remote sensing

MDPI

Article

Compositing Strategies for the Dynamic World Land-Cover Product

Yan Xu [1,2], Yongming Xu [1,*], Yutong Sun [1], Yiqi Wang [3] and Shanyou Zhu [1]

[1] School of Remote Sensing and Geomatics Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China; 261302173@njnu.edu.cn (Y.X.); sunyutong@niglas.edu.cn (Y.S.); 001646@nuist.edu.cn (S.Z.)
[2] School of Geography, Nanjing Normal University, Nanjing 210023, China
[3] College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China; 12531016@zju.edu.cn
* Correspondence: xym30@nuist.edu.cn

Highlights

What are the main findings?

- The mean spatial coverage achieved by the monthly Dynamic World composites was 68.78%; coverage increased to 98.52% and 98.57% with the growing-season and annual windows, respectively.
- Growing-season cumulative-probability and mode compositing achieved the highest OA and macro-F1, respectively. Common-sample analysis favored cumulative-probability compositing.

What are the implications of the main findings?

- Extending the compositing window from the growing season to the full year provided only a marginal gain in spatial coverage but reduced classification accuracy because of seasonal variations in land cover.
- Compositing strategies should be tailored to the temporal window and target land-cover class; growing-season cumulative-probability compositing is recommended when both spatial completeness and classification accuracy are needed.
- Annual compositing is most useful when residual data gaps must be filled. In such cases, mode compositing is preferable, and winter-related Snow & Ice misclassification should be carefully screened.

Abstract

Dynamic World provides near-real-time 10 m land-cover data, but scene-level classifications are affected by observation conditions and seasonal changes, causing missing valid data and unstable class labels. Using 2020 Dynamic World data, this study evaluated different temporal windows and compositing methods to identify suitable strategies for mapping stable annual land-cover types in China. We compared mode, cumulative-probability, and maximum-confidence composites at monthly, growing-season (April–October), and annual scales. Spatial coverage was assessed, and classification accuracy was evaluated using 14,042 reference samples. Effective observation rates showed clear spatial and temporal differences. Rates were lower in southwestern China, higher in northwestern and northern China, and generally lower in summer. Longer compositing windows significantly reduced data gaps. The monthly scale spatial coverage was 68.78%, while the growing-season and full-year windows increased spatial coverage to 98.52% and 98.57%, respectively. However, full-year coverage gains over the growing-season window were limited. Growing-season mode and cumulative-probability composites outperformed the other strategies, with the

Check for updates

Academic Editors: Rong Shang, Wang Li, Xiaobin Guan, Su Ye and Feng Zhao

Received: 30 September 2026
Revised: 3 November 2026
Accepted: 5 November 2026
Published: 7 November 2026

Copyright: © 2026 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.

highest macro-F1 (64.96%) and OA (76.19%), respectively. However, performance varied by time window and land-cover type, and no method consistently performed best. These results can guide the selection of time windows and compositing methods for regional Dynamic World mapping.

Keywords: dynamic world; land cover; multi-temporal compositing; accuracy assessment

1. Introduction

Land cover describes the basic natural attributes of the Earth's surface and human activities. Changes in land cover affect the surface energy balance and the biogeochemical water and carbon cycles, which are key variables in global change research [1–4]. Land-cover products are an important source of spatial data for climate and land-surface process simulation, ecosystem service assessment, biodiversity conservation, and hydrological research [5–8]. Land-cover products with high spatial resolution can more accurately characterize complex and broken land-cover patterns and their dynamic changes, which is of great significance for monitoring environmental changes, supporting land spatial planning and government decision-making, and promoting the realization of sustainable development goals [9–12].

Remote sensing technology allows for large-scale real-time observation of the surface environment and is an important means to obtain land-cover and environmental information [13,14]. With the development of high-resolution remote sensing sensors, cloud computing, and artificial intelligence technology, the spatial resolution of global land-cover products has been continuously improved, gradually developing from the kilometer level to the 10 m level [15–20]. Gong et al. constructed the 2017 FROM-GLC10 product based on Sentinel-2 images and migration samples, and achieved global 10 m resolution land-cover mapping [9,21]. Karra et al. trained deep learning models using Sentinel-2 time series images and large-scale manually labeled samples to generate Esri Land Cover products containing 10 land types [10]. The European Space Agency produced WorldCover2020 products by combining Sentinel-1 radar images and Sentinel-2 optical images, providing global mapping results for 11 land-cover types at a 10 m resolution [18].

Unlike the above land-cover products, which are published in the form of annual maps for a specific year, Google has partnered with the World Resources Institute (WRI) to develop a 10 m resolution near-real-time global land-cover product, Dynamic World. The product generates land-cover classification results based on the qualified Sentinel-2 L1C image and realizes near real-time update. It shows important application value in the fields of ecological environment change monitoring, disaster assessment, and land-use dynamic analysis [8,22–24]. However, due to the influence of clouds and other factors, there is a lack of spatial coverage in the scene-by-scene products, and the observation conditions and seasonal changes may cause instability in ground object recognition [25,26]. Therefore, it is necessary to make full use of effective observations of different phases through multi-temporal compositing, fill in missing areas, and weaken category fluctuations caused by factors such as clouds and seasonal changes to improve the spatial integrity and category stability of land-cover results.

