



Article Classifying Forest Types over a Mountainous Area in Southwest China with Landsat Data Composites and Multiple Environmental Factors

Ruonan Li ^{1,2,*}, Panfei Fang ¹, Weiheng Xu ^{3,4}, Leiguang Wang ^{3,4}, Guanglong Ou ⁵, Wanqiu Zhang ¹ and Xin Huang ⁶

- ¹ Faculty of Forestry, Southwest Forestry University, Kunming 650024, China; panfeiFang@swfu.edu.cn (P.F.); wanqiu_mou@hotmail.com (W.Z.)
- ² Urban and Rural Planning and Design Institute of Fuyang, Fuyang 236000, China
- ³ Institutes of Big Data and Artificial Intelligence, Southwest Forestry University, Kunming 650024, China; xwh@swfu.edu.cn (W.X.); wlgbain@126.com (L.W.)
- ⁴ Key Laboratory for Forestry and Ecological Big Data of National Forestry and Grassland Administration, Southwest Forestry University, Kunming 650024, China
- ⁵ Key Laboratory of State Forestry Administration on Biodiversity Conservation in Southwest China, Kunming 650024, China; olg2007621@163.com
- ⁶ School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China; xhuang@whu.edu.cn
- Correspondence: liruonan02@outlook.com

Abstract: Accurate information about forest type and distribution is critical for many scientific applications. It is possible to make a forest type map from the satellite data in a cost effective way. However, forest type mapping over a large and mountainous geographic area is still challenging, due to complex forest type compositions, spectral similarity among various forest types, poor quality images with clouds or cloud shadows and difficulties in managing and processing large amount data. Based on the Google Earth Engine (GEE) cloud platform, a method of forest types mapping using Landsat-8 OLI imagery and multiple environmental factors was developed and tested within Yunnan Province (about 390,000 km²) of China. The proposed approach employed a pixel-based seasonal image compositing method to produce two types of seasonal composite images, i.e., four 7-spectralband composite images and four 5-VI-band composite images associated in spring, summer, autumn, and winter. Then, single-season feature bands and multi-seasonal feature bands were combined with the feature bands of topography, temperature, and precipitation, respectively, and resulting in 17 feature combinations. Finally, using a random forest (RF) classifier, 17 feature combinations were separately experimented to classify the forest type over the study area. The study area was firstly classified into the forest and the non-forest, and then the forest was sub-classified into five forest types (evergreen needleleaf forest, deciduous needleleaf forest, evergreen broadleaf forest, deciduous broadleaf forest, and mixed forest). The results showed that the pixel-based multi-seasonal median composite can produce a cloud-free image for the entire region and is suitable for forest type mapping. Compared with a single-season composite, a multi-seasonal composite can distinguish different forest types more effectively. The environmental factors also improve the accuracy of forest type mapping. With the ground survey samples as reference values, the classification performance of 17 feature combinations was compared, and the optimal feature combination was found out. For the optimal feature combination, its overall accuracy of the forest/non-forest cover map and the forest type map reached 97.57% (Kappa = 0.950) and 70.30% (Kappa = 0.628), respectively. The proposed approach has demonstrated strong potential of high classification accuracy and convenient calculation when mapping forest types over a national or global scale, and its product of 30 m resolution forest type map is capable of contributing to forest resource management.

Keywords: forest type classification; multi-seasonal image composite; Landsat; Google Earth Engine; random forest



Citation: Li, R.; Fang, P.; Xu, W.; Wang, L.; Ou, G.; Zhang, W.; Huang, X. Classifying Forest Types over a Mountainous Area in Southwest China with Landsat Data Composites and Multiple Environmental Factors. *Forests* **2022**, *13*, 135. https://doi.org/ 10.3390/f13010135

Academic Editors: Panayotis Dimopoulos and Ioannis P. Kokkoris

Received: 26 November 2021 Accepted: 6 January 2022 Published: 17 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).

1. Introduction

As the largest biological resource bank on the Earth, forests cover about 30% of the land surface and play an irreplaceable role in mitigating climate change, improving the environment, and ensuring ecological security [1]. Through increasing carbon storage and water holding capacity and decreasing soil erosion, the ecological effects of forests contribute to the biodiversity of the Earth [2]. At the same time, the ecological function of forest ecosystems varies with forest types, and the extent of different forest types is significant for the ecological effectiveness of forests [3]. Yunnan Province, located in the southwest of China, is rich in forest resources, and forest covers 65.04% of the whole area [4]. Because of the rugged terrain and diverse climates, Yunnan Province owns an extremely rich biodiversity and a wide variety of vegetation types in China [5]. Therefore, information about the distribution of different forest types in this area is important for a variety of forest research.

The ground survey is a traditional method to collect forest stand and composition information. Usually, forest field inventory data are collected and aggregated by measuring sampling plots, and therefore the final ground survey data are generally at the forest stand level [6]. However, most forest areas, especially in mountainous regions, are often remote and even inaccessible [7], and thus the ground survey is costly, time-consuming, and spatially restricted [6]. In recent decades, remote sensing technology has become an important and efficient approach to obtain forest spatial distribution information at the regional and national levels in a more efficient way and it allows a higher update frequency [6,8]. Thanks to the free access to Landsat and Sentinel, as well as rapid improvements in data-storage and computing capabilities, a Chinese forest type product (Forest2010) [9] and several global land-cover products (FROM_GLC2010 [10], FROM_GLC2015 [11], GlobeLand30 [12] and FROM_GLC10 [13]) at fine spatial resolution (30 m and 10 m) have been successfully developed. However, the current main source of 30 m forest type data is global land cover products, and these land cover products did not focus on forest type classification and do not make better use of phenological information, which limits the research and application of forest type classification.

In the research of forest type mapping, phenological characteristics of forests are often used as useful features to discriminate forest types. Phenology includes obvious processes such as the coloring of leaves in deciduous temperate forests in autumn due to leaf senescence, the intense green colors of fresh leaves and needles in springtime as well as flowering events [14]. The seasonal variation of visible spectral response in forest species and the phenological differences in senescence among tree species could present unique forest classification opportunities [15]. The research on forest type classification using phenological characteristics is mainly divided into three types: (1) Identify one forest type by the key phenology phases, which are extracted through analyzing the vegetation index time series of different land cover types [16]. (2) Map multiple forest types by collecting a set of cloudless images from specific phenological windows, which show different phenological characteristics for different types. For this type of research, it not only requires good prior knowledge about the phenological characteristics of different forests, but also faces challenges to acquire cloudless imagery for specific phenological events, especially in areas with frequent cloud cover [8,15]. (3) Map multiple forest types using spectral information from time series data. This type of research does not require image selection and good prior knowledge about the phenological characteristics of different forests. For instance, Li et al. used forest phenological characteristics from MODIS EVI time-series data to map forests of China [9]. However, the methods, which use dense timeseries data to map forest types over large area, usually only use one or several vegetation index time-series to classify forests without considering the original band features; and they will lose the multi-seasonal spectral band information that may be more effective for forest classification. In addition, when mapping forest types on a large scale, especially

in large mountainous areas, clouds and rain often reduce image quality and then greatly impact the accuracy of forest type mapping.

In previous studies, topographic factors have been proven to improve the accuracy of forest type classification [6,17]. Topographic factors can be represented by DEM or variables derived from a DEM [18]. In addition, climate plays an important role in the dynamic change of forests, so the impact of climate factors on forest type could not be neglected especially in mountainous areas [8]. Among many climate factors, temperature and precipitation are very critical for local forest types [19]. However, climate factors were less considered in previous studies because climate factors with low spatial resolution, e.g., precipitation, cannot be measured or deduced easily. In addition, forest type mapping over a large scale requires powerful geo-big data processing capabilities. The production of products over a large scale implemented on a supercomputer or processor PC is timeconsuming and laborious and is not easy to update in time. Google Earth Engine (GEE) is a planetary-scale platform for large-scale Earth observation data and geographical analysis, and it can provide massive geospatial data and machine learning algorithms [20]. GEE has the entire Landsat archive along with many free raster datasets from NASA, European Space Agency (ESA), and other imagery [21]. The remote sensing cloud computing platforms can effectively solve limitations such as data collection, managing, and processing [20,22]. GEE has been a study platform for many studies [23–29]. There are also many studies on forest-type mapping based on GEE, such as extracting bamboo using Landsat time series [30] and mapping tree species using Sentinel-2 imagery [31]. However, there are still very limited ways to combine multiple environmental factors with a cloud platform to map forest types in large mountainous areas using multi-seasonal composites of one year of Landsat imagery [6].

In this context, to meet the needs of various fields for forest type maps over a large area, this study aims to explore methods of mapping forest types over large, complicated mountain areas by using one year of Landsat imagery and multiple environmental factors, based on the Google Earth Engine cloud platform. The study tries to address the impact of multi-seasonal and climate-related variables and determine the optimal feature combination for the forest type classification.

2. Materials and Methods

2.1. Study Area

As shown in Figure 1a, the study area is situated in the southwest of China, Yunnan Province. With a substantially large area of 390,000 km² and 16 administrative divisions, Yunnan Province stretches over more than 8 latitudes between tropical and subtropical climate zones. This region is featured as a mountainous plateau and has diversified terrain types. The whole terrain inclines from north to the south and southeast [32], with an elevation difference of more than 6000 m (Figure 1b). Different elevational belts can be found within the study area (montane, alpine and sub-alpine zone, and basin). The climate of Yunnan belongs to the plateau-type tropical monsoon climate, with annual average temperature of 15° and annual mean precipitation of over 1000 mm [33]. However, the distribution of annual precipitation dramatically decreases from west to east, with rainy areas in the west and south and dry areas in the northwest (Figure 1d). Due to its complex and varied terrain, the inner local climates of Yunnan vary greatly, including hot-dry, tropical, and subtropical, temperate, and cold climates.



Figure 1. The geography of the study area. (**a**) the location; (**b**) the elevation; (**c**) the annual average temperature (MOD11A2 V6 product [34]); and (**d**) the annual total precipitation (OpenLandMap Precipitation monthly product [35]).

2.2. Data

2.2.1. Landsat OLI Imagery

The research used the Landsat 8 surface reflectance (SR) Tier 1 images archived in the GEE platform (Table 1). According to the path/row of Landsat data, Yunnan Province is fully covered by 35 scenes of Landsat images. Given the frequent cloud coverage, especially in a mountainous region, all the available Landsat 8 images from 1 March 2017 to 1 March 2018 were collected to ensure that each pixel location has as much cloud-free data as possible. In total, 530 valid images were collected during the study period, with 155 images in spring, 91 images in summer, 133 images in autumn, and 151 images in winter (Figure 2).

