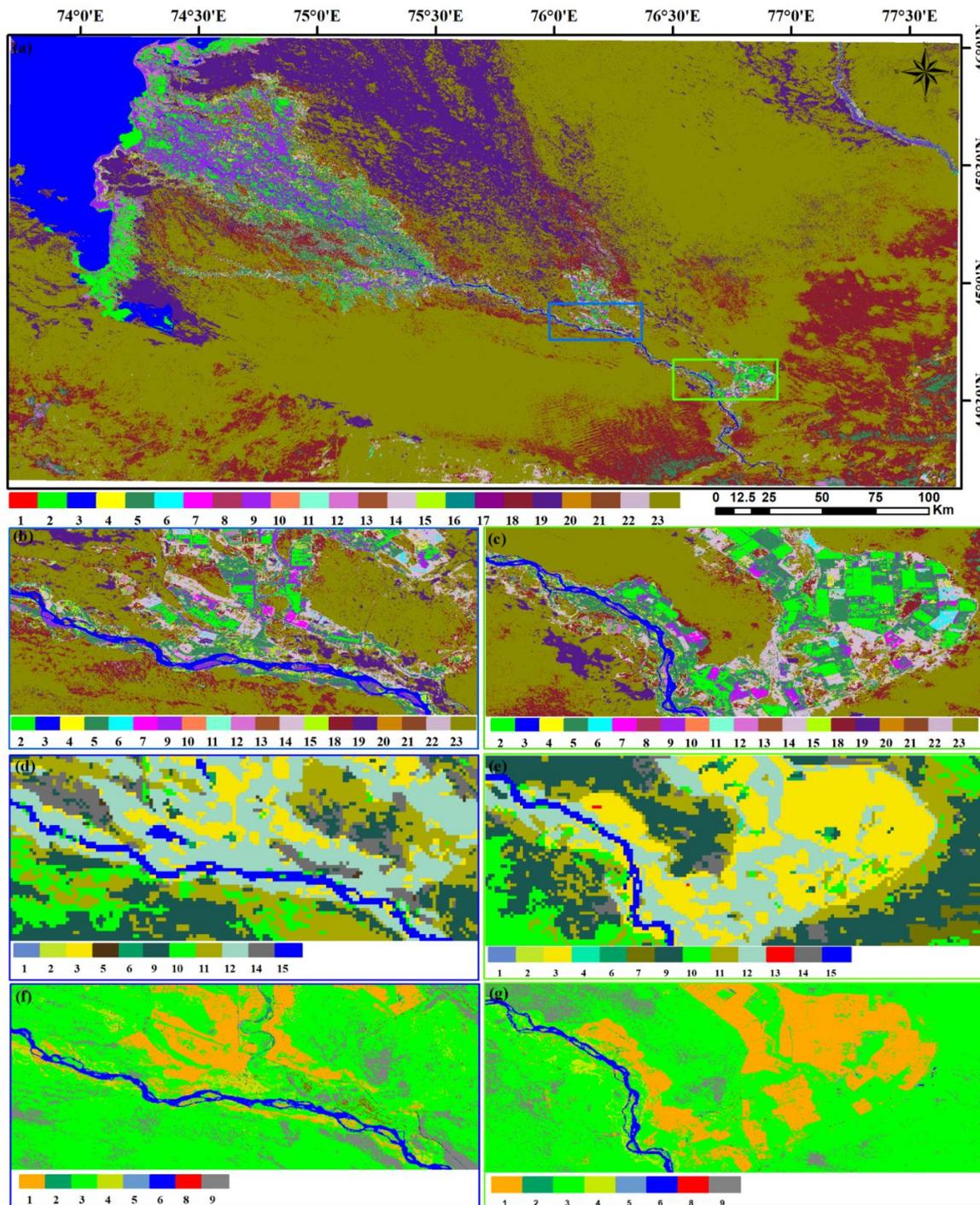
### 3.3. Final Vegetation Map

Figure 10 shows the classification map using the proposed method and the considered products to show the superiority of the Sentinel-2 MIL1C products over the MODIS LUC and GLC30 datasets in arid region. To further compare the findings of END-ERDT capable of reaching the best classification accuracy with a very high computational efficiency, Table 3 reports the classification accuracy values (the user accuracy (UA), AA, OA, and kappa statistics) with CPUTime in seconds for END-ERDT and ECOC:1vsAll with SVM-G optimization.

According to the results in Figure 10, it is apparent that Sentinel-2A MIL1C is better than MODIS LUC and GLC30 for vegetation diversity mapping in arid regions in Central Asia. Specifically, 19 different vegetation types were recorded by Sentinel-2A MIL1C for our study area, while 15 and eight land cover types were recorded by MODIS LUC and GLC30 products without specific vegetation taxonomic names. For instance, vegetation species, such as *Alhagi sparsifolia*, *Haloxylon ammodendron*, and *Artemisia lavandulaefolia* are classified as shrubs or herbaceous, while *Iris lactea* Pall. & *Sophora alopecuroides* and *Sophora alopecuroides* are classified into grassland in the MODIS LUC and GLC30 products. From a vegetation species taxonomy and distribution mapping point of view, the land cover taxonomy classification system might not be appropriate. For example, the vegetation species richness, which is defined as the numbers of different species that are present in a certain study zone, for the Bakanas and Bakbakty irrigation zones that are depicted by blue and green rectangles, respectively, in Figure 10a is four (crops, forest, grass, and shrubs) from the GLC30 product (see Figures 10f and 10g), five (crops, tree, grass, herbaceous, shrubs) from the MODIS LUC product, and 12 (rice, cloves, wheat, corn, reeds, *Alhagi sparsifolia*, *Carex duriuscula*, shrubs, *Haloxylon ammodendron*, grass, tamarisk, *Iris lactea* Pall., and *Sophora*), and 13 (rice, cloves, wheat, corn, desert steppe, reeds, *Alhagi sparsifolia*, *Carex duriuscula*, shrubs, *Haloxylon ammodendron*, grass, tamarisk, *Iris lactea* Pall., and *Sophora alopecuroides*) from the Sentinel-2 MIL1C classification with END-ERDT while using spectral and spatial features.

Based on the results in Table 3, again, it can be clearly seen that the END-ERDT method is capable of achieving the best results (OA = 99.85%) while using the stacked raw and EOMPs features with the highest model training efficiency (15.20 seconds) with respect to the results from RBF kernel-based SVM-G optimization teaching in the ECOC:1vsAll multiclass classification framework, which confirms the previous findings that END-ERDT could be an alternative to an SVM for generalized classification accuracy, computationally efficient operations, and easy to deploy points of view, especially in the case of sufficient samples with advanced features that are readily available. When the original raw data were adopted, OA values of 87.80% and 88.71% were achieved by the END-ERDT and SVM classifiers, respectively. Furthermore, END-ERDT showed the worst UA of 15.14% for *Alhagi sparsifolia*, while SVM showed the worst UA values of 2.75% for *Alhagi sparsifolia* and 1.79% for

*Sophora alopecuroides*. After the advanced features were included, almost all of the land cover classes were correctly classified with a > 95% UA value by both classifiers, and especially after the OO and EOMP's were included. However, only on the raw data, the END-ERDT model was trained in several to more than ten seconds, the optimum RBF kernel-based SVM model took more than ten thousand seconds.