Previous studies have generally conducted multi-temporal compositing of Dynamic World imagery over specific temporal windows tailored to particular application needs. Venter et al. processed Dynamic World data using mode compositing, mean class-probability compositing, and median compositing and compared the results with WorldCover and Esri Land Cover [18]. Myers et al. applied mode compositing separately to Dynamic World classes for the growing and non-growing seasons and examined how seasonal land-cover

estimates affected hydrological and water-quality models [27]. Rangel et al. used mode compositing on Dynamic World imagery acquired from July to September to construct a U.S. land-cover change dataset for 2016–2024 [28]. Li et al. used median compositing of November–December Dynamic World data to identify forest-disturbance types from 2017 to 2023 [29]. Rashid et al. applied mode compositing to Dynamic World to generate a 2021 land-cover map for flood-susceptibility assessment in the Mohmand Dam watershed [30]. These studies demonstrate the value of multi-temporal Dynamic World compositing for product assessment and environmental monitoring. Nevertheless, systematic comparisons of temporal windows and compositing methods within a unified evaluation framework remain limited, particularly with respect to spatial completeness, classification accuracy, and class-specific differences.

Therefore, this study utilized the 2020 Dynamic World data for China to conduct Dynamic World product compositing using different time scales and methods, in order to systematically evaluate the spatial integrity and classification accuracy of the results using different compositing strategies, clarify the performance and applicable scenarios of these different compositing strategies, and provide a scientific reference for time-window and method selection for the application of regional-scale Dynamic World land-cover products.

2. Study Area and Datasets

2.1. Study Area

China is located between 3°51′ and 53°33′ north latitude and 73°33′ and 135°05′ east longitude. It has a land area of about 9.6 million square kilometers, and its topography includes mountains, plateaus, hills, plains, and basins. Affected by factors such as latitude, land and sea distribution, and monsoon circulation, China spans tropical, subtropical, temperate, and plateau climate regions. Figure 1 presents satellite imagery of China and its climatic regionalization. Complex natural geographical conditions and high-intensity human activities have jointly shaped a variety of land-cover patterns such as forests, grasslands, shrubs, wetlands, cultivated land, bare land, and construction land. Cultivated land phenology, vegetation seasonal differences, and water fluctuations will also cause obvious time series changes. China's complex physiography and diverse land-cover types provide an ideal region for comparing composite results under different regional and seasonal observation conditions and for assessing class-specific responses to different temporal windows and compositing methods.

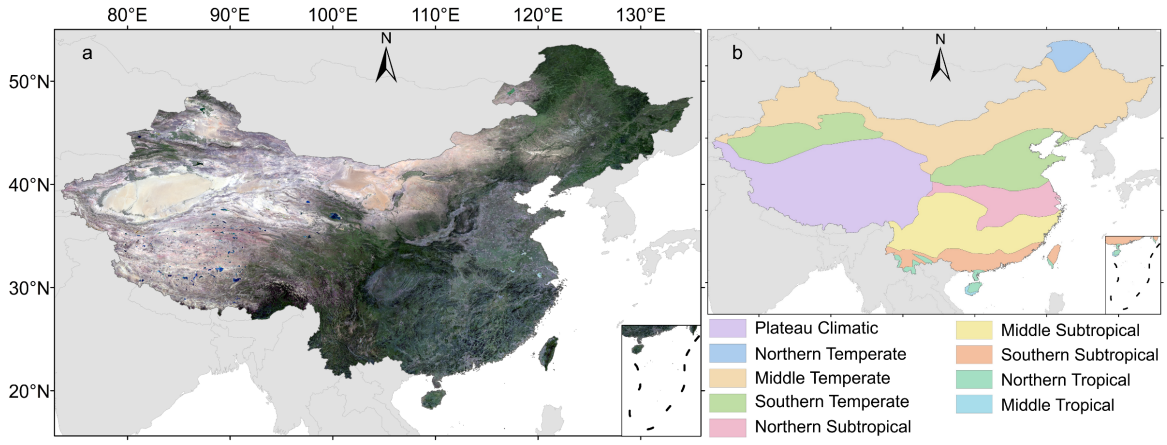


Figure 1. Overview of the study area: (**a**) true-color satellite imagery; (**b**) climate zones (source: GISRS).

2.2. Study Datasets

2.2.1. Dynamic World Land-Cover Product

This study is based on the 2020 Dynamic World product. Dynamic World takes Sentinel-2 L1C top-of-atmosphere reflectance imagery as input, uses a Fully Convolutional Neural Network (FCNN) to composite the spectral characteristics and neighborhood spatial information of pixels, and generates land-cover classification results scene by scene [8]. Only Sentinel-2 images with a scene-level cloud cover of 35% or less are classified. Clouds and cloud shadows are masked using Sentinel-2 cloud probability, the Cloud Displacement Index, and a directional distance transform, leaving some spatial gaps in the product.

Dynamic World uses a classification system consisting of nine land-cover types: Water, Trees, Grass, Flooded Vegetation, Crops, Shrub & Scrub, Built-Up Area, Bare Ground, and Snow & Ice. In addition to the category label band, it also provides nine categories of probability bands and gives the estimated probability that the pixel belongs to each category. The sum of the probabilities of each category in the same pixel is 1.

2.2.2. Reference Samples

In order to take into account the large area differences and sample representativeness of different land-cover types, the candidate samples were obtained by stratified random sampling according to the category division of multi-source land-cover products, and the reference categories were determined via manual visual interpretation. Specifically, the total sample size was determined according to the comprehensive area ratio of each category, which was then assigned to different products to randomly sample the corresponding category in each product.