Dataset	GEE ID	Dataset Provider	Period	Spatial Resolution
USGS Landsat 8 Level 2, Collection 2, Tier 1	LANDSAT/LC08/C02/T1_L2	USGS	1 March 2017–1 March 2018	30 m
OpenLandMap Precipitation monthly	OpenLandMap/CLM/CLM_ PRECIPITATION_SM2RAIN_M/v01	EnvirometriX Ltd.	2007–2019	1 Km
MOD11A2.006 Terra Land Surface Temperature and Emissivity 8-Day Global 1km	MODIS/006/MOD11A2	NASA LP DAAC at the USGS EROS Center	1 March 2017–1 March 2018	1 Km
SRTM Digital Elevation Data	USGS/SRTMGL1_003	NASA/USGS/JPL-Caltech	2000	30 m





Figure 2. The number of collected Landsat 8 surface reflectance images in 12 months with five cloud cover percent levels. Spring is from April to May; Summer is from June to August; Autumn is from September to November; Winter is from December to February of next year (Spring: 155, Summer: 91, Autumn: 133, Winter: 151).

Because Yunnan is located between tropical and subtropical zones and has a large area, there are few cloud-free Landsat images available in the growing season. Figure 2 shows the number of images with different cloud cover percentage levels in each month from 1 March 2017 to 1 March 2018. Images with a cloud coverage of less than 10% are mainly from January to May and November to December. The five months from June to October are the rainy season in Yunnan, so the images available in this period are less than the rest of the months because some poor-quality images are taken as invalid. It is so cloudy from June to October in Yunnan that most images in this period have more than 50% cloud coverage and few have less than 10% cloud coverage and this makes it difficult to capture the phenological change of forests during the growing season, which usually last from June to September in Yunnan.

2.2.2. Environmental Factors

The growth of forest relies on environmental conditions, so forest types are closely related to environmental factors. Adding environmental factors into remote sensing classification is bound to improve the accuracy of forest type mapping. Many studies have shown that topographic data, especially elevation, slope, and aspect, could increase the estimation accuracy on the land cover when classifying with spectral bands. At the same time, temperature and precipitation are critical for the growth of forests. Some studies have shown that rising temperatures could lead to a significant advancement of the spring phenology of plants in temperate and frigid zones, leading to a prolonged plant growth season [36]. Precipitation indirectly affects forest phenology by changing accumulated temperature and radiation. Hence, environmental factors, including terrain factors (elevation, slope, and aspect), temperature, and precipitation were also considered in the study.

However, it is worth noting that the temperature and precipitation data used in the study both are 1-km spatial resolution data products (Table 1). In addition, the temperature data were introduced from the MOD11A2 V6 product [34], which provides an average 8-day land surface temperature (LST). The monthly precipitation dataset used in the research was from OpenLandMap Precipitation monthly product [35]. Here, topographic data were derived from the Shuttle Radar Topography Mission (SRTM) with 30 m resolution [37]. In addition, the two environmental factors were obtained and resampled to 30 m resolution by the GEE platform.

2.2.3. Select of Sample Site

In order to have a reasonable coverage of samples, samples were collected from eight sites through the evenly distributed sampling scheme, as shown in Figure 3. The spatial locations of the eight sites are roughly located in the east, southeast, south, west, northwest, north, northeast, and central of Yunnan Province, and they include all climate types from the cold temperate zone to subtropical zone. There are significant differences in temperature, precipitation, and altitude among eight sites.

There are multiple climate types in the eight counties (cities) due to the complex geographic environment. In terms of temperature, the monthly average temperatures of site 5 and site 6 are significantly lower than other sites; in particular, the average monthly temperature is in the range of 12.9–22.2 °C; and the average monthly temperature of other sites is in the range of 14.2–31.4 °C. From Figure 3d, it can be said that the precipitation of site 2, site 3, and site 5 are higher than that of other regions in most months of a year; the monthly averages of precipitation of these three regions are extremely uneven in different seasons, with 208–324 mm in July and only 15–43 mm in February. Moreover, the monthly average precipitation of site 6, with a range of 3 mm (December)–186 mm (July), is lower than other sites. The climate of site 6 is relatively dry and cold. The eight sites also have an obvious difference in altitude. Compared with other sites, site 1 has the smallest altitude differences between the lowest and highest altitudes at site 2, site 3, site 5, site 7, and site 8 are in the range of 1776–2974 m, and the altitude differences at site 4 and site 6 are 3434 m and 3894 m, respectively.

The eight sites include evergreen needleleaf forests, deciduous needleleaf forests, evergreen broadleaf forests, deciduous broadleaf forests, and mixed forests. The tree species of evergreen coniferous forests in Yunnan Province mainly include spruce (Picea asperata Mast.), fir (Abies fabr (Mast.) Craib), cypress (Cupressus funebris Endl.), Chinese white pine (Pinus armandii Franch.), Yunnan pine (Pinus yunnanensis), sikang pine (Pinus densata Mast.), khasi pine (Pinus kesiya var. langbianensis), and so on. Deciduous broadleaf forest is the most common forest type in the temperate zone. Evergreen broad-leaved forest is also called Sub-tropical evergreen broad-leaved forest because it occurs in the areas dominated by a subtropical monsoon climate [38,39]. The evergreen broad-leaved forests in Yunnan distribute from 800 to about 3000 m and most are abundant between 1100 and 2700 m. In Yunnan, evergreen broad-leaved forest almost covers the entire tropical mountains and the subtropical area, except for the alpine and subalpine areas above 3000 m asl in northwestern Yunnan [39]. The distribution of deciduous broad-leaved forests is widely distributed, mainly distributed in the area north of 23°39′ N, but the area is small and sporadic [39]; this type of forest is all deciduous in winter, with obvious seasonal changes in the community physiognomy. Mixed forests occur in the transition zone between the evergreen broad-leaved forest belt and the subalpine coniferous forest, mainly in the mountainous areas of 2400 m and 3100 m [40]; this type of forest is concentrated on both sides of ditches and gentle slopes within a small area.



Figure 3. The spatial distribution of samples and the location, temperature, precipitation, and elevation characteristics of eight sites for sample collection. (**a**) 5321 sample plots for the forest/the non-forest are depicted in Section 2.3.2; (**b**) 8701 sample pixels for five forest types are depicted in Section 2.3.2. ENF means evergreen needleleaf forest; DNF means deciduous needleleaf forest; EBF means evergreen broadleaf forest; DBF means deciduous broadleaf forest; MF means mixed forest; (**c**) monthly average temperature; (**d**) monthly average precipitation; (**e**) minimum and maximum altitude. Eight black-border regions are the location of eight administrative counties (cities) for sample collection. Eight red rectangles are the case regions used for comparison.

2.3. Methods

The methodological workflow of this study consists of the following steps (Figure 4): (1) extracting seven spectral reflectance bands and five vegetation indexes for all the images; (2) extracting three environmental factors; (3) implementing seasonal image composite; (4) building 17 feature combinations using all existing features; (5) implementing forest/non-forest classification and forest type classification using RF classifier based on the 17 feature combinations; (6) accuracy assessment and comparison of 17 results; (7) improving forest type classification accuracy based on the feature combination with the best accuracy in (6); (8) analyzing the results using the information of RF feature importance and feature group ranking; (9) comparing the forest/non-forest map and the forest type map with four public reference products, Forest2010, FROM_GLC2015, GLC_FCS2020 and MCD12Q1.



Figure 4. Methodological workflow of forest classification using a Random Forest classifier based on the Google Earth Engine cloud platform.

2.3.1. Classification Scheme

Due to the diverse climate types and topographical conditions in Yunnan Province, its forest has a complex stand structure and environmental conditions, which makes it difficult to improve the accuracy of forest types classification. The spectral characteristics of forest types and other vegetation types (such as crops, shrubs and grasslands) are easy to be confused in the classification of land cover. In addition, the spectral characteristics of forest types and non-vegetation land cover types (such as bare land, construction land, water body, snow and cloud) are obviously different, which is easy to classify. To reduce the impact of other land cover types on forest classification, the study adopts two-level classification and distinguishes forests and non-forests firstly.

For the classification of forest type, the forest areas were further classified into 5 types, including evergreen needleleaf forest (ENF), deciduous needleleaf forest (DNF), evergreen broadleaf forest (EBF), deciduous broadleaf forest (DBF), and mixed forest (MF). The definitions of the five forest types are drawn from the International Geosphere-Biosphere Program (IGBP) classification schemes [41]. Table 2 lists the forest type classification scheme adopted by the study.

First-Level Class	Second-Level Class	Abbreviation	Description
	Evergreen Needleleaf Forests	ENF	Dominated by evergreen conifer trees (canopy > 2 m). Tree cover >60%.
Forest	Evergreen Broadleaf Forests	EBF	Dominated by evergreen broadleaf and palmate trees (canopy > 2 m). Tree cover >60%.
	Deciduous Needleleaf Forests	DNF	Dominated by deciduous needleleaf (larch) trees (canopy > 2 m). Tree cover >60%.
	Deciduous Broadleaf Forests	DBF	Dominated by deciduous broadleaf trees (canopy > 2 m). Tree cover >60%.
	Mixed Forests	MF	Dominated by neither deciduous nor evergreen (40–60% of each) tree type (canopy > 2 m). Tree cover >60%.
Non-forest	-	-	

Table 2. The classification scheme of this study, and the definitions of forest types are drawn from the International Geosphere-Biosphere Program (IGBP) classification schemes.

2.3.2. Sampling Strategy

Supervised classification methods rely on training samples [42]. The influence of training sample size on classification accuracy is greater than that of the adopted algorithm [42,43]. Thus, the size and quality of training samples are the key to classification [42]. Representative samples were collected in the above eight counties (cities) (Figure 3) from Chinese Forest Management Inventory (FMI) data and high-resolution images from Google Earth. Chinese FMI is a ground survey of every ten years at forest stand level. FMI collects forest information about forest management type, dominant tree, average height, the average diameter at breast height, age, canopy density, topographic factors, and soil type, etc. [1,44]. FMI is extremely laborious and time-consuming, and the latest one was carried out in 2016. Here, it is assumed that there were no dramatic changes in the forest types from the year of FMI data (2016) to the year of acquisition of remote sensing images (March 2017–March 2018). Moreover, with exhaustive comparison with high-spatial-resolution images from Google Earth, the quality of reference samples can be guaranteed.

Based on the consideration of spectral characteristics of different classes involved in the classification scheme (Table 2), two levels of samples with different sample size were collected in the research. For the first level, the non-forest classes consisted of multiple land cover types, which involved multiple spectral. Therefore, samples of non-forest types are needed to fully mix the spectra of multiple land cover types, and a large sample plot with 300 m \times 300 m was employed. For the forest-type classification level, the aim is the division of internal natural attributes of the forest land. The sample unit with a larger coverage area may lead to more mixed pixels, which represent the spectral mixture of different forest types. Hence, 30 m \times 30 m sample pixels, rather than 300 m \times 300 m, were involved in the classification of forest type.