**Figure 10.** Final vegetation distribution map using Sentinel-2 MSIL1C products with the END-ERDT classifier for our study area (a) and subareas (b, c) and corresponding examples from the 2015 MODIS LUCCL products (d, e) and the 2017 GLC30 (f, g) products (for the legends, refer to that in Figure 1).

**Table 3.** Classification accuracy values (user accuracy (UA), average accuracy (AA), OA, and kappa) for the considered methods in the study area.

Class No.	END-ERDT							ECOC:1vsAll (SVM-G)						
	Raw	Raw_MPs	Raw_MPPR	Raw_OMP	Raw_OMP	Raw_OMP	Raw_OMP	Raw	Raw_MPs	Raw_MPPR	Raw_OMP	Raw_OMP	Raw_OMP	Raw_OMP
1	54.01	92.27	96.21	96.06	98.50	99.64	99.64	62.53	93.52	93.59	92.95	96.39	97.57	98.85
2	98.12	99.63	99.84	100.00	100.00	100.00	100.00	97.64	99.59	99.86	100.00	99.91	100.00	100.00
3	99.86	99.86	99.88	99.95	99.96	100.00	100.00	99.66	99.52	100.00	99.88	99.86	100.00	100.00
4	63.65	92.99	97.62	98.31	99.85	100.00	100.00	55.58	97.80	97.10	98.86	99.96	100.00	100.00
5	79.79	96.25	98.03	97.59	99.97	100.00	100.00	82.44	96.73	97.98	97.61	99.92	99.92	99.92
6	95.46	99.32	99.88	100.00	100.00	100.00	100.00	97.77	99.11	99.98	100.00	100.00	100.00	100.00
7	89.62	97.65	98.64	99.14	99.14	97.57	98.02	94.28	98.27	99.05	99.63	98.89	97.32	97.32
8	25.28	92.90	97.44	100.00	100.00	100.00	100.00	37.50	99.43	98.86	99.15	99.72	100.00	100.00
9	88.68	98.60	99.03	99.90	99.97	100.00	100.00	88.18	98.75	98.44	99.77	99.69	100.00	100.00
10	15.14	75.23	95.41	94.95	96.79	98.17	100.00	2.75	92.66	97.71	94.95	92.20	100.00	100.00
11	79.52	98.40	99.32	99.54	99.89	100.00	100.00	88.82	98.75	99.83	99.77	99.89	100.00	100.00
12	77.42	94.96	93.61	97.59	98.49	98.33	99.97	79.25	94.47	93.90	94.67	99.16	98.30	98.30
13	58.19	97.23	97.86	99.37	99.24	99.87	99.87	87.41	97.98	96.98	99.12	99.37	99.87	99.87
14	90.56	97.03	98.21	98.99	99.66	99.90	99.96	91.86	98.12	98.82	99.09	99.71	100.00	100.00
15	43.17	91.06	95.46	97.84	99.42	99.28	99.28	37.11	93.89	94.12	98.88	99.46	98.97	99.06
16	53.40	93.54	96.48	97.07	98.71	99.77	99.77	56.57	94.13	95.66	96.71	99.41	99.77	99.77
17	33.33	76.74	95.16	90.50	93.99	95.54	96.12	58.91	91.47	96.32	88.18	92.05	95.16	95.16
18	75.86	97.90	93.98	98.39	99.72	100.00	100.00	73.20	98.25	99.16	98.53	99.51	100.00	100.00
19	99.43	99.94	99.89	99.90	99.98	99.98	100.00	99.36	99.94	99.95	99.92	99.98	100.00	100.00
20	96.12	98.95	99.32	99.63	99.32	99.54	99.51	97.13	98.86	99.17	99.88	99.45	99.54	99.54
21	20.89	76.79	76.79	93.39	97.32	97.50	97.50	1.79	87.50	90.00	90.71	96.61	97.50	97.50
22	89.61	99.31	99.34	99.73	99.96	99.82	99.93	89.94	99.47	99.89	99.78	99.82	99.76	99.89
23	99.92	99.94	99.93	99.93	100.00	100.00	100.00	99.99	100.00	100.00	100.00	100.00	100.00	100.00
AA	70.74	94.20	96.84	98.16	99.13	99.34	99.55	73.03	96.88	97.67	97.74	98.74	99.29	99.36
OA	87.80	97.82	98.62	99.17	99.71	99.75	99.85	88.71	98.42	98.74	99.01	99.60	99.67	99.72
Kappa	0.87	0.98	0.98	0.99	1.00	1.00	1.00	0.88	0.98	0.99	0.99	1.00	1.00	1.00
CPUTime	1.14	2.61	2.16	2.32	3.52	11.22	15.20	11333.50	18625.80	18746.10	20524.90	22582.60	11898.40	62736.90

#### 4. Discussion

For arid land vegetation mapping while using Sentinel-2 MSIL1C image task, the superior performance of the proposed EOMPs over conventional OO, MPs, MPPR, OMPs, and OMPsM is confirmed, both statistically and visually. Additionally, as expected, possible side effects from very large segments could be controlled and even overcome by simply containing the mean pixel values of the objects and the object profiles, such as the compactness, roundness, and shape index. To overcome these potential drawbacks, multiple scale parameter  $\lambda$  should be provided in OOB and OCB. On the other hand, those that have been repeatedly proven effective object profiles should also be considered.

With respect to the results from various classifiers in direct, ECOC, ND, and END frameworks, END with ERDT (END-ERDT) always capable of reaching the highest OA values. This finding could be explained by the “diversity” foundation for constructing an effective EL system, which says that “weaker” classifiers (ERDT here) always have a better chance of reaching the trade-off between diversity and accuracy than “stronger” classifiers (C4.5 here) [85,128,129]. Additionally, according to statistical theory regarding EL, reduced numbers of dichotomies in ENDCB, ENDDB, and ENDRPS are still large enough to ensure that there is a high level of diversity among END members to facilitate improvement by the ensemble. Hence, investigating the performance of other weak classifiers in END framework will be an interesting topic.