In this study, four kinds of global 10 m resolution land-cover data (WorldCover, Esri Land Cover, Dynamic World, and FROM-GLC10) were integrated into the Dynamic World classification system, and the area of each land-cover category in the four products was determined. The initial sample size for land-cover class i was calculated from the ratio of the summed area of that class across the four products to the summed area of all classes:

$$n_i = N \frac{\sum_{j=1}^{4} A_{ij}}{\sum_{i=1}^{9} \sum_{j=1}^{4} A_{ij}} \quad (1)$$

where n_i is the initial total number of samples for class i; A_{ij} is the area of class i in land-cover product j; and N is the planned total sample size.

The sample size was then allocated among the products according to the area proportion of that class in each product:

$$n_{ij} = n_i \frac{A_{ij}}{\sum_{j=1}^{4} A_{ij}} \quad (2)$$

where n_{ij} is the number of samples to be drawn for class i from product j.

A total of 15,000 candidate samples were extracted, and the area ratio of each category and the initial sample allocation are shown in Table 1. After determining the number of samples for each category in each product, all valid pixels of the corresponding category in each product were taken as the sampling range of the category, a specified number of pixels were randomly selected, and their central positions were used as candidate sample points. Then, the candidate samples generated by the four products were summarized, and artificial visual interpretation and quality inspection were carried out. No duplicate geographic locations were found among the randomly generated candidate sample points.

Table 1. Initial number of sample points allocated to each class.

Land-Cover Class	Proportion (%)	Initially Allocated Samples
Water	2.09	314
Trees	26.04	3906
Grass	12.36	1854
Flooded Vegetation	0.18	27
Crops	16.16	2424
Shrub & Scrub	10.40	1560
Built-Up Area	3.99	599
Bare Ground	28.12	4218
Snow & Ice	0.66	99
Total	100	15,001

Due to the small proportion of wetland and ice/snow areas, strict proportional sampling may lead to an insufficient sample size, so supplementary sampling of these two types was carried out based on the initial allocation. The reference class of each candidate sample was independently determined through visual interpretation of Google Earth high-resolution imagery, supported by auxiliary data. The class labels of the four land-cover products, including Dynamic World, were used for sampling only and did not determine the final reference labels. The imagery used for interpretation was primarily acquired in 2020. If no imagery from 2020 was available for a sample location, imagery from the years before and after 2020 was used instead. These images were compared to verify temporal stability, and only samples showing no evidence of land-cover change over the period examined were retained. The reference labels represent stable annual land-cover classes and were kept the same for all temporal windows. Differences between the composites and reference labels may reflect seasonal changes, such as snow cover over cropland, rather than classification errors.

We used the following criteria to guide sample screening and interpretation: (1) Each sample point had to lie within a homogeneous 10 × 10 m area and have clear high-resolution imagery. Points that did not meet these conditions were moved or excluded. (2) For classes with seasonal changes, such as Trees, Grass, and Bare Ground, Google Earth images from different dates were checked. Summer images helped distinguish permanent snow and glaciers from seasonal snow. Unclear cases were checked against China's 1:1 million vegetation map or Google Earth street-level photographs. Only points with the same underlying land-cover class across dates were retained. (3) Multiple interpreters checked each point and discussed any disagreements. If they could not agree, the point was moved or excluded. A total of 14,042 reference samples were retained. The numbers of samples added or removed for each class are provided in Table A2. Figure 2 shows the spatial distribution and number of reference sample points.

2.3. Data Processing and Sample Extraction

Dynamic World images were retrieved from the GOOGLE/DYNAMICWORLD/V1 collection in Google Earth Engine. For each temporal window, we selected images covering the reference sample locations, with the start date included and the end date excluded. We retained the masks provided by Dynamic World without applying additional cloud masking or probability thresholds and used all unmasked observations for compositing. Land-cover classes were then extracted at the reference sample locations at a spatial scale of 10 m using the projection of each composite.

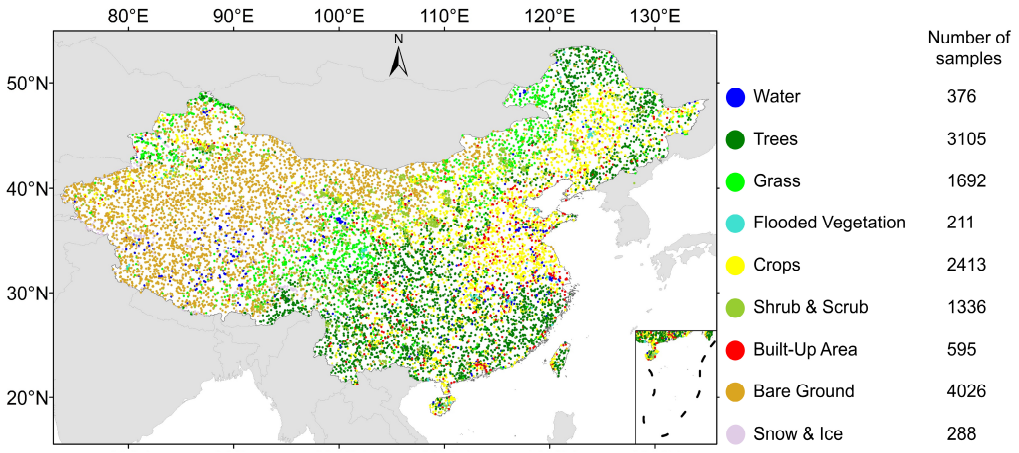


Figure 2. Spatial distribution and class counts of the reference sample points.