The 300 m \times 300 m sample plots were produced by buffering sample points that generated randomly and distributed evenly in the eight sites (Figure 3a). Each sample plot was reviewed and categorized into the forest and the non-forest using high-resolution images from Google Earth. Sample plots that caused the problems with spectral confusion would be removed or manually modified. Sample pixels for five forest types were also generated randomly, and the forest type of each pixel was specified according to the dominant tree information of FMI. An internal of at least 90 m was set for adjacent sample pixels (3 pixels at least on Landsat 8 images). At the same time, sample pixels located within 600 m (20 pixels) of the adjacent area of different forest types were eliminated. The remaining sample pixels were also reviewed using field work data, high spatial resolution Google Earth images, and the Landsat NDVI time series curve. Unfortunately, it is not easy to check every forest type sample due to the huge quantity of samples and only parts of the forest type sample pixels were randomly selected for checking. Finally, the remaining sample plots and pixels were randomly divided into training and validation sets according

to the ratio of 7 to 3, respectively (Table 3). There is no overlap between the training sample and validation sample sets.

Table 3. The number of samples in the two classification levels. The sample unit are with size of 300-m by 300-m for the forest/non-forest classification and 30-m by 30-m for the forest type classification.

Class Name	Forest/Non-Forest (Plot)				Forest Types (Pixel)				
	Non-Forest	Forest	Total	ENF	DNF	EBF	DBF	MF	Total
Train Samples	1750	1988	3738	1240	914	1315	1346	1478	6293
Test Samples	772	861	1633	451	441	530	484	502	2408

2.3.3. Image Composite and Feature Extraction

Although the long-term plans for periodic and systematic data acquisitions have guaranteed large amounts of Landsat data available, the availability of Landsat data is not ideal for persistently cloud-contaminated areas, i.e., the mountain area, is less than ideal [45,46]. Most of the Landsat 8 images for Yunnan are not free of clouds, and images from June to September are practically affected by clouds and cloud shadows (Figure 2). In this research, pixels affected by cloud and cloud shadows in each Landsat imagery (530 images in total) were identified and eliminated by Pixel Quality Assessment (QA) band from Landsat 8 SR collection [47–49]. The QA band generated from the CFMASK algorithm [48], and each pixel in the QA band contains a value that represents bit-packed combinations of conditions, such as cloud, cloud shadow, snow, and water, that can affect the given pixel [50]. Therefore, cloud and cloud shadow-affected pixels were masked.

After the elimination of cloud-affected pixels, seven spectral bands were derived for each image from the nine bands of original OLI image. Seven spectral bands are blue, green, red, NIR, SWIR1, SWIR2, and NIR_SWIR1. Because the spectral reflectance difference of different forest types in the NIR and SWIR1 is obvious, a compound band, NIR_SWIR1, is calculated by adding NIR and SWRI1. Phenology is a useful trait for forest type mapping by using the multi-temporal vegetation index approach [51]. The link between vegetation indexes (VIs) and vegetation phenological variability is more robust than that of original single spectral bands and vegetation phenological variability [52]. Therefore, five vegetation indexes, NDVI [53], EVI [54], SAVI [55], LSWI [56,57], and NBR2 [21] were calculated for each image using formulas shown in Table 4.

Table 4. Vegetation indexes used in the study.

Index	Abbreviation	Formula
Normalized Difference Vegetation Index	NDVI	$(ho_{NIR} - ho_{red})/(ho_{NIR} + ho_{red})$
Enhanced Vegetation Index	EVI	$2.5 imes rac{(ho_{NIR} - ho_{Red})}{ ho_{NIR} - 6.0 imes ho_{Red} - 7.5 imes ho_{Red} + 1}$
Soil-adjusted Vegetation Index	SAVI ¹	$(\rho_{NIR} - \rho_{Red})(1+C)/(\rho_{NIR} + \rho_{Red} + C)$
Land Surface Water Index	LSWI	$(\rho_{NIR} - \rho_{SWIR1})/(\rho_{NIR} + \rho_{SWIR1})$
Normalized Burn Ratio 2	NBR2	$(\rho_{SWIR1} - \rho_{SWIR2})/(\rho_{SWIR1} + \rho_{SWIR2})$

 1 C= 0.8.

Then, a pixel-based multi-seasonal composite method was used to produce two types of seasonal composite images, i.e., four 7-spectral-band composite images and four 5-VI-bands composite images corresponding to spring, summer, autumn and winter, respectively. Firstly, the time series were divided into four seasonal periods according to the seasonal characteristics of forest phenology and seasonal climate of Yunnan. Next, an annual composite was generated using all images in a year (from 1 March 2017 to 1 March 2018), and four seasonal composites for spring, summer, autumn, and winter were also respectively produced using temporal statistic operators. As for the production of four seasonal composites, if there was no observation due to the elimination of cloud-affected pixels, pixel values from the annual composite would be used instead. The temporal statistic operator,

the mean, and median values of time series pixels were explored to perform the seasonal image compositing. Here, the seasonal pixel-based composite method was applied to the time series of 7-spectral-band images and the time series of 5-VI-band images, respectively. As a result, four seasonal 7-spectral-band composites (Spr_Ref, Sum_Ref, Aut_Ref, and Win_Ref) and four seasonal 5-VI-band composites (Spr_VI, Sum_VI, Aut_VI, and Win_VI) were generated in the GEE cloud platform, respectively (shown in Table 5).

Table 5. Features, including four seasonal of spectral reflectance images and four seasonal vegetation index images, three terrain factors, temperature, and precipitation in 12 months, used in the research.

Data	Number of Feature Bands	Description
Ref_spr	7	A 7 reflectance-band seasonal image composited from the images of spring.
Ref_sum	7	A 7 reflectance-band seasonal image composited from the images of summer.
Ref_aut	7	A 7 reflectance-band seasonal image composited from the images of autumn.
Ref_win	7	A 7 reflectance-band seasonal image composited from the images of winter.
VI_spr	5	A 5-VI-band seasonal composited from the images of spring.
VI_sum	5	A 5-VI-band seasonal composited from the images of summer.
VI_aut	5	A 5-VI-band seasonal composited from the images of autumn.
VI_win	5	A 5-VI-band seasonal composited from the images of winter.
DEM, Slope, Aspect	3	Elevation, slope, aspect
T_12s	12	The monthly mean temperature for 12 months.
P_12s	12	The monthly mean precipitation for 12 months.

In addition, three commonly used terrain factors, elevation, slope, and aspect, were derived from the SRTM image. The average monthly temperature and average monthly precipitation data for 12 months were also produced from MOD11A2 Terra Land Surface Temperature and Emissivity and OpenLandMap Precipitation monthly data, respectively. The temperature and precipitation data were resampled from 1 km to 30 m resolution using nearest neighbor resampling to match the spatial resolution of other data. In summary, three types of feature sets, a total 75 bands, were derived, including four seasonal spectral bands sets, four seasonal vegetation index sets, and environmental factors (Table 5). The code for the implementing multi-seasonal image composite is provided in Supplementary Materials.

2.3.4. Random Forest Classification

Random Forest (RF) is an integrated learning method based on a decision tree, which is combined with many ensembles regression or classification trees [58], and has become increasingly common in remote sensing applications due to its flexible, nonparametric nature and ability to limit overfitting [59,60]. Bremian found that this method performs better than many other classifiers, such as discriminant analysis, support vector machines, and neural networks [58]. Many studies have shown that the RF classifier can successfully process high-dimensional data with fast calculating speed and is insensitive to over fitting [61]. In the past 20 years, RF classifier has attracted much attention and has been widely used in the processing remote sensing images [21,61].

Different features make the final result of classification different in accuracy [62]. In this study, based on the RF classifier, 17 feature combinations were built to perform hierarchical classification. As shown in Table 6, 17 feature combinations can be included in several types, which consist of single season features (R_{SS} , VI_{SS} , and RVI_{SS}), multi-seasonal features (R_{4S} , VI_{4S} , and RVI_{4S}), and the combination of multi-seasonal features and environmental factors ($RVI_{4S}T$ and $RVI_{4S}TTP$).

No.	Set Name	Feature Description	Dimension of Features	Set Type (Abbreviation)
1	Ref_spr	7 spectral reflectance ¹ in spring.		
2	Ref_sum	7 spectral reflectance ¹ in summer.	_	Reflectance in Single Season
3	Ref_aut	7 spectral reflectance ¹ in autumn.	7	(R _{SS})
4	Ref_win	7 spectral reflectance ¹ in winter.		
5	VI_spr	5 vegetation Indexes ² in spring.		
6	VI_sum	5 vegetation Indexes ² in summer.	_	Vegetation Index in Single
7	VI_aut	5 vegetation Indexes ² in autumn.	5	Season (VISS)
8	VI_win	5 vegetation Indexes ² in winter.		× ,
9	Ref_VI_spr	7 spectral reflectance and 5 vegetation Indexes in spring.		
10	Ref_VI_sum	7 spectral reflectance and 5 vegetation Indexes in summer.	12	Reflectance and Vegetation
11	Ref_VI_aut	7 spectral reflectance and 5 vegetation Indexes in autumn.		Index in Single season (RVISS)
12	Ref_VI_win	7 spectral reflectance and 5 vegetation Indexes in winter.		
13	Ref _{4S}	7 spectral reflectance in spring, summer, autumn and winter, respectively.	28	Reflectance in 4 seasons (R_{4S})
14	VI _{4S}	5 vegetation Indexes in spring, summer, autumn and winter, respectively.	20	Vegetation Index in 4 seasons (VI _{4S})
15	Ref_VI _{4S}	7 spectral reflectance and 5 vegetation Indexes in spring, summer, autumn and winter, respectively.	48	Reflectance_Vegetation Indexes in 4 seasons (RVI _{4S})
16	Ref_VI _{4S} _T	7 spectral reflectance and 5 vegetation Indexes in spring, summer, autumn and winter respectively & 3 terrain factors ³ .	51	Reflectance_Vegetation Indexes in 4 seasons & Terrain (RVI _{4S} T)
17	Ref_VI _{4S} _TTP	7 spectral reflectance and 5 vegetation Indexes in spring, summer, autumn and winter respectively & 3 terrain factors & 12 months Temperature & Precipitation ⁴ .	75	Reflectance_Vegetation Indexes in 4 seasons & Terrain & Temperature & Precipitation (RVI _{4S} TTP)

Table 6. The 17 feature combinations for RF classification experiments.

¹ Blue, Green, Red, NIR, SWIR1, SWIR2, NIR_SWIR1. ² NDVI, EVI, SAVI, LSWI, NBR2. ³ Elevation, Aspect and Slope. ⁴ The monthly mean temperature and the monthly mean precipitation for 12 months.

The hierarchical classification process was performed on GEE. In the first level, 17 feature combinations were separately carried out the forest/non-forest classification, and the accuracy of 17 forest/non-forest classification results were evaluated by the overall accuracy (OA) and Kappa. Then, the forest area of the result with the highest classification accuracy was taken as the final forest distribution area. In the second level, based on the forest cover produced from the first classification level, the forest was subdivided into five forest types (Table 2). Forest type classification was also conducted on 17 feature combinations, respectively, and the classification accuracy of 17 results were also evaluated and compared by OA and Kappa.

The classification accuracy of RF classifier is insensitive to two adjustable parameters: the number of trees and the number of prediction variables [45,63]. Based on the balance between classification accuracy and computation time, a final RF model with 100 trees for forest/non-forest classification and 200 trees for forest type classification was set up in the study. The number of selected prediction variables were set to the square root of the number of variables for both forest/non-forest classification.