Ensembling randomly generated ND is an effective approach to multiclass classification problems, as proven by the results in Figure 6 and by the works of Frank and Kramer [85]. However, the equal sampling strategy that was adopted in END could limit the classification accuracy by generating a very limited depth of trees that is controlled by the number of classes; moreover, a very unbalanced tree can negatively affect the runtime. To remedy such limitations, NDCB, NDDB, and their ensemble versions (ENDCB and ENDDB, respectively) were proposed by Dong et al. [75]. According to their results, the runtime efficiency of ENDCB and ENDDB were slightly better than that of END in the same cases, and no obvious improvements were observed by setting the ensemble size to a constant value. Hence, it is of interest to comparatively investigate the performances of END, ENDCB, ENDDB, and ENDRPS with various sets of ensemble sizes. Our experiments confirmed that the positive effects from ensemble size are larger for END frameworks with weak classifiers than those with strong classifiers.

In studies that involve RS for biodiversity searches, land cover classification is considered the first-order analysis for species occurrence and mapping [10,23]. In general, coarse-spatial-resolution satellite imagery (e.g., MODIS, TM, and ETM+) and land cover products (e.g., the MODIS land use and cover change (LUCC) and Global Land Cover 30 (GLC30) datasets) are useful in detecting and evaluating ecosystems and habitat structures on a large scale, while HR/VHR satellite imagery products are useful for estimating habitat quality, predicting taxonomic groups, determining species richness, and mapping diversity [130–132]. In arid and semiarid regions, sparsely distributed vegetation species are crucially important in regional ecosystems, but they are easily mixed into dominant land cover types (e.g., bare land) in coarse-resolution satellite imagery. Spectral unmixing and subpixel mapping techniques could eventually solve these problems; however, vegetation species diversity mapping at a fine scale using coarse-resolution satellite images from MODIS, TM, ETM+, and OLI sensors is still quite challenging. For example, spectral unmixing can determine the fractions of classes within mixed pixels, but it fails to predict the spatial location. Our experiment also showed that the Sentinel-2 MIL1C products proved to be a more valuable data source than MODIS LUCC and GLC30 datasets for arid land vegetation mapping. Hence, Sentinel-2 products with 10m, 20m, and 60m spatial resolution, 13 bands spanning from the visible and the near-infrared (VNIR) to the short-wave infrared (SWIR) portion of the spectrum 12 spectral bands, and with a five-day revisit time over land and coastal areas, are better choice than MODIS, TM, ETM+, and OLI for arid land vegetation mapping.

Based on this work, we also envisage future perspectives. Further calculating more advanced vegetation species diversity indices, such as the spectral variation hypothesis (SVH), alpha-diversity, and beta-diversity, to show the superiority of Sentinel-2 MSIL1C images over MODIS and Landsat images should be an interesting future direction, especially at large regional or national scales.

Since END-ERDT showed a state-of-the-art classification performance, statistical and more empirical experiments should both also be conducted. Finally, we will deploy the END-ERDT method on a Spark platform to support big data processing to facilitate its application.

## 5. Conclusions

Sentinel-2 MSIL1C images of the Ili River delta region of Kazakhstan were classified while using spectral and EOMPs to investigate the performance of the Sentinel-2A MSIL1C products for vegetation mapping in an arid land environment with respect to land cover products from MODIS and Landsat and to answer the question of “is ND and END are superior to state-of-the-art direct and ECOC-based-multiclass classification approaches?” and an accurate classification purposes.

According to the results, several conclusions can be drawn. First and foremost, the proposed EOMP features are better than all of the features, while the OO features are better than the spatial features from the MPs, MPPR, OMPs, and OMPsM for Sentinel-2 MSIL1C image classification. Furthermore, some previous findings of the ND, NDCB, NDDB, NDRPS, and NDFC frameworks showed superiority to direct multiclass classification, and the ECOC approaches are arguably useful in the Sentinel-2 MSIL1C image classification task. This finding can be explained by the fact that the final classification performance is controlled not only by the robustness of the adopted classifier but also by the discrimination capable of providing data. Additionally, the superiority of the END, ENDRPS, ENDCB, and ENDDDB frameworks over the ND, NDCB, NDDB, and NDRPS frameworks is confirmed, and one can obtain compatible and even better OA results than the direct and ECOC frameworks by using weak and simple classifiers in the END, ENDRPS, ENDCB, and ENDDDB frameworks. For example, END-ERDT can be an alternative to RBF kernel-based SVM in the ECOC framework from the generalized classification accuracy, computationally efficient model training, and easy deployment points of view. Finally, from both greater numbers of species identification and a high classification accuracy point of view, the Sentinel-2A MSIL1C product is more suitable than the global land cover products that are generated from MODIS and Landsat imagery for arid-land vegetation species mapping.

**Author Contributions:** Conceptualization, A.S. and P.D.; Methodology, A.S.; Situ data collection: A.S., Y.G., L.M. and G.I.; Sentinel 2 data collection and processing: A.S.; Land cover type check using VHR image from Google Earth: C.L.; Original draft preparation: A.S.; Review and editing: A.S., N.Y., P.D., and S.L.; Project admission: A.S. and J.A.; Funding: A.S. and J.A.

**Funding:** This work was partially supported by the National Natural Science Foundation of China (grant nos. U1603242, 41601440), the Youth Innovation Promotion Association Foundation of the Chinese Academy of Sciences (2018476), and the West Light Foundation of the Chinese Academy of Sciences (2016-QNXZ-B-11).

**Acknowledgments:** The authors would like to acknowledge the use of free open-access Copernicus Sentinel-2 data (<https://scihub.copernicus.eu>) and the use of SNAP-ESA Sentinel Application Platform v6.0.0 (<http://step.esa.int>).

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

## References

1. Chen, Z.; Elvidge, C.D.; Groeneveld, D.P. Monitoring seasonal dynamics of arid land vegetation using AVIRIS data. *Remote Sens. Environ.* **1998**, *65*, 255–266. [[CrossRef](#)]
2. Clark, J.S.; Bell, D.M.; Kwit, M.C.; Zhu, K. Competition-interaction landscapes for the joint response of forests to climate change. *Glob. Chang. Biol.* **2014**, *20*, 1979–1991. [[CrossRef](#)] [[PubMed](#)]
3. Wu, D.; Zhao, X.; Liang, S.; Zhou, T.; Huang, K.; Tang, B.; Zhao, W. Time-lag effects of global vegetation responses to climate change. *Glob. Chang. Biol.* **2015**, *21*, 3520–3531. [[CrossRef](#)] [[PubMed](#)]
4. Olefeldt, D.; Goswami, S.; Grosse, G.; Hayes, D.; Hugelius, G.; Kuhry, P.; McGuire, A.D.; Romanovsky, V.E.; Sannel, A.B.K.; Schuur, E.A.G.; et al. Circumpolar distribution and carbon storage of thermokarst landscapes. *Nat. Commun.* **2016**, *7*, 13043. [[CrossRef](#)]