3. Methods

3.1. Analysis of Valid Dynamic World Observations

To quantify valid observations in Dynamic World, the pixel-level valid-observation rate was calculated for China in 2020. This metric was defined as the ratio of the number of valid Dynamic World classifications to the number of Sentinel-2 observations meeting the scene-level cloud-cover criterion during the same period (Equation (3)):

$$M_{i,t} = \frac{N_{valid,i,t}}{N_{total,i,t}} \quad (3)$$

where $M_{i,t}$ is the valid-observation rate of pixel i during period t; $N_{total,i,t}$ is the number of Sentinel-2 images containing pixel i during period t for which the scene-level cloud cover was below 35%; and $N_{valid,i,t}$ is the number of times that pixel i received a valid classification in Dynamic World during the same period. To avoid division by zero, the valid-observation rate was only calculated for pixels with at least one Sentinel-2 observation satisfying the Dynamic World classification criteria.

The spatial distribution of the Dynamic World valid-observation rate was then analyzed, and monthly rates were used to calculate the intra-annual variation (Equation (4)):

$$\overline{M}_m = \frac{1}{n_m} \sum_{i=1}^{n_m} M_{i,m} \quad (4)$$

where $\overline{M}_m$ is the mean valid-observation rate for month m in the study area; $M_{i,m}$ is the valid-observation rate of pixel i in month m; and n_m is the number of valid pixels included in the statistics for month m.

3.2. Multi-Temporal Compositing

3.2.1. Compositing Windows

To analyze how different time windows affect the compositing results of the Dynamic World land-cover product, this study focused on the year 2020 and used three time windows: monthly, growing-season, and annual. For each window, all valid observations within the time range were used for mode, cumulative-probability, and maximum-confidence compositing to compare compositing performance across different observation periods.

Each method generated 12 monthly land-cover maps using all valid observations within each month. These maps were used to compare spatial completeness, overall classification accuracy, and class-specific seasonal differences.

The growing-season window takes April–October as the compositing period, and all effective observations in this period were used for compositing to investigate the classifica-

tion performance using the longer seasonal window. It should be noted that China has a vast territory, and there are differences in climatic conditions, vegetation phenology, and agricultural cropping systems in different regions, and the starting and ending times of the actual growing season are not consistent. For convenience, this study uses April–October as the compositing window of the growing season, rather than delimitation for each region. To assess this choice, we also evaluated March–September, April–September, and May–October using the three compositing methods.

The annual window covers January to December, incorporating all valid observations from the year into the compositing. Compared with the growing-season window, this setup adds observation information from other months, offering more chances to fill in missing spatial data, but it might also introduce seasonal category differences. By comparing the results of annual and growing-season compositing, we can investigate if extending the time window further improves spatial coverage and how the classification performance for different land types changes.

3.2.2. Compositing Methods

Three methods were used for multi-temporal compositing: mode, cumulative-probability, and maximum-confidence compositing. Each method determines the final land-cover class of a pixel based on the frequency of class occurrence, the cumulative result of multi-temporal classification probabilities, and the classification confidence, respectively.

Mode compositing assigns each pixel the most frequent class across observation dates. When two or more classes have the same highest frequency, the class with the highest mean probability is selected.

Cumulative-probability compositing sums the probabilities of each class across observation dates and assigns each pixel the class with the highest total:

$$\hat{c}(i) = \text{argmax}_{c \in C} \sum_{t \in T} P_t(c|i) \quad (5)$$

where $\hat{c}(i)$ is the final composite class of pixel i, and $P\square(c|i)$ is the probability that pixel i belongs to class c at time t.

Maximum-confidence compositing treats the class probabilities output at each time point as the classification confidence, which represents the reliability of the classification at that time, and selects the class with the highest confidence across all times and classes as the final result:

$$\hat{c}(i) = \text{argmax}_{c \in C}[\max_{t \in T} P_t(c|i)] \quad (6)$$

The inner max operator searches across all temporal observations to obtain the maximum confidence for each class, whereas the outer argmax operator selects the class with the highest maximum confidence among all classes as the final composite class.

For cumulative-probability and maximum-confidence compositing, if probabilities were equal, the more frequent class within the temporal window was selected.

3.3. Evaluation

We evaluated the composites in terms of spatial completeness, classification accuracy, and spatial consistency. Spatial completeness was measured as the proportion of the study area with valid classifications, using the same total area for all temporal windows.

Of the 14,042 reference samples, only 3641 had valid classifications across all temporal windows and compositing methods because of extensive missing classifications in the monthly composites. Accuracy was evaluated using the available valid samples from the full reference dataset for each composite and, separately, using the common set of 3641 samples. The true classes of the reference samples were linked to the classified results of

their corresponding pixels to create a confusion matrix. Overall Accuracy (OA), Kappa coefficient, User's Accuracy (UA), Producer's Accuracy (PA), and F1 score were calculated. OA and Kappa give a sense of the overall classification performance. UA and PA indicate the misclassification and missed classification rates for each class, and F1, as their harmonic mean, gives a balanced view of the classification accuracy for all classes. The initial sample allocation was based on class area proportions. Accuracy metrics were calculated from the confusion matrix, with each evaluated sample given equal weight. Thus, the reported OA and class-specific metrics are unweighted sample statistics. They are not design-based, area-weighted accuracy estimates for the entire study area.