2.3.5. Feature Importance Assessment

Feature importance assessment is a vital step when high-dimensional datasets are used [64]. In the research, in addition to considering the impact of different feature combinations on forest classification, we also used "Gini importance" to find out which spectral, vegetation index, and environmental factor features were most important for classification of different forest types. "Gini importance" is a feature importance measure method embedded in Random Forest classifier [58]. The Gini importance score describes the relative importance of features after comparing the input features, and it corresponds to features that are consistently found more often and higher up in the split of individual decision trees [64]. Therefore, based on GEE, feature importance score was calculated by an additional RF classification with all the 75 features (The set Ref_VI_{4S}_TTP shown in Table 6) using "Gini importance".

Furthermore, although including more predictor variables potentially adds additional information to separate classes, this increased dimensionality and complexity may result in a decrease in classification accuracy [42]. Using only the most important features may result in higher classification accuracies than using all available features [8]. Even if the accuracy is not improved by only the most important features, some studies have shown that the loss of the overall classification accuracy is relatively small when the number of features was reduced, during mapping land cover with RF [42,65]. In addition, the reduction of predictor variables may help to simplify the model, shorten data processing times, and make the experiment reproducible. In this study, the Random Forest recursive feature elimination (RFRFE) algorithm was used to perform feature selection (the algorithm principle is demonstrated in [18,66]). We performed RFRFE on GEE.

2.3.6. Accuracy Evaluation and Comparison

Accuracy assessment of our resultant forest maps include two aspects: (1) evaluation by calculating confusion matrix with the test samples collected from FMI, and (2) visual and quantitate comparison with four public forest maps.

The confusion matrix is often used to evaluate the classification accuracy of remote sensing image, and it is able to identify the confusion among categories and the possible sources of errors. In this study, the confusion matrix was calculated by using the test samples that came from FMI, which is built through a field survey. The overall accuracy (OA), user accuracy (UA), producer accuracy (PA), commission error (CE), omission error (OE), and Kappa coefficient were calculated for the classification results of both levels.

Furthermore, the forest/non-forest classification result with the best OA and the forest type classification result with the best OA were selected to compare with four commonly used forest type products, including Forest2010 [9], FROM_GLC2015 [11], GLC_FCS2020 [67], and MCD12Q1 [68], to assess the quality of the forest type mapping product of this study. Among them, Forest2010 is a product focusing on mapping forest types over China, and the others are global land cover products with 30 m resolution and 500 m resolution. Forest2010 and FROM_GLC2015 also use IGBP classification schemes to define their six types of forest [9,11,68]. Compared to the five forest types in this study, Forest 2010 has one more type of bamboo forest, while FROM_GLC2015 divides mixed forest into evergreen mixed and deciduous mixed forests. GLC_FCS2020 has ten forest types according to forest canopy density and forest types (Table 7). In this study, the ten forest types of GLC_FCS2020 has been categorized into five forest types used in this study for comparison. MCD12Q1 includes a land cover classification scheme from IGBP classification, and this research used the MCD12Q1 products produced in 2017. The comparisons were conducted on the Yunnan Province as a whole and eight selected case regions (Figure 3a, red rectangle), respectively.

Table 7. Forest type comparison of different products.

Products	Forest Type
This study	evergreen needleleaf forest, deciduous needleleaf forest, evergreen broadleaf forest, deciduous broadleaf forest and mixed forest
Forest2010	evergreen needleleaf forest, deciduous needleleaf forest, evergreen broadleaf forest, deciduous broadleaf forest, mixed forest, and bamboo forest
FROM_GLC2015	evergreen needleleaf forest, deciduous needleleaf forest, evergreen broadleaf forest, deciduous broadleaf forest, evergreen mixed forest, deciduous mixed forest
GLC_FCS2020	open evergreen broadleaved forest $(0.15 < \text{fc}^{-1} < 0.4)$, closed evergreen broadleaved forest $(\text{fc}^{-1} > 0.4)$, open deciduous broadleaved forest $(0.15 < \text{fc}^{-1} < 0.4)$, closed deciduous broadleaved forest $(\text{fc}^{-1} > 0.4)$, open evergreen needle-leaved forest $(0.15 < \text{fc}^{-1} < 0.4)$, closed evergreen needle-leaved forest $(\text{fc}^{-1} > 0.4)$, open deciduous needle-leaved forest $(0.15 < \text{fc}^{-1} < 0.4)$, closed deciduous needle-leaved forest $(\text{fc}^{-1} > 0.4)$, open mixed leaf forest $(0.15 < \text{fc}^{-1} < 0.4)$, closed mixed leaf forest $(\text{fc}^{-1} > 0.4)$
MCD12Q1	evergreen needleleaf forest, deciduous needleleaf forest, evergreen broadleaf forest, deciduous broadleaf forest, and mixed forest

 1 fc = forest canopy.

3. Results

In order to evaluate the effectiveness of the proposed method, many experiments were conducted. Firstly, the mean and median statistical operators were tested on four seasonal composites and a multi-seasonal composite image by comparing forest/non-forest classification accuracy and forest classification accuracy, to determine the optimal statistics operator (Section 3.1). Secondly, based on the optimal multi-seasonal image composite, by comparing the accuracies of hierarchical classification of 17 feature combinations, the feature combination obtained the best accuracy was fixed (Section 3.2). Thirdly, by performing feature importance analysis, an optimal feature subset for forest type classification were selected from the feature combination (Section 3.3). In addition, finally, the forest and the non-forest classification and forest type classification results were produced by using the optimal feature combination and subset, respectively (Section 3.4).

3.1. Input Seasonal Image Composite

As highlighted by Phan et al. [69], the choice of the image composite method can influence the accuracy of the classification process. Different statistical operation leads to different accuracy values at the end of the classification process. As mentioned in Materials and Methods, the mean and median statistical operators were tested on four seasonal composites with 7-spectral-band composites (Spr_Ref, Sum_Ref, Aut_Ref, and Win_Ref) and a multi-seasonal 28-spectral-band composite (R_{4S}), and the two operators were compared by assessing the forest type classification accuracy of their corresponding composites.

Table 8 reports the main results of two statistical operations on four seasonal composites. In the first level where forest and non-forest classification was performed, the median-based composites achieved higher accuracy than mean-based composites, and the best result of forest/non-forest classification was obtained using the median-based multi-seasonal composite image, with a Kappa of 0.945, and OA of 97.33%. As far as the forest type classification was concerned, the results of mean-based composites in spring, summer, and autumn obtained better accuracy than that of median-based composites; however, the median-based multi-seasonal composite performed better than mean-based multi-seasonal composite, with a Kappa of 0.442, and OA of 58.81%. Therefore, a median statistic operator was finally chosen to produce four seasonal 7-spectral-band composites and a multi-seasonal 28-spectral-band composite image.

Input Image		FNF				Forest Type			
	Mean		Median		Mean		Median		
	Kappa	OA (%)	Kappa	OA (%)	Kappa	OA (%)	Kappa	OA (%)	
Spr_Ref	0.917	95.94	0.922	96.22	0.325	50.26	0.323	50.25	
Sum_Ref	0.836	92.01	0.839	92.12	0.286	47.37	0.279	46.90	
Aut_Ref	0.886	94.44	0.890	94.63	0.320	49.84	0.314	49.44	
Win_Ref	0.920	96.10	0.924	96.29	0.388	54.94	0.389	55.02	
R _{4S}	0.944	97.28	0.945	97.33	0.438	58.48	0.442	58.81	

Table 8. The classification accuracy (OA and Kappa) comparison of composite images produced by two statistical operators (mean and median) of image composites tested in the proposed method.

3.2. Random Forest Classification Results

Different features lead to different accuracy values at the end of the classification process. To assess the best achievable accuracy, based on the RF classifier, 17 feature combinations were separately carried out the forest classification on GEE (Table 6). Table 9 compares the classification accuracy of all the experimental results with OA and Kappa coefficient. For the forest/non-forest classification, $RVI_{4S}TTP$ obtained the best accuracy with OA and Kappa values of 97.64% and 0.952. Therefore, the forest/non-forest map from set RVI_{4STTP} was used as the basis of forest type classification.

Cat Trues		G	Forest/N	on-Forest	Fores	Forest Type	
Set Type	Set Name	Season —	Kappa	OA (%)	Kappa	OA (%)	
	Ref_spr	Spring	0.922	96.22	0.387	51.12	
р	Ref_sum	Summer	0.839	92.12	0.335	46.78	
K _{SS}	Ref_aut	Autumn	0.890	94.63	0.410	53.03	
	Ref_win	Winter	Forest/Non-Forest Kappa OA (%) Spring 0.922 96.22 ummer 0.839 92.12 uutumn 0.890 94.63 Winter 0.924 96.29 Spring 0.893 94.76 ummer 0.791 89.81 utumn 0.855 92.70 Winter 0.885 94.39 Spring 0.937 96.91 ummer 0.866 93.47 wutumn 0.913 95.77 Winter 0.926 96.64 ti-Seasonal 0.929 96.52 ti-Seasonal 0.929 96.52 ti-Seasonal 0.946 97.10 ti-Seasonal 0.949 97.51 ti-Seasonal 0.949 97.51	0.466	57.46		
	VI_spr	Spring	0.893	94.76	0.365	49.45	
VI	VI_sum	Summer	0.791	89.81	0.303	44.59	
VISS	VI_aut	Autumn	0.85	92.70	0.365	49.50	
	VI_win	Winter	0.885	94.39	0.415	53.41	
	RVI_spr	Spring	0.937	96.91	0.402	52.46	
DVI	RVI_sum	Summer	0.866	93.47	0.362	49.26	
KV ISS	RVI_aut	Autumn	0.913	95.77	0.416	53.55	
RVI _{SS}	RVI_win	Winter	0.926	96.64	0.470	57.80	
R _{4S}	R _{4S}	Multi-Seasonal	0.945	97.33	0.513	61.23	
VI _{4S}	VI _{4S}	Multi-Seasonal	0.929	96.52	0.524	62.09	
RVI _{4S}	RVI _{4S}	Multi-Seasonal	0.946	97.10	0.538	63.28	
RVI _{4S} T	RVI _{4S} T	Multi-Seasonal	0.949	97.51	0.564	65.28	
RVI _{4S} TTP	RVI _{4S} TTP	Multi-Seasonal	0.950	97.57	0.638	71.14	

Table 9. The OA and Kappa obtained with different feature combinations.

For the forest type classification, the best accuracy value was reached when using 75-feature set $RVI_{4S}TTP$ again (71.14% and 0.638 of OA and Kappa, respectively). Using only single season features (R_{SS} , VI_{SS} , and RVI_{SS}) resulted in their OA and Kappa lower than 60% and 0.500. The usage of multi-seasonal spectral reflectance features (R_{4S}) achieved an OA of 61.23%, Kappa of 0.513, and RVI_{4S} , which was formed by adding vegetation indexes to R_{4S} , displayed an OA to 63. 82%. Moreover, $RVI_{4S}T$, which had three additional terrain features (elevation, slope, and aspect) than RVI_{4S} , resulted in an obvious improvement in accuracy (65.28% and 0.564 of OA and Kappa, respectively), so it could be concluded that the additional environmental factors also improved the OA and Kappa of RVI_{4S}TTP.