5. Fleischer, E.; Khashimov, I.; Hölzel, N.; Klemm, O. Carbon exchange fluxes over peatlands in Western Siberia: Possible feedback between land-use change and climate change. *Sci. Total Environ.* **2016**, *545*, 424–433. [[CrossRef](#)] [[PubMed](#)]
6. Tian, H.; Cao, C.; Chen, W.; Bao, S.; Yang, B.; Myneni, R.B. Response of vegetation activity dynamic to climatic change and ecological restoration programs in Inner Mongolia from 2000 to 2012. *Ecol. Eng.* **2015**, *82*, 276–289. [[CrossRef](#)]
7. Jiang, L.; Bao, A.; Guo, H.; Ndayisaba, F. Vegetation dynamics and responses to climate change and human activities in Central Asia. *Sci. Total Environ.* **2017**, *599*, 967–980. [[CrossRef](#)] [[PubMed](#)]
8. Zhang, Y.; Zhang, C.; Wang, Z.; Chen, Y.; Gang, C.; An, R.; Li, J. Vegetation dynamics and its driving forces from climate change and human activities in the Three-River Source Region, China from 1982 to 2012. *Sci. Total Environ.* **2016**, *563*, 210–220. [[CrossRef](#)]
9. Kerr, J.T.; Ostrovsky, M. From space to species: Ecological applications for remote sensing. *Trends Ecol. Evol.* **2003**, *18*, 299–305. [[CrossRef](#)]
10. Turner, W.; Spector, S.; Gardiner, N.; Fladeland, M.; Sterling, E.; Steininger, M. Remote sensing for biodiversity science and conservation. *Trends Ecol. Evol.* **2003**, *18*, 306–314. [[CrossRef](#)]
11. Madonsela, S.; Cho, M.A.; Ramoelo, A.; Mutanga, O. Remote sensing of species diversity using Landsat 8 spectral variables. *ISPRS J. Photogramm. Remote Sens.* **2017**, *133*, 116–127. [[CrossRef](#)]
12. Gould, W. Remote sensing of vegetation, plant species richness, and regional biodiversity hotspots. *Ecol. Appl.* **2000**, *10*, 1861–1870. [[CrossRef](#)]
13. Barbier, N.; Coutron, P.; Lejoly, J.; Deblauwe, V.; Lejeune, O. Self-organized vegetation patterning as a fingerprint of climate and human impact on semi-arid ecosystems. *J. Ecol.* **2006**, *94*, 537–547. [[CrossRef](#)]
14. Zimmermann, N.E.; Edwards, T.C.; Moisen, G.G.; Frescino, T.S.; Blackard, J.A. Remote sensing-based predictors improve distribution models of rare, early successional and broadleaf tree species in Utah. *J. Appl. Ecol.* **2007**, *44*, 1057–1067. [[CrossRef](#)] [[PubMed](#)]
15. Xie, Y.; Sha, Z.; Yu, M. Remote sensing imagery in vegetation mapping: A review. *J. Plant Ecol.* **2008**, *1*, 9–23. [[CrossRef](#)]
16. Gaitan, J.J.; Oliva, G.E.; Bran, D.E.; Maestre, F.T.; Aguiar, M.R.; Jobbagy, E.G.; Buono, G.G.; Ferrante, D.; Nakamatsu, V.B.; Ciari, G.; et al. Vegetation structure is as important as climate for explaining ecosystem function across Patagonian rangelands. *J. Ecol.* **2014**, *102*, 1419–1428. [[CrossRef](#)]
17. Waltari, E.; Schroeder, R.; McDonald, K.; Anderson, R.P.; Carnaval, A. Bioclimatic variables derived from remote sensing: Assessment and application for species distribution modelling. *Methods Ecol. Evol.* **2014**, *5*, 1033–1042. [[CrossRef](#)]
18. Tian, F.; Brandt, M.; Liu, Y.Y.; Verger, A.; Tagesson, T.; Diouf, A.A.; Rasmussen, K.; Mbow, C.; Wang, Y.; Fensholt, R. Remote sensing of vegetation dynamics in drylands: Evaluating vegetation optical depth (VOD) using AVHRR NDVI and in situ green biomass data over West African Sahel. *Remote Sens. Environ.* **2016**, *177*, 265–276. [[CrossRef](#)]
19. Mildrexler, D.J.; Zhao, M.; Running, S.W. Testing a MODIS global disturbance index across North America. *Remote Sens. Environ.* **2009**, *113*, 2103–2117. [[CrossRef](#)]
20. Pettorelli, N.; Laurance, W.F.; O'Brien, T.G.; Wegmann, M.; Nagendra, H.; Turner, W. Satellite remote sensing for applied ecologists: Opportunities and challenges. *J. Appl. Ecol.* **2014**, *51*, 839–848. [[CrossRef](#)]
21. Sulla-Menashe, D.; Kennedy, R.E.; Yang, Z.; Braaten, J.; Krankina, O.N.; Friedl, M.A. Detecting forest disturbance in the Pacific Northwest from MODIS time series using temporal segmentation. *Remote Sens. Environ.* **2014**, *151*, 114–123. [[CrossRef](#)]
22. McDowell, N.G.; Coops, N.C.; Beck, P.S.; Chambers, J.Q.; Gangodagamage, C.; Hicke, J.A.; Huang, C.Y.; Kennedy, R.; Krofcheck, D.J.; Litvak, M.; et al. Global satellite monitoring of climate-induced vegetation disturbances. *Trends Plant Sci.* **2015**, *20*, 114–123. [[CrossRef](#)] [[PubMed](#)]
23. Rhodes, C.J.; Henrys, P.; Siriwardena, G.M.; Whittingham, M.J.; Norton, L.R. The relative value of field survey and remote sensing for biodiversity assessment. *Methods Ecol. Evol.* **2015**, *6*, 772–781. [[CrossRef](#)]
24. Assal, T.J.; Anderson, P.J.; Sibold, J. Spatial and temporal trends of drought effects in a heterogeneous semi-arid forest ecosystem. *For. Ecol. Manag.* **2016**, *365*, 137–151. [[CrossRef](#)]
25. Harvey, K.R.; Hill, G.J.E. Vegetation mapping of a tropical freshwater swamp in the Northern Territory, Australia: A comparison of aerial photography, Landsat TM and SPOT satellite imagery. *Int. J. Remote Sens.* **2001**, *22*, 2911–2925. [[CrossRef](#)]