4. Results

4.1. Spatiotemporal Distribution of Valid Dynamic World Observations

Figure 3 shows the spatial distribution of the valid-observation rate for the Dynamic World land-cover product across China in 2020. Rates ranged from 0 to 76.14% and exhibited pronounced spatial variation, with an overall pattern of lower rates in southwestern China, higher rates in northwestern and northern China, and intermediate rates in eastern China. Areas with low valid-observation rates were concentrated mainly on the Tibetan Plateau and its eastern margin, in the Sichuan Basin, and in parts of southwestern China, with particularly low values in some mountain–valley systems and along basin margins. Moderate rates predominated in southern China, the middle and lower reaches of the Yangtze River, and parts of northeastern China. In contrast, most of Xinjiang, the Inner Mongolian Plateau, the North China Plain, and the Northeast China Plain had relatively high rates and more complete observations. Overall, the Dynamic World valid-observation rate showed pronounced spatial variation that was broadly consistent with regional differences in climate and topography.

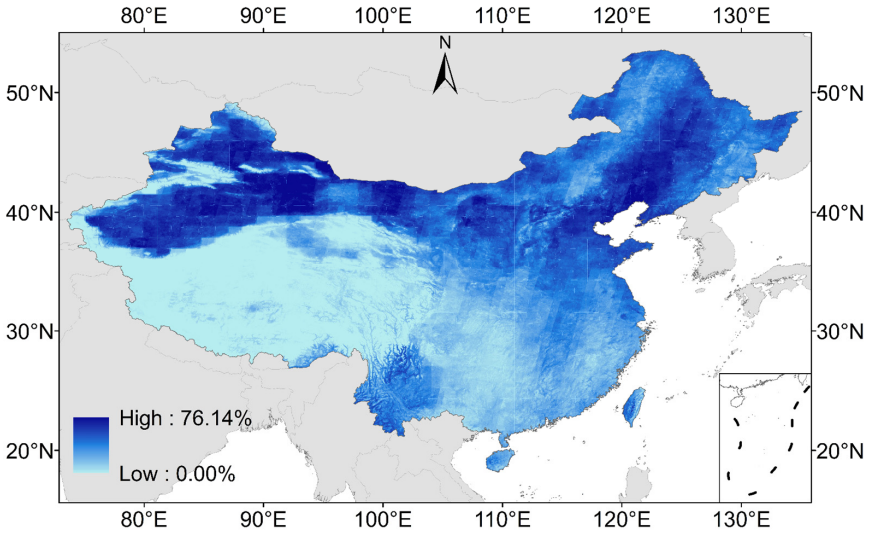


Figure 3. Spatial distribution of the annual valid-observation rate of Dynamic World data in 2020.

Figure 4 shows the monthly variation in the national average valid-observation rate. The average valid-observation rate remained at about 30% from January to May, decreased to its lowest value in June, about 24%, and remained at a low level of about 25–26% from July to August. After September, it gradually increased, reached its highest value of about 41% in November, and decreased to about 35% in December. There were obvious seasonal differences in the valid-observation rate, which was characterized by low values in summer and increasing values in autumn. There were obvious spatial and temporal differences in the valid-observation rate of Dynamic World, and the observation conditions that could be

used for compositing in different regions and months were not balanced. Therefore, it was necessary to further compare the spatial completeness and classification accuracy of the composited results using different time windows.

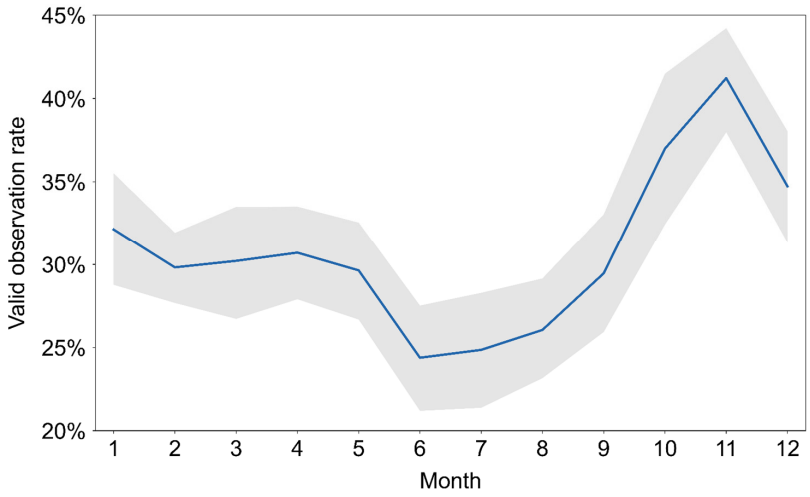


Figure 4. Monthly variation in the Dynamic World valid-observation rate (light-blue shading denotes the 90% confidence interval).

4.2. Performance of Different Compositing Strategies

4.2.1. Monthly Compositing

Figure 5 shows the spatial coverage of the three compositing methods at the monthly scale and the area proportion of each land-cover class. The three methods produced the same spatial coverage in each month, with a mean monthly spatial coverage of 68.78% and a range of 59.06–80.84%; the minimum occurred in June, and the maximum occurred in August. Although class-area proportions differed slightly among methods, the monthly patterns were similar. The variations in Bare Ground and Snow & Ice were the most prominent. The area proportion of Bare Ground was highest in summer and autumn, reaching 27.49–27.78% in September, whereas Snow & Ice showed the opposite pattern, increasing substantially in winter and reaching 27.08–30.61% in January. In contrast, Water, Flooded Vegetation, and Built-Up Area varied relatively little throughout the year.