The lowest accuracy found in the result from VI_sum which has five VI bands in summer, showing 44.54% and 0.304 of OA and Kappa, respectively.

3.3. Feature Importance

Feature importance score was used to find out which features contributed more in the classification of forest types. In addition, it was performed using the Gini criterion in the RF algorithm of GEE platform. Figure 5 shows the feature importance of 75 features in the classification of five forest types. The variables are ranked by the score of importance. According to feature importance score, vegetation index features and environmental features show significant importance. The top 10 features were six vegetation index features and four environmental features, and the elevation contributed the most to the classification, which is consistent with [6]. Among all the vegetation index features, the VI features in winter composite image and autumn composite image contributed the most to the classification of five forest types. Furthermore, the top 36 features were 14 vegetation index variables, 12 temperature features, six precipitation features, three terrain features, and only one spectral reflectance feature. Therefore, besides vegetation index features and terrain features, the monthly mean temperature and monthly mean precipitation over 12 months also show their importance. It could be found that spectral reflectance features (blue in Figure 5) show less importance since they mainly ranked in the last 30 variables.



Figure 5. The feature importance score for five forest types classified using the combination of multiseasonal spectral reflectance and vegetation index, terrain (elevation, slope, and aspect), temperature, and precipitation features. The features are colored according to the feature type: orange indicates environmental factors (27 variables), green indicates vegetation indexes (20 variables), and blue indicates spectral reflectance features (28 variables).

From the feature with the highest score to the one with the lowest score, RFRFE starts with all the features to execute forest type classification and recursively removes one insignificant feature each time. Figure 6 shows how the OA and Kappa in the forest type classification increase with the increase in the number of features. According to Figure 6, the highest accuracy was obtained when using 44 variables. The accuracy obviously increases from using four features (60.63% of OA) to using 16 features used (70.03% of OA). After reaching the peak when 44 features are used (72.17% and 0.657 of OA and Kappa, respectively), the resulting differences among the following results are minor, with OA values oscillating between 69.33 and 72.17%. Considering achieving a better OA with a faster process speed on GEE, the forest type mapping result from 16 features were used. The 16 features are elevation, Winter_LSWI, Winter_NBR2, Winter_NDVI, Autumn_NDVI, Autumn_LSWI, Autumn_NBR2, LST_12, Slope, Precip_7, Precip_8, LST_7, Precip_4, Spring_NBR2, Summer_LSWI, Precip_6.



Figure 6. Results of Random Forest recursive feature elimination: number of variables and OA and Kappa of the mode. Three vertical red lines represented 16 variables (the minimum number of features to achieve an accuracy higher than 70%), 44 variables (the number of features to obtain the highest OA and Kappa), and 75 variables (all the features) were considered, respectively.

3.4. Final Maps of Forest/Non-Forest and Forest Type

Figure 7 illustrates the final classification outputs of this study. The forest/non-forest classification results were finally produced by using 75 features since the highest accuracy was achieved when using 75 variables according to Table 9. Specifically, 28 features are multispectral bands from seasonal composite images in four seasons, 28 features are VIs from seasonal composite images in four seasons, 3 features are from terrain features, and 24 features are from monthly mean precipitation and monthly mean temperature over 12 months. The forest type classification results were finally produced by using 16 features, considering achieving a better OA and a faster process speed on GEE (depicted in Section 3.3).



Figure 7. The classification results of this study. (**a**) the forest/non-forest map of Yunnan; (**b**) the forest type map of Yunnan.

Through the evaluation of sufficient forest and non-forest sample plots that covers more than 150,000 pixels, the results show that these features can get 97.64% OA for distinguishing the forest and non-forest area (Table 10), and the class of forest also obtained high UA (97.72%) and PA (98.22%) for the class of forest. Therefore, the identified forest cover could be the basis of forest type classification. Figure 7a shows the final map result of forest/non-forest mapping in the southwest of Yunnan Province. We can see that the nonforest land covers, including urban, cropland, water, and grass types, are mainly located in the flat areas and valleys, while the forests are mainly located in the mountainous areas with an altitude of less than 4000 m. The area of forest cover in the southwest half of Yunnan is more than that of the northeast half, and this is consistent with the precipitation trend. Most of the forests are distributed in the mountainous areas where the precipitation is more than 400 mm. Table 11 reports the confusion matrix for the second stage classification. Forest type classification achieved an overall accuracy of 70.30% and a Kappa coefficient of 0.628. The PA of the five forest types ranged from 56.99% (mixed forest) to 95.44% (deciduous needleleaf forest), while UA ranged from 62.21% (mixed forest) to 94.49% (deciduous needleleaf forest). The identification of deciduous needleleaf forest is better than other forest types. Mixed forests had the highest CE (37.79%) and OE (43.01%), and it had been found to be mostly misclassified as evergreen needleleaf forest and evergreen broadleaf forest. In addition, there was confusion among evergreen coniferous forests and evergreen broad-leaved forests, probably due to these two forest types belong to evergreen forests, and the seasonal variation of phenology between the two forest types is relatively small.

 Table 10. Confusion matrix for the forest/non-forest classification map in Yunnan.

	True Class										
		Forest	Non-Forest	Total	UA (%)	CE (%)					
-	Forest	86,330	1909	88,239	97.84	2.16					
ioi s	Non-forest	1568	61,779	63,347	97.52	2.48					
lict	Total	87,898	63,793	151,691							
lred	PA (%)	98.22	96.68								
с С	OE (%)	1.78	3.32								
OA = 97.57% Kappa = 0.950											

Table 11. Confusion matrix for the forest type classification map in Yunnan.

	True Class										
		ENF	DNF	EBF	DBF	MF	Total	UA (%)	CE (%)		
	ENF	281	7	58	29	67	442	63.57	36.43		
	DNF	5	377	7	2	8	399	94.49	5.51		
u	EBF	45	10	311	30	75	471	66.03	33.97		
ctic ss	DBF	27	0	37	229	50	343	66.76	33.24		
edi	MF	53	1	77	30	265	426	62.21	37.79		
Pre	Total	411	395	490	320	465	2081				
	PA (%)	68.37	95.44	63.47	71.56	56.99					
	OE (%)	31.63	4.56	36.53	28.44	43.01					
	OA = 70.30 Kappa = 0.628										

Figure 7b shows the result of forest type classification. The forest type distribution map shows that needleleaf forests are mainly distributed in northwestern and northeast of Yunnan, and mixed forests are mainly distributed in southern Yunnan and the western boundary of Yunnan. It also displays that the forest types in Yunnan Province gradually change from coniferous forests in the northwest to broadleaf forests in the middle, and then to mixed forests in the south. Deciduous needleleaf forest is mainly distributed in the northwestern part of Yunnan Province and less in other regions. Comparing the forest type map (Figure 7b) to the elevation map of Yunnan (Figure 1b), it could be found that the transition of forest types is consistent with the transition trend of altitude.

4. Discussion

4.1. Seasonal Composite Image

A key challenge in mapping forest type over a large region with remote sensed data is that the coverage of valid data is limited both time and in area due to the presence of clouds. Due to the complex topography and climate of Yunnan Province, there are more clouds and rain from June to October, which leads to fewer high-quality images (Figure 2). Except for the overlapping area of the adjacent scenes, the original observation frequency of 35 scenes of images covering Yunnan is in the range of 10 to 20 times from 1 March 2017 to 1 March 2018 (Figure 8a). However, after cloud elimination, the number of observations per pixel in a 12-month period is mostly in the range of 0 to 10 times (Figure 8b). Therefore, it is difficult to obtain a high-quality single-date image at every location of 35 scenes within a season, especially in summer and autumn (Figure 9a–d). In addition, for a large research area, it takes more time to select the images with less amounts of clouds.

Different from the traditional way, there is no valid image selection in the production of the median-based seasonal/annual composite images on GEE. All images in the 12-month period were used to produce the composite images, making sure all cloud-free pixels can be used. As shown from Figure 9e–h, four seasonal cloud-free images covering the whole area were generated using the pixel-based median composite method. The colors of spring, autumn, and winter composite images look much smoother across the whole area, while the summer composite image has obvious color boundaries. This is mainly because all the images for summer composite output values in these areas are derived from the annual composite image rather than the median values from the images for summer. As a result, the large spectral differences of some land covers between summer and other seasons make the color boundaries in the summer composite image. In summary, the seasonal image composite method using 12-month time series images offers an opportunity to overcome the limited availability of valid data.



Figure 8. The observation frequency of 35 scenes of Landsat 8 images over the whole region. (**a**) initial observation frequency; (**b**) observation frequency after the elimination of cloud-affected pixels.

(a)





Single-date image of spring



Median composite of spring









Median composite of summer Median composite of autumn

Median composite of winter

Figure 9. Single-date images and median-based seasonal composite images of four seasons. Four seasonal single-date images were obtained by selecting and mosaicking 35 images with the least amount of clouds in each season and the pixels with cloud were eliminated during mosaicking. False color composite of SWIR1/NIR/Red as RGB. (a) Single-date image of spring, (b) Single-date image of summer, (c) Single-date image of autumn, (d) Single-date image of winter, (e) Median composite of spring, (f) Median composite of summer, (g) Median composite of autumn, (h) Median composite of winter.

Figure 10 shows the spectral reflectance curves of five forest types in a summer singledate image and summer composite image. As shown in Figure 10, due to the median statistic operation for the composite image, the spectral reflectance values of five forest types from composite image are lightly lower than that of from the single-date image. In general, the spectral reflectance values of five forest types in the summer composite image are close to the corresponding values in the cloudless single-date image for five forest types. Analysis of these spectral curve suggests that the median-based seasonal composite values are credible.



Figure 10. The spectral reflectance curves of five forest type from the summer composite image and the single-date image for summer. The squares represent the values from the single-date image of summer, while the asterisks are the summer reflectance from the composite image.

4.2. The Analysis of Different Feature Combinations

(1) The influence of different single-season composite images

In the experiments of R_{SS}, VI_{SS} and RVI_{SS}, four single-season composite images were separately used to classify forest types. Through the analysis of the OA and Kappa of the three sets of experimental results, it is found that the season of input image has a significant impact on the forest classification accuracy. All the experimental results from R_{SS}, VI_{SS} and RVI_{SS} showed a similar pattern that the composite images from winter had the highest OA, followed by the autumn composite images, and the classification accuracy of the summer composite images were the lowest (Table 9). The OAs of forest type classification using winter composite in R_{SS}, VI_{SS} and RVI_{SS} are 8.82%, 9.54%, and 10.21%, higher than that of the summer composites, respectively (Table 9). Given the best season of composite image to extract each forest type may vary with forest type, the PA and UA of five forest types were compared in Figure 11. For deciduous needleleaf forests, composite images from autumn produced the highest PA and UA in all the single-season feature combinations. In addition, when classifying evergreen needleleaf forests, deciduous broadleaf forests, and mixed forests, composite images from winter produced the highest PA and UA in all single-season feature combinations. The results of this study are different from that of [18] who found forest species can be best separated using image in the growing season. One of the reasons is that the levels of forest type classification focused on the two studies are different. For example, we classified the forest area into five broad forest categories and [18] classified forest types in more detail and determined the forest types at the individual tree level. Another possible explanation for the inconsistent result is that the study areas adopted by the two studies have different climates characteries. Yunnan Province located in the tropical and subtropical climate zones and the phenological difference of different forest types are more obvious in the leaf-off season. For instance, the length of season life of broad-leaved forest is longer than that of needleleaf forest; therefore, the vegetation index value of broad-leaved forest decreases later than that of needleleaf forest. Furthermore, the vegetation index values of deciduous forests are significantly reduced in winter, while that of evergreen forests decreased only slightly in winter.