26. Brown, M.E.; Pinzón, J.E.; Didan, K.; Morisette, J.T.; Tucker, C.J. Evaluation of the consistency of long-term NDVI time series derived from AVHRR, SPOT-vegetation, SeaWiFS, MODIS, and Landsat ETM+ sensors. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 1787–1793. [[CrossRef](#)]
27. Vieira, M.A.; Formaggio, A.R.; Rennó, C.D.; Atzberger, C.; Aguiar, D.A.; Mello, M.P. Object based image analysis and data mining applied to a remotely sensed Landsat time-series to map sugarcane over large areas. *Remote Sens. Environ.* **2012**, *123*, 553–562. [[CrossRef](#)]
28. Dubovyk, O.; Landmann, T.; Erasmus, B.F.; Tewes, A.; Schellberg, J. Monitoring vegetation dynamics with medium resolution MODIS-EVI time series at sub-regional scale in southern Africa. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *38*, 175–183. [[CrossRef](#)]
29. Anchang, J.Y.; Ananga, E.O.; Pu, R. An efficient unsupervised index based approach for mapping urban vegetation from IKONOS imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *50*, 211–220. [[CrossRef](#)]
30. Su, Y.; Guo, Q.; Fry, D.L.; Collins, B.M.; Kelly, M.; Flanagan, J.P.; Battles, J.J. A vegetation mapping strategy for conifer forests by combining airborne LiDAR data and aerial imagery. *Can. J. Remote Sens.* **2016**, *42*, 15. [[CrossRef](#)]
31. da Silveira, H.L.F.; Galvão, L.S.; Sanches, I.D.A.; de Sá, I.B.; Taura, T.A. Use of MSI/Sentinel-2 and airborne LiDAR data for mapping vegetation and studying the relationships with soil attributes in the Brazilian semi-arid region. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *73*, 179–190. [[CrossRef](#)]
32. Vrieling, A.; Meroni, M.; Darvishzadeh, R.; Skidmore, A.K.; Wang, T.; Zurita-Milla, R.; Oosterbeek, K.; O'Connor, B.; Paganini, M. Vegetation phenology from Sentinel-2 and field cameras for a Dutch barrier island. *Remote Sens. Environ.* **2018**, *215*, 517–529. [[CrossRef](#)]
33. Saveraid, E.H.; Debinski, D.M.; Kindscher, K.; Jakubauskas, M.E. A comparison of satellite data and landscape variables in predicting bird species occurrences in the Greater Yellowstone Ecosystem, USA. *Landsc. Ecol.* **2001**, *16*, 71–83. [[CrossRef](#)]
34. Giri, C.; Ochieng, E.; Tieszen, L.L.; Zhu, Z.; Singh, A.; Loveland, T.; Masek, J.; Duke, N. Status and distribution of mangrove forests of the world using earth observation satellite data. *Glob. Ecol. Biogeogr.* **2011**, *20*, 154–159. [[CrossRef](#)]
35. Kachelriess, D.; Wegmann, M.; Gollock, M.; Pettorelli, N. The application of remote sensing for marine protected area management. *Ecol. Indic.* **2014**, *36*, 169–177. [[CrossRef](#)]
36. Hansen, M.C.; DeFries, R.S.; Townshend, J.R.; Sohlberg, R. Global land cover classification at 1 km spatial resolution using a classification tree approach. *Int. J. Remote Sens.* **2000**, *21*, 1331–1364. [[CrossRef](#)]
37. Bartholome, E.; Belward, A.S. GLC2000: A new approach to global land cover mapping from Earth observation data. *Int. J. Remote Sens.* **2005**, *26*, 1959–1977. [[CrossRef](#)]
38. Tateishi, R.; Uriyangqai, B.; Al-Bilbisi, H.; Ghar, M.A.; Tsend-Ayush, J.; Kobayashi, T.; Kasimu, A.; Hoan, N.T.; Shalaby, A.; Alsaaidh, B.; et al. Production of global land cover data–GLCNMO. *Int. J. Digit. Earth* **2011**, *4*, 22–49. [[CrossRef](#)]
39. Friedl, M.A.; Sulla-Menashe, D.; Tan, B.; Schneider, A.; Ramankutty, N.; Sibley, A.; Huang, X. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sens. Environ.* **2010**, *114*, 168–182. [[CrossRef](#)]
40. Arino, O.; Perez, R.; Ramos Perez, J.J.; Kalogirou, V.; Bontemps, S.; Defourny, P.; Van Bogaert, E. Global land cover map for 2009, European Space Agency (ESA) & Université catholique de Louvain (UCL), PANGAEA. 2012. Available online: <https://doi.pangaea.de/10.1594/PANGAEA.787668> (accessed on 30 July 2019).
41. Chen, J.; Liao, A.; Cao, X.; Chen, L.; Chen, X.; He, C.; Han, G.; Peng, S.; Lu, M.; Zhang, W. Global land cover mapping at 30 m resolution: A POK-based operational approach. *ISPRS J. Photogramm. Remote Sens.* **2015**, *103*, 7–27. [[CrossRef](#)]
42. Hansen, M.C.; Reed, B. A comparison of the IGBP DISCover and University of Maryland 1 km global land cover products. *Int. J. Remote Sens.* **2000**, *21*, 1365–1373. [[CrossRef](#)]
43. Myneni, R.B.; Hoffman, S.; Knyazikhin, Y.; Privette, J.L.; Glassy, J.; Tian, Y.; Wang, Y.; Song, X.; Zhang, Y.; Smith, G.R.; et al. Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. *Remote Sens. Environ.* **2002**, *83*, 214–231. [[CrossRef](#)]
44. Hansen, M.C.; DeFries, R.S.; Townshend, J.R.G.; Carroll, M.; Dimiceli, C.; Sohlberg, R.A. Global percent tree cover at a spatial resolution of 500 meters: First results of the MODIS vegetation continuous fields algorithm. *Earth Interact.* **2003**, *7*, 1–15. [[CrossRef](#)]