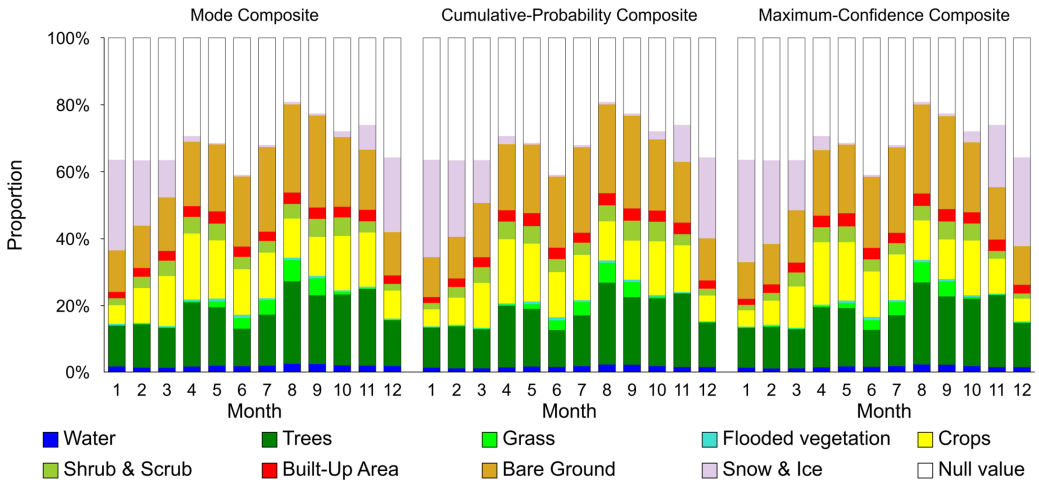


Figure 5. Monthly proportions of valid pixels and area fractions of individual land-cover classes for the three compositing methods.

Table 2 summarizes the overall accuracy (OA) and Kappa coefficient of the monthly composites produced by the three methods. Mode compositing achieved the highest mean OA (63.50%), with a mean Kappa coefficient of 0.56. Cumulative-probability compositing ranked second, with corresponding values of 63.11% and 0.56, whereas maximum-confidence compositing performed less well, at 61.99% and 0.55, respectively. For all three methods, accuracy increased from winter to summer and then declined. OA was lowest in January, ranging from 39.59% to 41.91%, and highest in July, ranging from 74.33% to 75.09%. The relatively low winter accuracy may be attributable to insufficient valid observations and to widespread snow and ice, which can obscure the spectral signatures of underlying surface features.

Table 2. Overall accuracy and Kappa coefficient of the monthly composite results.

Month	Mode Compositing		Cumulative-Probability Compositing		Maximum-Confidence Compositing	
	OA (%)	Kappa	OA (%)	Kappa	OA (%)	Kappa
1	41.91	0.34	40.67	0.33	39.59	0.32
2	51.92	0.44	49.40	0.42	47.45	0.40
3	58.12	0.50	57.29	0.49	56.23	0.48
4	68.12	0.60	67.99	0.60	67.29	0.59
5	70.97	0.64	70.92	0.64	70.75	0.64
6	71.45	0.64	71.80	0.65	71.62	0.65
7	74.33	0.68	75.09	0.69	74.67	0.68
8	72.84	0.66	73.51	0.67	73.31	0.67
9	71.84	0.65	72.49	0.66	72.01	0.65
10	67.70	0.60	67.49	0.60	67.00	0.60
11	64.00	0.56	62.99	0.55	57.95	0.50
12	48.74	0.41	47.62	0.40	45.97	0.38
Mean	63.50	0.56	63.11	0.56	61.99	0.55

Due to the large number of monthly scale compositing results, the compositing results for January and July were selected for analysis (Figure 6). The spatial land-cover patterns obtained with the three compositing methods were generally consistent, and the difference between the methods was less than the difference between months. Data gaps in January were mainly concentrated in the Qinghai–Tibet Plateau, accompanied by some strip-shaped and irregular data gaps; in July, missing data in Central China, South China, and parts of Southwest China was more obvious. This shows that the spatial coverage of the monthly composite results varied with season and region.

4.2.2. Growing-Season and Annual Compositing

The spatial coverage of the three compositing methods was 98.52% and 98.57% using the growing-season and annual windows, respectively. Both are significantly higher than the monthly compositing results. However, the difference between annual and growing-season compositing was only about 0.05%, indicating that extending the window further brings only a small gain in spatial coverage.

To justify the choice of the April to October timeframe, we compared the accuracy of composite results generated using three alternative windows (Table 3). Both cumulative-probability and maximum-confidence compositing achieved the highest OA and Kappa values from April to October, while mode compositing performed best from April to September. Therefore, selecting April to October as the growing season balances overall classification accuracy and provides a consistent temporal baseline for subsequent analysis.