Figure 11. The accuracy comparison of four seasonal single-season composite images on 7 spectralreflectance bands, 5 VI bands, and the combination of bands of spectral reflectance and VI, respectively.

(2) The influence of multi-seasonal

Previous research has clearly demonstrated the effectiveness of using multi-seasonal images to distinguish among forest types [1,15,64,70]. It was found that that multi-seasonal feature sets consistently and significantly outperformed all single season feature sets across the result comparison between R_{4S}, VI_{4S}, RVI_{4S} and Ref_{SS}, VI_{SS}, RVI_{SS}. For all the classifications using spectral reflectance features alone, vegetation index features alone, or a combination of spectral reflectance and VI features, their OAs showed that the combined use of multi-seasonal images features (R_{4S}, VI_{4S}, RVI_{4S}) could achieve higher accuracy than using single-season image features alone (Ref_{SS}, VI_{SS}, RVI_{SS}). From R_{4S} to Ref_{SS}, VI_{4S} to VI_{SS}, RVI_{4S} to RVI_{5S}, their accuracy improvements of three feature sets are 4.86–15.07%, 8.68–17.50%, and 4.96–14.50%, respectively, reaffirming the importance of phenological information contained in multi-seasonal images for forest type mapping.

Furthermore, the strategy of using multi-seasonal images varies in different research. Ref. [1] selected images from the leaf-off, growing, and leaf-on season, and then extracted spectral reflectance and vegetation index features from these images. Ref. [15] screened the image from the period which is the best phenological state for the classification of different tree species; therefore, five images in different seasons were used. Ref. [64] collected all available Landsat observations from 1985 to 2015 and then extracted harmonic and phenology features from time series of all cloud-clear observations. For the forest type mapping over a large region, as suggested in [64], multi-seasonal feature composites tend to show more spatial variability compared to individual date images.

(3) The influence of spectral reflectance and vegetation index

The influence of spectral band and vegetation index on forest type classification accuracy is related to the season of the image. For example, the specific single-season spectral reflectance features in R_{SS} performs better than their corresponding single-season vegetation index features in VI_{SS}, and the OAs of four single-season experiments from set R_{SS} are 1.81–3.58% higher than that of from VI_{SS}. In contrast, the OA of multi-seasonal spectral reflectance composite (R_{4S}) is 2.04% lower than that of multi-seasonal VI composite (VI_{4S}). This indicates that multi-seasonal VI features are crucial for the forest type classification because multi-seasonal VI can reflect the seasonal phenology differences of different forests. In general, the combination of the spectral reflectance and vegetation index features has a better classification result than either of them alone, regardless of whether the combination is used in a single season or multiple seasons, i.e., RVI_{SS} is better than R_{SS} or VI_{SS} , and RVI_{4S} is better than R_{4S} or VI_{4S} .

(4) The influence of environmental factors

In this research, three topographic features, elevation, slope, and aspect were chosen as features to classify forest. According to the feature importance score, elevation is the most important features among all features involved in forest type classification (Figure 5). In the classification of five forest types, the addition of three terrain features into the feature combination for forest type classification resulted in an improvement of PA of 2.54–3.65%, and UA of 0.44–3.71% (Figure 12). Elevation and slope are also reported as important features in mapping tree species distribution in [1,8] as they can provide geographical characteristics. In addition, topographic data, especially elevation and slope, are frequently used in mapping land use and land cover over a regional or global scale [10,45,71]. Therefore, topographical features should also be considered when mapping global forest types.

Temperature and precipitation are crucial factors that cause the change of phenology with time [72]. Our results show that the features of precipitation and temperature characteristics also have obvious effects on improving forest type classification accuracy. Figure 12 compared RVI_{4S} and RVI_{4S}TTP, and it could be found that the addition of the features of precipitation and temperature factors brought a significant improvement of PA of 3.37-12.28%, and UA of 4.85-10.01% on evergreen broadleaf forest, deciduous broadleaf forest, and mixed forest. The accuracy increases of the classification results on evergreen needleleaf and deciduous needleleaf forest are not obvious. This is because the distribution

of forest types is related to the distribution of the trend of temperature and precipitation and broad-leaved forests require more water and are more sensitive to the variation of precipitation. In the research, due to the lack of temperature and precipitation data with 30 m resolution over a large region, the 1-km temperature and precipitation data were to resample by the reproject function embedded in GEE. This processing might influence the classification accuracy. In this study, the impact of resampling method on forest type classification was not discussed.



Figure 12. The PA and UA comparison of multi-seasonal composite images with spectral reflectance, VI, and the combination of spectral reflectance and VI, respectively.

4.3. Comparisons to Other Products

As mentioned in the Introduction, there are some 30 m resolution products with fined forest types, including Forest2010 [9], FROM_GLC2015 [11], GLC_FCS2020 [67], and MCD12Q1 [68]. The methods of forest type mapping or land use and cover mapping were compared, and there are four products adopting different strategy to select remote sensing data and features (Table 12). The early product, Forest2010, was derived from single-date Landsat images circa 2010 [9]. FROM_GLC2015 was produced after the launch and free access of Landsat 8, and therefore it could use multi-seasonal Landsat images [11]. FROM_GLC2015 chose one image with the least cloud for each season (spring, summer, autumn, and winter) and the vegetation growing season, so there are five multi-seasonal images for each path/row position, which can better distinguish vegetation types with similar spectral characteristics in a particular season and improve mapping accuracy. GLC_FCS2020 adopted a new data selection method, time series Landsat data, which uses all Landsat imagery from specific period [45,67]. With the advancement of computer computing ability, the production of forest type mapping over large areas has also shifted from local computers to cloud computing platforms. Taking FMI as "true" data, the forest/non-forest classification and the forest type classification results of this study were further compared with four reference products, respectively.

4.3.1. Comparisons on Forest/Non-Forest

The forest/forests classification result of this study were compared with the reference products in two aspects. On one hand, the comparison was made at the administrative division level of Yunnan Province; and on the other hand, it was carried out on the eight selected case regions selected from Forest Management Inventory. Because the four reference products have many forest types and non-forest types, all the non-forest types of each of the four products were combined into one type named as non-forest, and different forest types were combined as forest.

Yunnan Province has a total of 16 administrative divisions. This study calculated the forest area of each division in the products of this study and the four references. Taking

FMI data as a standard, the forest area of 16 divisions in each product were respectively compared to FMI. As shown in Figure 13, the estimated forest area of this study had the best correlation with FMI with $R^2 = 0.9816$, followed by Forest2010 with $R^2 = 0.977$ and GLC_FCS2020 with $R^2 = 0.9742$. FROM_GLC2015 has the lowest correlation with FMI with an $R^2 = 0.8341$.

Table 12. Comparisons of four public products.

Product Name	Classifier	Parameters	Landsat/MODIS Images	Environmental Factors	Platform
Forest 2010	Random Forest	Ntree = 200	520 Landsat TM with fewer clouds obtained in the growing season	DEM (90 m)	-
FROM_GLC 2015	Random Forest	Ntree = 200	60,998 multi-seasonal Landsat images primarily acquired with Landsat 8 between 2013–2015;	Geographical coordinate, elevation (30 m)	Supercom-puter
GLC_FCS30 2020	-	-	Landsat surface reflectance imagery from 2019–2020	DEM (30 m);	GEE
MCD12Q1	Random Forest	-	C6 MODIS Normalized BRDF-Adjusted Reflectance (NBAR) product (MCD43A2, MCD43A4)	Worldclim climatology	-



Figure 13. Comparison of forest coverage area between FMI and the products of this study and four references. (**a**) This study; (**b**) Forest 2010; (**c**) FROM_GLC2015; (**d**) GLC_FCS2020; and (**e**) MCD12Q1 2017. FMI: Forest Management Inventory statistics.

Besides the above quantitative assessment, Figure 14 showed the comparisons of the classification performance of each product with FMI over eight case regions (Figure 3a), which covers various climate and landscape environment. The forest/non-forest map of this study has great performance in Regions 3, 4, 5, 6, and 8. In Regions 1, 2, and 7, there are some obvious non-forest cover lands that were classified as forest in this study. As for Forest 2010, it performs well in Regions 4, 6, 7, and 8, but its forest distribution is significantly sparser than FMI in Regions 1, 2, and 3, and some of the forest on the 8 side of the its Region 5 has not been identified. The distribution of forests in FROM_GLC2015 is significantly more than FMI, which is obvious in Regions 1, 2, and 7. Compared with FMI, FLC_FCS2020 also has great performance in Regions 4, 5, 6, and 8, and forest distribution is significantly more than FMI in Regions 1, 2, and 7, and the details of Region 3 are not as good as our products. Because of the coarse resolution, MCD12Q1 shows less detail in all eight case regions. Based on the above comparisons, it could be found that the forest/non-forest



mapping result of this study is superior to other existing products, and therefore it can be used as the basis for forest type classification.

Figure 14. Regional comparison of forest coverage among the product of FMI, this study, Forest 2010, FROM_GLC2015, GLC_FCS2020, and MCD12Q1 2017 over eight case regions. The first column are maps from FMI as a standard.

4.3.2. Comparisons on Forest Types

By comparing the distribution of forest types of different products over the whole Yunnan Province, it can be found that the differences between different products are very large [73]. Due to the lack of FMI statistical data on the distribution area of different forest types in the Yunnan Province, we only compared the classification details of this study and the reference products in eight case regions. Figure 15 shows that: (1) In the results of this study, Regions 3, 4, and 8 have achieved satisfactory classification results, while Region 2 and Region 7 have serious confusion between DBF and MF, the classification in Region 5 is not as good as the result in Forest2010, and, in Region 6, some ENF are divided into EBF. (2) As for Forest2010, it works well in Region 5 and Region 8, but the difference between Forest2010 and FMI is obvious in the other six regions. In addition, there is no deciduous needleleaf forest in Yunnan Province in Forest2010. (3) The pattern of FROM_GLC2015 is relatively fragmented, which is quite different from FMI. The most obvious differences between FROM_GLC2015 and FMI occur in Regions 2, 3, 4, and 8. In Region 2, many pixels with deciduous broadleaf forest are classified as evergreen broadleaf forest or evergreen deciduous forest. In Region 3, the mixed forest is not distinguished. In Region 4, the mixed forest is classified as evergreen broadleaf forest. In Region 8, evergreen coniferous forest is classified as mixed forests. (4) GLC_FCS2020 works well in Region 6 and Region 8, and obvious difference exists in other six regions by comparing with FMI. (5) MCD12Q1 product has poor visual effects due to its low resolution. It could be found that the forest product of this study is better than that of others by comparing with FMI. This study and Forest 2010 are forest products that focus on forest type mapping. FROMGLC 2015 and MCD12Q1 are land cover products that focus on the classification of all land cover types, and forest cover or forest types is one of all types. Therefore, the results of this study and Forest2010 are more accurate that FROM_GLC2015 and MCD12Q1. Since FMI is a ground survey result, it is taken as a correct data for comparison. Overall, the result of this study has a better performance.