45. Ganguly, S.; Friedl, M.A.; Tan, B.; Zhang, X.; Verma, M. Land surface phenology from MODIS: Characterization of the Collection 5 global land cover dynamics product. *Remote Sens. Environ.* **2010**, *114*, 1805–1816. [[CrossRef](#)]
46. Gong, P.; Wang, J.; Yu, L.; Zhao, Y.; Zhao, Y.; Liang, L.; Niu, Z.; Huang, X.; Fu, H.; Liu, S.; et al. Finer resolution observation and monitoring of global land cover: First mapping results with Landsat TM and ETM+ data. *Int. J. Remote Sens.* **2013**, *34*, 2607–2654. [[CrossRef](#)]
47. Tuanmu, M.N.; Jetz, W. A global 1-km consensus land-cover product for biodiversity and ecosystem modelling. *Glob. Ecol. Biogeogr.* **2014**, *23*, 1031–1045. [[CrossRef](#)]
48. Zhang, H.K.; Roy, D.P. Using the 500 m MODIS land cover product to derive a consistent continental scale 30 m Landsat land cover classification. *Remote Sens. Environ.* **2017**, *197*, 15–34. [[CrossRef](#)]
49. Yu, Q.; Hu, Q.; van Vliet, J.; Verburg, P.H.; Wu, W. GlobeLand30 shows little cropland area loss but greater fragmentation in China. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *66*, 37–45. [[CrossRef](#)]
50. Drusch, M.; Del Bello, U.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti, P.; Meygret, A.; et al. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sens. Environ.* **2012**, *120*, 25–36. [[CrossRef](#)]
51. Frampton, W.J.; Dash, J.; Watmough, G.; Milton, E.J. Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. *ISPRS J. Photogramm. Remote Sens.* **2013**, *82*, 83–92. [[CrossRef](#)]
52. Verrelst, J.; Rivera, J.P.; Leonenko, G.; Alonso, L.; Moreno, J. Optimizing LUT-based RTM inversion for semiautomatic mapping of crop biophysical parameters from Sentinel-2 and-3 data: Role of cost functions. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 257–269. [[CrossRef](#)]
53. Käab, A.; Winsvold, S.H.; Altena, B.; Nuth, C.; Nagler, T.; Wuite, J. Glacier remote sensing using Sentinel-2. part I: Radiometric and geometric performance, and application to ice velocity. *Remote Sens.* **2016**, *8*, 598.
54. Novelli, A.; Aguilar, M.A.; Nemmaoui, A.; Aguilar, F.J.; Tarantino, E. Performance evaluation of object based greenhouse detection from Sentinel-2 MSI and Landsat 8 OLI data: A case study from Almería (Spain). *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *52*, 403–411. [[CrossRef](#)]
55. Belgiu, M.; Csillik, O. Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. *Remote Sens. Environ.* **2018**, *204*, 509–523. [[CrossRef](#)]
56. Griffiths, P.; Nendel, C.; Hostert, P. Intra-annual reflectance composites from Sentinel-2 and Landsat for national-scale crop and land cover mapping. *Remote Sens. Environ.* **2019**, *220*, 135–151. [[CrossRef](#)]
57. Chan, J.C.W.; Paelinckx, D. Evaluation of Random Forest and Adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery. *Remote Sens. Environ.* **2008**, *112*, 2999–3011. [[CrossRef](#)]
58. Naidoo, L.; Cho, M.A.; Mathieu, R.; Asner, G. Classification of savanna tree species, in the Greater Kruger National Park region, by integrating hyperspectral and LiDAR data in a Random Forest data mining environment. *ISPRS J. Photogramm. Remote Sens.* **2012**, *69*, 167–179. [[CrossRef](#)]
59. Kumar, P.; Gupta, D.K.; Mishra, V.N.; Prasad, R. Comparison of support vector machine, artificial neural network, and spectral angle mapper algorithms for crop classification using LISS IV data. *Int. J. Remote Sens.* **2015**, *36*, 1604–1617. [[CrossRef](#)]
60. Omer, G.; Mutanga, O.; Abdel-Rahman, E.M.; Adam, E. Performance of support vector machines and artificial neural network for mapping endangered tree species using WorldView-2 data in Dukuduku forest, South Africa. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 4825–4840. [[CrossRef](#)]
61. Olofsson, P.; Foody, G.M.; Herold, M.; Stehman, S.V.; Woodcock, C.E.; Wulder, M.A. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* **2014**, *148*, 42–57. [[CrossRef](#)]
62. Samat, A.; Li, J.; Liu, S.; Du, P.; Miao, Z.; Luo, J. Improved hyperspectral image classification by active learning using pre-designed mixed pixels. *Pattern Recognit.* **2016**, *51*, 43–58. [[CrossRef](#)]
63. Mas, J.F.; Flores, J.J. The application of artificial neural networks to the analysis of remotely sensed data. *Int. J. Remote Sens.* **2008**, *29*, 617–663. [[CrossRef](#)]
64. Mountrakis, G.; Im, J.; Ogole, C. Support vector machines in remote sensing: A review. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 247–259. [[CrossRef](#)]
65. Maulik, U.; Chakraborty, D. Remote sensing image classification: A survey of support-vector-machine-based advanced techniques. *IEEE Geosci. Remote Sens. Mag.* **2017**, *5*, 33–52. [[CrossRef](#)]
66. Samat, A.; Du, P.; Liu, S.; Li, J.; Cheng, L. E2LMs: Ensemble Extreme Learning Machines for Hyperspectral Image Classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 1060–1069. [[CrossRef](#)]