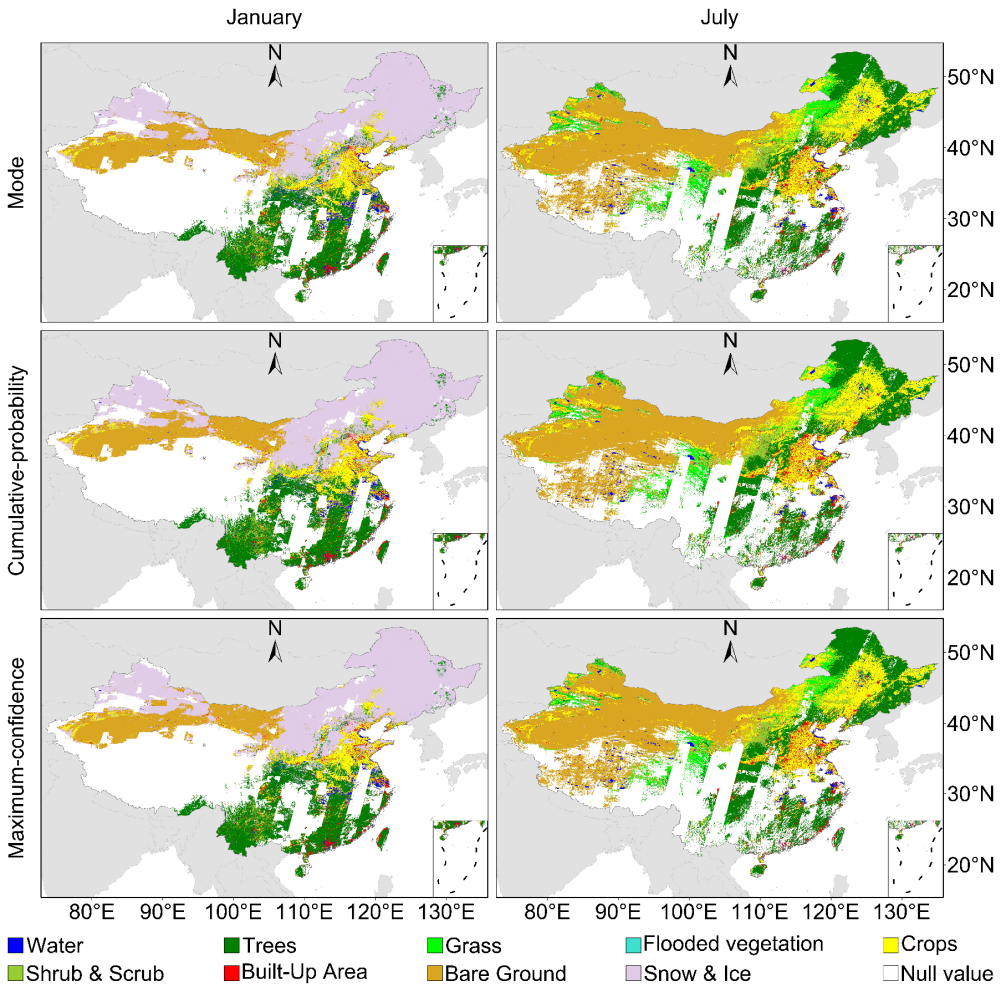


Figure 6. Monthly land-cover composites generated using the three compositing methods.

Table 3. Overall accuracy and Kappa coefficients for composites generated using four candidate growing-season windows.

Compositing Method	Temporal Window	OA (%)	Kappa
Mode compositing	4–10	75.94	0.7007
	3–9	75.92	0.7005
	4–9	76.05	0.7021
	5–10	75.75	0.6985
Cumulative-probability compositing	4–10	76.19	0.7058
	3–9	75.95	0.7012
	4–9	76.11	0.7030
	5–10	76.06	0.7024
Maximum-confidence compositing	4–10	73.27	0.6682
	3–9	70.89	0.6424
	4–9	73.18	0.6666
	5–10	73.21	0.6674

Table 4 presents the overall accuracy (OA) and Kappa coefficient for the three compositing methods using the growing-season and annual windows. The confusion matrices for the six compositing strategies are provided in Tables A3–A8. Using the growing-season window, cumulative-probability compositing achieved the highest OA (76.19%) and Kappa coefficient (0.71), followed by mode compositing (OA = 75.94%; Kappa = 0.70). Maximum-

confidence compositing performed less well (OA = 73.27%; Kappa = 0.67). However, growing-season mode compositing achieved a higher macro-F1 score than cumulative-probability compositing. Using the annual window, mode compositing achieved the highest OA (72.48%), with a Kappa coefficient of 0.66, while cumulative-probability compositing achieved an OA of 72.07% and a Kappa coefficient of 0.66. Maximum-confidence compositing declined to an OA of 60.63% and a Kappa coefficient of 0.54.

Table 4. Overall classification accuracy, Kappa coefficient, and macro-F1 score for the growing-season and annual composites.

Temporal Window	Compositing Method	OA (%)	Kappa	Macro-F1 (%)
Growing season	Mode compositing	75.94	0.70	64.96
	Cumulative-probability compositing	76.19	0.71	64.18
	Maximum-confidence compositing	73.27	0.67	57.00
Annual	Mode compositing	72.48	0.66	56.97
	Cumulative-probability compositing	72.07	0.66	56.30
	Maximum-confidence compositing	60.63	0.54	49.85

It should be noted that the decline in the accuracy of annual maximum-confidence compositing results from the combined effects of seasonal changes and the compositing method. Although multi-season observations include useful phenological information, these three methods only summarize category labels or probabilities and do not make use of the chronological sequence of phenological changes. Mode and cumulative-probability compositing combine multiple observations, whereas maximum-confidence compositing may assign the final class based on a single high-confidence observation of winter snow. The latter may therefore be more sensitive to pronounced but short-lived seasonal changes, consistent with its larger accuracy decline using the annual window. All three methods had lower OAs using the annual window, and the decrease was even more pronounced for maximum-confidence compositing. Overall, extending the compositing period to the full year slightly increased spatial coverage but reduced classification accuracy.