In addition, the samples in this study are distributed in eight sites instead of the entire Yunnan Province because it is difficult to obtain samples from the entire study area. As a result, the classification is inevitably affected by the local distribution of samples. Therefore, we will use samples evenly distributed in the entire study area for further research. GEE provides great convenience for massive data managing, processing, and analysis. The large amount of calculation that originally took several days can be completed in a few hours, which greatly saves work time. The using of Landsat time-series imagery and GEE provides a new direction for the improvement of large-scale forest type mapping.



Figure 15. Local comparison for forest coverage developed by FMI, this study, Forest 2010, FROM_GLC2015, GLC_FCS 2020 and MCD12Q1 2017 in eight case regions. FMI: Forest management inventory.

5. Conclusions

In the research, a method of forest types mapping using Landsat-8 OLI imagery and multiple environment factors has been developed and tested within Yunnan Province of China, a large mountainous region with landscape of high geographic and climate heterogeneity. The approach employs a seasonal image compositing methodology adapted for incomplete annual data availability and frequent cloud cover. By comparing with a single-date image, the multi-seasonal image composite method could generate a cloud-free image over the whole mountainous region, Yunnan Province. The results of comparative experiments, based on a two-step classification strategy combining a feature combination procedure, confirmed that the multi-seasonal features can distinguish forest types more accurately than only single-seasonal feature. Seasonal characteristics quantify meaningful spectral and temporal variability in the complex forested landscape of Yunnan Province. Our results also demonstrated that adding environmental factors can significantly improve the classification accuracy of forest types. The feature combination of multi-seasonal spectral reflectance features, and vegetable index features, and three environmental factors (topography, temperature data, and precipitation data) obtained the highest accuracy on both forest/non-forest and forest type classification. Through the method of this study, an updated annual forest type distribution product with a resolution of 30 m can be obtained. In addition, benefitting from the powerful computing abilities and archive of Google Earth Engine cloud platform, the production of the products in this study is convenient and conducive to the real-time update of future data. Based on Google Earth Engine, the approach has the potential to be used in other areas and applied to remotely sensed images from different sources to map forest types at regional or global scales.

Supplementary Materials: The code for implementing median-based multi-seasonal image composite for forest type mapping using Landsat 8 images is available in GEE: https://code.earthengine. google.com/d47a5f8e8489f866c916ec211119ac01 (accessed on 25 November 2021).

Author Contributions: Conceptualization, L.W. and W.X.; methodology, R.L. and L.W.; software, R.L. and P.F.; validation, L.W. and G.O.; formal analysis, R.L., W.Z., and L.W.; investigation, G.O.; resources, L.W. and G.O.; data curation, R.L., P.F., and L.W.; writing—original draft preparation, R.L. and L.W.; writing—review and editing, L.W., W.Z., and X.H.; visualization, R.L. and P.F.; supervision, L.W. and W.Z.; project administration, L.W.; funding acquisition, L.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partly funded by the National Natural Science Foundation of China, Grant Nos. 31860182, 32160369, 31860181 and 41961053, partly funded by the Fund of Reserve Talents for Young and Middle-Aged Academic and Technological Leaders of Yunnan Province, Grant No. 2018HB026.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy restrictions.

Acknowledgments: The authors gratefully acknowledge the valuable assistance provided by Baokun Feng of Yunnan Normal University for his generous offer to review this manuscript.

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

References

- Liu, Y.; Gong, W.; Hu, X.; Gong, J. Forest Type Identification with Random Forest Using Sentinel-1A, Sentinel-2A, Multi-Temporal Landsat-8 and DEM Data. *Remote Sens.* 2018, 10, 946. [CrossRef]
- Liu, S.; Dong, Y.; Sun, Y.; Li, J.; An, Y.; Shi, F. Modelling the spatial pattern of biodiversity utilizing the high-resolution tree cover data at large scale: Case study in Yunnan province, Southwest China. *Ecol. Eng.* 2019, 134, 1–8. [CrossRef]
- Tøttrup, C. Improving tropical forest mapping using multi-date Landsat TM data and pre-classification image smoothing. *Int. J. Remote Sens.* 2004, 25, 717–730. [CrossRef]
- 4. Peiran, Z. Yunnan's Forest Coverage Rate Reaches 65.04%; Xinhuanet: Beijing, China, 2021.
- 5. He, Y.; Cao, M.; Tang, Y.; Yang, G. A comparative study on the tree species diversity of evergreen broad-leaved forests in central Yunnan. *J. Mt. Sci.* 2000, *18*, 322. [CrossRef]

- 6. Hościło, A.; Lewandowska, A. Mapping Forest Type and Tree Species on a Regional Scale Using Multi-Temporal Sentinel-2 Data. *Remote Sens.* **2019**, *11*, 929. [CrossRef]
- Young, B.; Yarie, J.; Verbyla, D.; Huettmann, F.; Herrick, K.; Chapin, F.S. Modeling and mapping forest diversity in the boreal forest of interior Alaska. *Landsc. Ecol.* 2017, 32, 397–413. [CrossRef]
- 8. Grabska, E.; Frantz, D.; Ostapowicz, K. Evaluation of machine learning algorithms for forest stand species mapping using Sentinel-2 imagery and environmental data in the Polish Carpathians. *Remote Sens. Environ.* **2020**, *251*, 112103. [CrossRef]
- 9. Li, C.; Wang, J.; Hu, L.; Yu, L.; Clinton, N.; Huang, H.; Yang, J.; Gong, P. A Circa 2010 Thirty Meter Resolution Forest Map for China. *Remote Sens.* 2014, *6*, 5325–5343. [CrossRef]
- 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]
- Gong, P.; Wang, J.; Li, C.; Ji, L.; Huang, H.; Clinton, N.; Cheng, Y.; Li, W.; Zhang, M.; Bai, Y.; et al. Automated Global Land Cover Mapping–FROM-GLC Version 2: The Production of the 30 M Circa 2015 Global Land Cover Map; Tsinghua University: Beijing, China, 2017. [CrossRef]
- Chen, J.; Chen, J.; Liao, A.; Cao, X.; Chen, L.; Chen, X.; He, C.; Han, G.; Peng, S.; Lu, M.; et al. Global land cover mapping at 30 m resolution: A POK-based operational approach. *ISPRS J. Photogramm. Remote Sens.* 2015, 103, 7–27. [CrossRef]
- Li, C.; Gong, P.; Wang, J.; Zhu, Z.; Biging, G.S.; Yuan, C.; Hu, T.; Zhang, H.; Wang, Q.; Li, X.; et al. The first all-season sample set for mapping global land cover with Landsat-8 data. *Sci. Bull.* 2017, *62*, 508–515. [CrossRef]
- 14. Fassnacht, F.E.; Latifi, H.; Stereńczak, K.; Modzelewska, A.; Lefsky, M.; Waser, L.T.; Straub, C.; Ghosh, A. Review of studies on tree species classification from remotely sensed data. *Remote Sens. Environ.* **2016**, *186*, 64–87. [CrossRef]
- 15. Lanari, R. A new method for the compensation of the SAR range cell migration based on the chirp z-transform. *IEEE Trans. Geosci. Remote Sens.* **1995**, *33*, 1296–1299. [CrossRef]
- 16. Kou, W.; Liang, C.; Wei, L.; Hernandez, A.J.; Yang, X. Phenology-Based Method for Mapping Tropical Evergreen Forests by Integrating of MODIS and Landsat Imagery. *Forests* **2017**, *8*, 34. [CrossRef]
- 17. Dorren, L.K.; Maier, B.; Seijmonsbergen, A.C. Improved Landsat-based forest mapping in steep mountainous terrain using object-based classification. *For. Ecol. Manag.* 2003, 183, 31–46. [CrossRef]
- Zhu, X.; Liu, D. Accurate mapping of forest types using dense seasonal Landsat time-series. *ISPRS J. Photogramm. Remote Sens.* 2014, 96, 1–11. [CrossRef]
- 19. Xu, K.; Zeng, H.; Chang, C.-T.; Xie, J.; Yang, Y. Relating forest phenology to temperature and precipitation in subtropical Fujian based on time-series MODIS-NDVI. *J. Geo-Inf. Sci.* 2015, *17*, 1249–1259. [CrossRef]
- Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* 2017, 202, 18–27. [CrossRef]
- Oliphant, A.J.; Thenkabail, P.S.; Teluguntla, P.; Xiong, J.; Gumma, M.K.; Congalton, R.G.; Yadav, K. Mapping cropland extent of Southeast and Northeast Asia using multi-year time-series Landsat 30-m data using a random forest classifier on the Google Earth Engine cloud. *Int. J. App. Earth Observ. Geoinf.* 2019, 81, 110–124. [CrossRef]
- 22. Cheng, K.; Wang, J.; Yan, X. Mapping Forest Types in China with 10 m Resolution Based on Spectral–Spatial–Temporal Features. *Remote Sens.* **2021**, *13*, 973. [CrossRef]
- Coltin, B.; McMichael, S.; Smith, T.; Fong, T. Automatic boosted flood mapping from satellite data. *Int. J. Remote Sens.* 2016, 37, 993–1015. [CrossRef] [PubMed]
- Dong, J.; Xiao, X.; Menarguez, M.A.; Zhang, G.; Qin, Y.; Thau, D.; Biradar, C.; Moore, B., III. Mapping paddy rice planting area in northeastern Asia with Landsat 8 images, phenology-based algorithm and Google Earth Engine. *Remote Sens. Environ.* 2016, 185, 142–154. [CrossRef]
- Hansen, M.C.; Potapov, P.V.; Moore, R.; Hancher, M.; Turubanova, S.A.; Tyukavina, A.; Thau, D.; Stehman, S.V.; Goetz, S.J.; Loveland, T.R.; et al. High-resolution global maps of 21st-century forest cover change. *Science* 2013, 342, 850–853. [CrossRef] [PubMed]
- 26. Lobell, D.B.; Thau, D.; Seifert, C.; Engle, E.; Little, B. A scalable satellite-based crop yield mapper. *Remote Sens. Environ.* **2015**, *164*, 324–333. [CrossRef]
- 27. Patel, N.N.; Angiuli, E.; Gamba, P.; Gaughan, A.; Lisini, G.; Stevens, F.; Tatem, A.J.; Trianni, G. Multitemporal settlement and population mapping from Landsat using Google Earth Engine. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, 35, 199–208. [CrossRef]
- Pekel, J.-F.; Cottam, A.; Gorelick, N.; Belward, A.S. High-resolution mapping of global surface water and its long-term changes. *Nature* 2016, 540, 418–422. [CrossRef] [PubMed]
- Xie, Z.; Phinn, S.R.; Game, E.T.; Pannell, D.J.; Hobbs, R.J.; Briggs, P.R.; McDonald-Madden, E. Using Landsat observations (1988–2017) and Google Earth Engine to detect vegetation cover changes in rangelands—A first step towards identifying degraded lands for conservation. *Remote Sens. Environ.* 2019, 232, 111317. [CrossRef]
- 30. Zhang, M.; Gong, P.; Qi, S.; Liu, C.; Xiong, T. Mapping bamboo with regional phenological characteristics derived from dense Landsat time series using Google Earth Engine. *Int. J. Remote Sens.* **2019**, *40*, 9541–9555. [CrossRef]
- Kollert, A.; Bremer, M.; Löw, M.; Rutzinger, M. Exploring the potential of land surface phenology and seasonal cloud free composites of one year of Sentinel-2 imagery for tree species mapping in a mountainous region. *Int. J. Appl. Earth Obs. Geoinf.* 2021, 94, 102208. [CrossRef]