67. Xu, M.; Watanachaturaporn, P.; Varshney, P.K.; Arora, M.K. Decision tree regression for soft classification of remote sensing data. *Remote Sens. Environ.* **2005**, *97*, 322–336. [[CrossRef](#)]
68. Du, P.; Samat, A.; Waske, B.; Liu, S.; Li, Z. Random forest and rotation forest for fully polarized SAR image classification using polarimetric and spatial features. *ISPRS J. Photogramm. Remote Sens.* **2015**, *105*, 38–53. [[CrossRef](#)]
69. 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](#)]
70. Zhang, L.; Zhang, L.; Du, B. Deep learning for remote sensing data: A technical tutorial on the state of the art. *IEEE Geosci. Remote Sens. Mag.* **2016**, *4*, 22–40. [[CrossRef](#)]
71. Zhu, X.X.; Tuia, D.; Mou, L.; Xia, G.S.; Zhang, L.; Xu, F.; Fraundorfer, F. Deep learning in remote sensing: A comprehensive review and list of resources. *IEEE Geosci. Remote Sens. Mag.* **2017**, *5*, 8–36. [[CrossRef](#)]
72. Dietterich, T.G.; Bakiri, G. Solving multiclass learning problems via error-correcting output codes. *J. Artif. Intell. Res.* **1994**, *2*, 263–286. [[CrossRef](#)]
73. Allwein, E.L.; Schapire, R.E.; Singer, Y. Reducing multiclass to binary: A unifying approach for margin classifiers. *J. Mach. Learn. Res.* **2000**, *1*, 113–141.
74. Duarte-Villaseñor, M.M.; Carrasco-Ochoa, J.A.; Martínez-Trinidad, J.F.; Flores-Garrido, M. Nested dichotomies based on clustering. In *Iberoamerican Congress on Pattern Recognition*; Springer: Berlin/Heidelberg, Germany, September 2012; pp. 162–169.
75. Dong, L.; Frank, E.; Kramer, S. Ensembles of balanced nested dichotomies for multi-class problems. In *European Conference on Principles of Data Mining and Knowledge Discovery*; Springer: Berlin/Heidelberg, Germany, October 2005; pp. 84–95.
76. Foody, G.M.; Mathur, A. A relative evaluation of multiclass image classification by support vector machines. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 1335–1343. [[CrossRef](#)]
77. Plaza, A.; Benediktsson, J.A.; Boardman, J.W.; Brazile, J.; Bruzzone, L.; Camps-Valls, G.; Chanussot, J.; Fauvel, M.; Gamba, P.; Gualtieri, A.; et al. Recent advances in techniques for hyperspectral image processing. *Remote Sens. Environ.* **2009**, *113*, S110–S122. [[CrossRef](#)]
78. Shao, Y.; Lunetta, R.S. Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points. *ISPRS J. Photogramm. Remote Sens.* **2012**, *70*, 78–87. [[CrossRef](#)]
79. Hüllermeier, E.; Vanderlooy, S. Combining predictions in pairwise classification: An optimal adaptive voting strategy and its relation to weighted voting. *Pattern Recognit.* **2010**, *43*, 128–142. [[CrossRef](#)]
80. Passerini, A.; Pontil, M.; Frasconi, P. New results on error correcting output codes of kernel machines. *IEEE Trans. Neural Netw.* **2004**, *15*, 45–54. [[CrossRef](#)]
81. Pujol, O.; Radeva, P.; Vitria, J. Discriminant ECOC: A heuristic method for application dependent design of error correcting output codes. *IEEE Trans. Pattern Anal. Mach. Intell.* **2006**, *28*, 1007–1012. [[CrossRef](#)]
82. Escalera, S.; Pujol, O.; Radeva, P. On the decoding process in ternary error-correcting output codes. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 120–134. [[CrossRef](#)]
83. Pal, M. Class decomposition Approaches for land cover classification: A comparative study. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2006, Denver, CO, USA, 31 July–4 August 2006*; pp. 2731–2733.
84. Mera, D.; Fernández-Delgado, M.; Cotos, J.M.; Viqueira, J.R.R.; Barro, S. Comparison of a massive and diverse collection of ensembles and other classifiers for oil spill detection in sar satellite images. *Neural Comput. Appl.* **2017**, *28*, 1101–1117. [[CrossRef](#)]
85. Frank, E.; Kramer, S. Ensembles of nested dichotomies for multi-class problems. In *Proceedings of the Twenty-First International Conference on Machine Learning, Banff, AB, Canada, 4–8 July 2004*; p. 39.
86. Rodríguez, J.J.; García-Osorio, C.; Maudes, J. Forests of nested dichotomies. *Pattern Recognit. Lett.* **2010**, *31*, 125–132. [[CrossRef](#)]
87. Quinlan, J.R. Bagging, boosting, and C4. 5. In *Proceedings of the AAAI'96 Proceedings of the Thirteenth National Conference on Artificial Intelligence, Portland, OR, USA, 4–8 August 1996; Volume 1*, pp. 725–730.
88. Rätsch, G.; Onoda, T.; Müller, K.R. Soft margins for AdaBoost. *Mach. Learn.* **2001**, *42*, 287–320. [[CrossRef](#)]
89. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
90. Rodríguez, J.J.; Kuncheva, L.I.; Alonso, C.J. Rotation forest: A new classifier ensemble method. *IEEE Trans. Pattern Anal. Mach. Intell.* **2006**, *28*, 1619–1630. [[CrossRef](#)] [[PubMed](#)]

91. Geurts, P.; Ernst, D.; Wehenkel, L. Extremely randomized trees. *Mach. Learn.* **2006**, *63*, 3–42. [[CrossRef](#)]
92. Cortes, C.; Vapnik, V. Support vector machine. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
93. Fauvel, M.; Tarabalka, Y.; Benediktsson, J.A.; Chanussot, J.; Tilton, J.C. Advances in spectral-spatial classification of hyperspectral images. *Proc. IEEE* **2013**, *101*, 652–675. [[CrossRef](#)]
94. Li, M.; Zang, S.; Zhang, B.; Li, S.; Wu, C. A review of remote sensing image classification techniques: The role of spatio-contextual information. *Eur. J. Remote Sens.* **2014**, *47*, 389–411. [[CrossRef](#)]
95. Chen, Y.; Zhao, X.; Jia, X. Spectral-spatial classification of hyperspectral data based on deep belief network. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 2381–2392. [[CrossRef](#)]
96. He, L.; Li, J.; Liu, C.; Li, S. Recent advances on spectral-spatial hyperspectral image classification: An overview and new guidelines. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 1579–1597. [[CrossRef](#)]
97. Liao, W.; Chanussot, J.; Dalla Mura, M.; Huang, X.; Bellens, R.; Gautama, S.; Philips, W. Taking Optimal Advantage of Fine Spatial Resolution: Promoting partial image reconstruction for the morphological analysis of very-high-resolution images. *IEEE Geosci. Remote Sens. Mag.* **2017**, *5*, 8–28. [[CrossRef](#)]
98. Samat, A.; Persello, C.; Liu, S.; Li, E.; Miao, Z.; Abuduwaili, J. Classification of VHR Multispectral Images Using ExtraTrees and Maximally Stable Extremal Region-Guided Morphological Profile. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 3179–3195. [[CrossRef](#)]
99. Samat, A.; Liu, S.; Persello, C.; Li, E.; Miao, Z.; Abuduwaili, J. Evaluation of ForestPA for VHR RS image classification using spectral and superpixel-guided morphological profiles. *Eur. J. Remote Sens.* **2019**, *52*, 107–121. [[CrossRef](#)]
100. Blaschke, T. Object based image analysis for remote sensing. *ISPRS J. Photogramm. Remote Sens.* **2010**, *65*, 2–16. [[CrossRef](#)]
101. Blaschke, T.; Hay, G.J.; Kelly, M.; Lang, S.; Hofmann, P.; Addink, E.; Queiroz Feitosa, R.; van der Meer, F.; van der Werff, H.; Tiede, D.; et al. Geographic object-based image analysis—towards a new paradigm. *ISPRS J. Photogramm. Remote Sens.* **2014**, *87*, 180–191. [[CrossRef](#)] [[PubMed](#)]
102. Ma, L.; Li, M.; Ma, X.; Cheng, L.; Du, P.; Liu, Y. A review of supervised object-based land-cover image classification. *ISPRS J. Photogramm. Remote Sens.* **2017**, *130*, 277–293. [[CrossRef](#)]
103. Kezer, K.; Matsuyama, H. Decrease of river runoff in the Lake Balkhash basin in Central Asia. *Hydrol. Process. Int. J.* **2006**, *20*, 1407–1423. [[CrossRef](#)]
104. Propastin, P.A. Simple model for monitoring Balkhash Lake water levels and Ili River discharges: Application of remote sensing. *Lakes Reserv. Res. Manag.* **2008**, *13*, 77–81. [[CrossRef](#)]
105. Propastin, P. Problems of water resources management in the drainage basin of Lake Balkhash with respect to political development. In *Climate Change and the Sustainable Use of Water Resources*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 449–461.
106. Petr, T.; Mitrofanov, V.P. The impact on fish stocks of river regulation in Central Asia and Kazakhstan. *Lakes Reserv. Res. Manag.* **1998**, *3*, 143–164. [[CrossRef](#)]
107. Bai, J.; Chen, X.; Li, J.; Yang, L.; Fang, H. Changes in the area of inland lakes in arid regions of central Asia during the past 30 years. *Environ. Monit. Assess.* **2011**, *178*, 247–256. [[CrossRef](#)]
108. Klein, I.; Gessner, U.; Kuenzer, C. Regional land cover mapping and change detection in Central Asia using MODIS time-series. *Appl. Geogr.* **2012**, *35*, 219–234. [[CrossRef](#)]
109. Chen, X.; Bai, J.; Li, X.; Luo, G.; Li, J.; Li, B.L. Changes in land use/land cover and ecosystem services in Central Asia during 1990. *Curr. Opin. Environ. Sustain.* **2013**, *5*, 116–127. [[CrossRef](#)]
110. De Beurs, K.M.; Henebry, G.M.; Owsley, B.C.; Sokolik, I. Using multiple remote sensing perspectives to identify and attribute land surface dynamics in Central Asia 2001. *Remote Sens. Environ.* **2015**, *170*, 48–61. [[CrossRef](#)]
111. Zhang, C.; Lu, D.; Chen, X.; Zhang, Y.; Maisupova, B.; Tao, Y. The spatiotemporal patterns of vegetation coverage and biomass of the temperate deserts in Central Asia and their relationships with climate controls. *Remote Sens. Environ.* **2016**, *175*, 271–281. [[CrossRef](#)]
112. Leathart, T.; Pfahringer, B.; Frank, E. Building ensembles of adaptive nested dichotomies with random-pair selection. In *Joint European Conference on Machine Learning and Knowledge Discovery in Data Bases*; Springer: Cham, Switzerland, 2016; pp. 179–194.
113. Leathart, T.; Frank, E.; Pfahringer, B.; Holmes, G. Ensembles of Nested Dichotomies with Multiple Subset Evaluation. *arXiv* **2018**, arXiv:1809.02740.

114. Wever, M.; Mohr, F.; Hüllermeier, E. Ensembles of evolved nested dichotomies for classification. In Proceedings of the Genetic and Evolutionary Computation Conference, Kyoto, Japan, 15–19 July 2018; pp. 561–568.
115. Myint, S.W.; Gober, P.; Brazel, A.; Grossman-Clarke, S.; Weng, Q. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sens. Environ.* **2011**, *115*, 1145–1161. [[CrossRef](#)]
116. Drăguț, L.; Eisank, C. Automated object-based classification of topography from SRTM data. *Geomorphology* **2012**, *141*, 21–33. [[CrossRef](#)]
117. Drăguț, L.; Csillik, O.; Eisank, C.; Tiede, D. Automated parameterisation for multi-scale image segmentation on multiple layers. *ISPRS J. Photogramm. Remote Sens.* **2014**, *88*, 119–127. [[CrossRef](#)]
118. Benz, U.C.; Hofmann, P.; Willhauck, G.; Lingenfelder, I.; Heynen, M. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS J. Photogramm. Remote Sens.* **2004**, *58*, 239–258. [[CrossRef](#)]
119. Kim, M.; Warner, T.A.; Madden, M.; Atkinson, D.S. Multi-scale GEOBIA with very high spatial resolution digital aerial imagery: Scale, texture and image objects. *Int. J. Remote Sens.* **2011**, *32*, 2825–2850. [[CrossRef](#)]
120. Drăguț, L.; Tiede, D.; Levick, S.R. ESP: A tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *Int. J. Geogr. Inf. Sci.* **2010**, *24*, 859–871. [[CrossRef](#)]
121. Benediktsson, J.A.; Palmason, J.A.; Sveinsson, J.R. Classification of hyperspectral data from urban areas based on extended morphological profiles. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 480–491. [[CrossRef](#)]
122. Fauvel, M.; Benediktsson, J.A.; Chanussot, J.; Sveinsson, J.R. Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 3804–3814. [[CrossRef](#)]
123. Dalla Mura, M.; Villa, A.; Benediktsson, J.A.; Chanussot, J.; Bruzzone, L. Classification of hyperspectral images by using extended morphological attribute profiles and independent component analysis. *IEEE Geosci. Remote Sens. Lett.* **2011**, *8*, 542–546. [[CrossRef](#)]
124. Plaza, A.; Martinez, P.; Perez, R.; Plaza, J. A new approach to mixed pixel classification of hyperspectral imagery based on extended morphological profiles. *Pattern Recognit.* **2004**, *37*, 1097–1116. [[CrossRef](#)]
125. Aptoula, E.; Lefèvre, S. A comparative study on multivariate mathematical morphology. *Pattern Recognit.* **2007**, *40*, 2914–2929. [[CrossRef](#)]
126. Samat, A.; Gamba, P.; Liu, S.; Miao, Z.; Li, E.; Abuduwaili, J. Quad-PolSAR data classification using modified random forest algorithms to map halophytic plants in arid areas. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *73*, 503–521. [[CrossRef](#)]
127. Snoek, J.; Larochelle, H.; Adams, R.P. Practical bayesian optimization of machine learning algorithms. In Proceedings of the Advances in Neural Information Processing Systems, Lake Tahoe, NV, USA, 3–6 December 2012; pp. 2951–2959, Curran Associates.
128. Du, P.; Xia, J.; Zhang, W.; Tan, K.; Liu, Y.; Liu, S. Multiple classifier system for remote sensing image classification: A Review. *Sensors (Basel)* **2012**, *12*, 4764–4792. [[CrossRef](#)]
129. Samat, A.; Gamba, P.; Du, P.; Luo, J. Active extreme learning machines for quad-polarimetric SAR imagery classification. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *35*, 305–319. [[CrossRef](#)]
130. Vihervaara, P.; Auvinen, A.P.; Mononen, L.; Törmä, M.; Ahlroth, P.; Anttila, S.; Böttcher, K.; Forsius, M.; Heino, J.; Koskelainen, M.; et al. How essential biodiversity variables and remote sensing can help national biodiversity monitoring. *Glob. Ecol. Conserv.* **2017**, *10*, 43–59. [[CrossRef](#)]
131. Gholizadeh, H.; Gamon, J.A.; Zygielbaum, A.I.; Wang, R.; Schweiger, A.K.; Cavender-Bares, J. Remote sensing of biodiversity: Soil correction and data dimension reduction methods improve assessment of  $\alpha$ -diversity (species richness) in prairie ecosystems. *Remote Sens. Environ.* **2018**, *206*, 240–253. [[CrossRef](#)]
132. Gholizadeh, H.; Gamon, J.A.; Townsend, P.A.; Zygielbaum, A.I.; Helzer, C.J.; Hmimina, G.Y.; Moore, R.M.; Schweiger, A.K.; Cavender-Bares, J. Detecting prairie biodiversity with airborne remote sensing. *Remote Sens. Environ.* **2019**, *221*, 38–49. [[CrossRef](#)]