- 32. Xiwen, L.; Walker, D. The Plant Geography of Yunnan Province, Southwest China. J. Biogeogr. 1986, 13, 367. [CrossRef]
- Wu, P.; Shen, H.; Cai, N.; Zeng, C.; Wu, Y.; Wang, B.; Wang, Y. Spatiotemporal analysis of water area annual variations using 33. a Landsat time series: A case study of nine plateau lakes in Yunnan province, China. Int. J. Remote.Sens. 2016, 37, 5826–5842. [CrossRef]
- Wan, Z.; Hook, S.; Hulley, G. MOD11A2 MODIS/Terra Land Surface Temperature/Emissivity 8-Day L3 Global 1 km SIN Grid V006; 34. NASA LP DAAC: Greenbelt, MD, USA, 2015. [CrossRef]
- Brocca, L.; Filippucci, P.; Hahn, S.; Ciabatta, L.; Massari, C.; Camici, S.; Schüller, L.; Bojkov, B.; Wagner, W. SM2RAIN–ASCAT 35. (2007–2018): Global daily satellite rainfall data from ASCAT soil moisture observations. Earth Syst. Sci. Data 2019, 11, 1583–1601. [CrossRef]
- 36. Fu, Y.; Li, X.; Zhou, X.; Geng, X.; Guo, Y.; Zhang, Y. Progress in plant phenology modeling under global climate change. Sci. China Earth Sci. 2020, 63, 1237–1247. [CrossRef]
- Farr, T.G.; Rosen, P.A.; Caro, E.; Crippen, R.; Duren, R.; Hensley, S.; Kobrick, M.; Paller, M.; Rodriguez, E.; Roth, L.; et al. The 37. Shuttle Radar Topography Mission. Rev. Geophys. 2007, 45, 1–33. [CrossRef]
- 38. Wang, X.-H.; Kent, M.; Fang, X.-F. Evergreen broad-leaved forest in Eastern China: Its ecology and conservation and the importance of resprouting in forest restoration. For. Ecol. Manag. 2007, 245, 76–87. [CrossRef]
- 39. Wu, Z.; Zhu, Y.; Jiang, H. Flora of Yunnan; Science Press: Beijing, China, 1987.
- 40. Zeng, J. The Classification System of Natural Forests and Its Geographic Distribution in Yunnan. J. Southwest For. Univ. 2018, 38, 1-18. [CrossRef]
- 41. Belward, A.S. The IGBP-DIS Global 1 Km Land Cover Data Set (DISCover): Proposal and Implementation Plans, IGBP-DIS Working Paper 13; International Geosphere–Biosphere Programme Data and Information System Office: Toulouse, France, 1996.
- Maxwell, A.E.; Warner, T.A.; Fang, F. Implementation of machine-learning classification in remote sensing: An applied review. 42. Int. J. Remote Sens. 2018, 39, 2784–2817. [CrossRef]
- 43. Huang, C.; Davis, L.S.; Townshend, J.R.G. An assessment of support vector machines for land cover classification. Int. J. Remote Sens. 2002, 23, 725–749. [CrossRef]
- 44. Tan, B.; Du, J. Research on Forest Resource Information Update and Mapping Method Combining Remote Sensing and GIS. For. Res. 2001, 14, 4. [CrossRef]
- Zhang, X.; Liu, L.; Chen, X.; Gao, Y.; Xie, S.; Mi, J. GLC_FCS30: Global land-cover product with fine classification system at 30 m 45. using time-series Landsat imagery. Earth Syst. Sci. Data 2021, 13, 2753–2776. [CrossRef]
- Griffiths, P.; van der Linden, S.; Kuemmerle, T.; Hostert, P. A Pixel-Based Landsat Compositing Algorithm for Large Area Land 46. Cover Mapping. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2013, 6, 2088–2101. [CrossRef]
- USGS. CFMask Algorithm. Available online: https://www.usgs.gov/core-science-systems/nli/landsat/cfmask-algorithm 47. (accessed on 13 August 2021).
- Foga, S.; Scaramuzza, P.L.; Guo, S.; Zhu, Z.; Dilley, R.D.; Beckmann, T.; Schmidt, G.L.; Dwyer, J.L.; Hughes, M.J.; Laue, B. Cloud 48. detection algorithm comparison and validation for operational Landsat data products. Remote Sens. Environ. 2017, 194, 379–390. [CrossRef]
- 49. Zhu, Z.; Wang, S.; Woodcock, C.E. Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. Remote Sens. Environ. 2015, 159, 269–277. [CrossRef]
- 50. USGS. USGS Landsat 8 Level 2, Collection 2, Tier 1. Available online: https://developers.google.com/earth-engine/datasets/ catalog/LANDSAT_LC08_C02_T1_L2 (accessed on 25 December 2021).
- Wang, X.; Zhang, S.; Feng, L.; Zhang, J.; Deng, F. Mapping Maize Cultivated Area Combining MODIS EVI Time Series and the 51. Spatial Variations of Phenology over Huanghuaihai Plain. Appl. Sci. 2020, 10, 2667. [CrossRef]
- Coppin, P.; Jonckheere, I.; Nackaerts, K.; Muys, B.; Lambin, E. Digital change detection methods in ecosystem monitoring: A 52. review. Int. J. Remote Sens. 2004, 25, 1565–1596. [CrossRef]
- 53. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. J. Remote Sens. Environ. 1979, 8, 127–150. [CrossRef]
- Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the radiometric and biophysical performance 54. of the MODIS vegetation indices. Remote Sens. Environ. 2002, 83, 195–213. [CrossRef]
- Huete, A.R. A soil-adjusted vegetation index (SAVI). Remote Sens. Environ. 1988, 25, 295–309. [CrossRef] 55.
- Gao, B.-C. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sens. 56. Environ. 1996, 58, 257–266. [CrossRef]
- 57. Xiao, X.; Zhang, Q.; Braswell, B.; Urbanski, S.; Boles, S.; Wofsy, S.; Moore, B.; Ojima, D. Modeling gross primary production of temperate deciduous broadleaf forest using satellite images and climate data. Remote Sens. Environ. 2004, 91, 256–270. [CrossRef] 58.
- Breiman, L. Random Forests. Mach. Learn. 2001, 45, 5–32. [CrossRef]
- 59. Belgiu, M.; Drăguţ, L. Random forest in remote sensing: A review of applications and future directions. ISPRS J. Photogramm. Remote Sens. 2016, 114, 24–31. [CrossRef]
- Gómez, C.; White, J.C.; Wulder, M.A. Optical remotely sensed time series data for land cover classification: A review. ISPRS J. 60. Photogramm. Remote Sens. 2016, 116, 55–72. [CrossRef]
- 61. Mellor, A.; Haywood, A.; Stone, C.; Jones, S. The Performance of Random Forests in an Operational Setting for Large Area Sclerophyll Forest Classification. Remote Sens. 2013, 5, 2838–2856. [CrossRef]

- Praticò, S.; Solano, F.; Di Fazio, S.; Modica, G. Machine Learning Classification of Mediterranean Forest Habitats in Google Earth Engine Based on Seasonal Sentinel-2 Time-Series and Input Image Composition Optimisation. *Remote Sens.* 2021, 13, 586. [CrossRef]
- 63. Gislason, P.O.; Benediktsson, J.A.; Sveinsson, J.R. Random Forests for land cover classification. *Pattern Recognit. Lett.* **2006**, 27, 294–300. [CrossRef]
- 64. Pasquarella, V.J.; Holden, C.E.; Woodcock, C.E. Improved mapping of forest type using spectral-temporal Landsat features. *Remote Sens. Environ.* **2018**, *210*, 193–207. [CrossRef]
- 65. Duro, D.C.; Franklin, S.; Dubé, M.G. Multi-scale object-based image analysis and feature selection of multi-sensor earth observation imagery using random forests. *Int. J. Remote Sens.* **2011**, *33*, 4502–4526. [CrossRef]
- 66. Genuer, R.; Poggi, J.-M.; Tuleau-Malot, C. VSURF: An R Package for Variable Selection Using Random Forests. *R J.* **2015**, *7*, 19–33. [CrossRef]
- Liu, L.; Zhang, X. Global Land Cover with Fine Classification System at 30m in 2020 (GLC_FCS30-2020) User Guides; Aerospace Information Research Institute, Chinese Academy of Sciences: Beijing, China, 2020.
- 68. Friedl, M.A.; Sulla-Menashe, D. User Guide to Collection 6 MODIS Land Cover (MCD12Q1 and MCD12C1) Product; NASA LP DAAC: Greenbelt, MD, USA, 2017.
- 69. Phan, T.N.; Kuch, V.; Lehnert, L.W. Land Cover Classification using Google Earth Engine and Random Forest Classifier—The Role of Image Composition. *Remote Sens.* **2020**, *12*, 2411. [CrossRef]
- Clark, M.L. Comparison of multi-seasonal Landsat 8, Sentinel-2 and hyperspectral images for mapping forest alliances in Northern California. *ISPRS J. Photogramm. Remote Sens.* 2020, 159, 26–40. [CrossRef]
- Liu, H.; Gong, P.; Wang, J.; Clinton, N.; Bai, Y.; Liang, S. Annual dynamics of global land cover and its long-term changes from 1982 to 2015. *Earth Syst. Sci. Data* 2020, 12, 1217–1243. [CrossRef]
- 72. Nietupski, T.C.; Kennedy, R.E.; Temesgen, H.; Kerns, B.K. Spatiotemporal image fusion in Google Earth Engine for annual estimates of land surface phenology in a heterogenous landscape. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *99*, 102323. [CrossRef]
- Li, R.; Wang, L.; Ou, G.; Xu, W.; Dai, Q. Mapping Forest Type with Multi-Seasonal Landsat Data and Multiple Environmental Factors in Yunnan Province Based on Google Earth Engine. In Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Brussels, Belgium, 11–16 July 2021; pp. 6468–6471. [CrossRef]