Using the growing-season window, cumulative-probability compositing had an OA only 0.25 percentage points higher than that of mode compositing, whereas mode compositing had a higher macro-F1 score. We therefore examined the three growing-season compositing methods further. For the 13,241 valid growing-season samples used in Table 4, the OA difference between cumulative-probability and mode compositing was not significant (McNemar test, $p = 0.1352$). Spatial block resampling gave 95% confidence intervals of −0.29–0.75 percentage points for the OA difference and −2.21–0.66 percentage points for the macro-F1 difference. Both intervals included zero, providing no clear evidence that either method performed better overall. To make the comparisons consistent across methods and temporal windows, we also evaluated the composites using 3641 common sample locations (Table A1). For these samples, growing-season cumulative-probability compositing had an OA that was 0.77 percentage points higher and a macro-F1 score that was 4.05 percentage points higher than those of mode compositing. The McNemar test showed a significant OA difference ($p = 0.0104$), assuming that sample points were independent. We also resampled groups of nearby points to check whether the differences remained when spatial dependence was considered. The 95% confidence intervals were 0.08–1.48 percentage points for the OA difference and 2.03–6.13 percentage points for the macro-F1 difference. Both intervals remained above zero for all the spatial groupings tested. These results support

the use of growing-season cumulative-probability compositing for the evaluated common samples; however, its advantage may vary among land-cover classes.

Table 5 shows the accuracy of the three growing-season compositing methods across climate zones. Although cumulative-probability compositing had a slightly higher national OA than mode compositing, its performance varied by region. It achieved higher OA, Kappa, and macro-F1 scores in the Northern Temperate, Southern Temperate, Northern Subtropical, Middle Subtropical, Southern Subtropical, Northern Tropical, and Middle Tropical zones, but lower values in the Middle Temperate and Plateau Climatic zones. Maximum-confidence compositing generally performed less well. These results suggest that the choice of compositing strategy should consider both the temporal window and regional differences.

Table 5. Classification accuracy of the three growing-season compositing methods across climate zones.

Climate Zone	Valid Samples	Mode Compositing			Cumulative-Probability Compositing			Maximum-Confidence Compositing		
		OA (%)	Kappa	Macro-F1 (%)	OA (%)	Kappa	Macro-F1 (%)	OA (%)	Kappa	Macro-F1 (%)
Northern Temperate	229	77.73	0.37	37.42	82.97	0.55	51.52	78.17	0.41	38.80
Middle Temperate	4397	69.29	0.62	61.75	68.82	0.61	60.43	63.53	0.55	50.81
Southern Temperate	2340	81.75	0.76	54.38	83.20	0.78	62.43	78.92	0.72	53.34
Northern Subtropical	802	81.05	0.74	42.16	82.67	0.76	44.97	81.30	0.74	41.98
Middle Subtropical	1781	80.56	0.63	44.60	81.18	0.64	48.17	80.17	0.62	43.42
Southern Subtropical	625	78.72	0.62	45.86	79.20	0.63	48.95	77.44	0.59	42.10
Northern Tropical	136	63.24	0.50	31.49	66.18	0.54	39.55	65.44	0.52	33.36
Middle Tropical	15	46.67	0.34	32.72	53.33	0.39	34.00	46.67	0.33	32.22
Plateau Climatic	2916	77.28	0.66	61.00	76.46	0.66	60.78	76.46	0.65	56.61

Figure 7 shows the compositing results of the three methods using the growing-season and annual windows. In general, the spatial patterns obtained by different time scales and compositing methods were relatively consistent. Bare Ground occurred mainly in northwestern China, Trees were concentrated in northeastern China and in the mountainous areas of southern China, and Crops were concentrated in the eastern plains. Differences mainly occurred in the distribution of Snow & Ice in parts of northern China and in local areas with interspersed land-cover classes. In the growing-season composites, the mode and cumulative-probability composites were similar, whereas the maximum-confidence composite contained more Snow & Ice in some high-latitude and high-elevation areas. Using the annual window, the mapped extent of Snow & Ice increased in parts of northeastern China in both the mode and cumulative-probability composites. The maximum-confidence composite showed more pronounced changes and differed substantially from the other two composites in parts of Inner Mongolia and northeastern China. To examine these spatial differences further, in the next subsection, we compare the performance of the compositing strategies based on class-specific classification accuracy.

To compare spatial differences among compositing methods within the same temporal window, spatial overlay analysis was used to assess agreement among the classification results produced by the three methods using the growing-season and annual windows. Pixels for which all three compositing methods assigned the same class were categorized as complete agreement; those for which only two methods agreed were categorized as partial agreement; and those for which all three methods assigned different classes were categorized as no agreement [31]. If one method returned a missing value, the pixel was categorized as partial agreement when the other two methods agreed and as no agreement

when they differed. Pixels with missing values from two or more methods were excluded from the analysis. Figure 8 shows the spatial distribution of agreement among the three compositing methods using the growing-season and annual windows. Using the growing-season window, pixels with complete agreement accounted for 85.65%, while pixels with partial agreement and no agreement accounted for 13.38% and 0.97%, respectively. Pixels with partial agreement occurred mainly in high-latitude regions, including northeastern China and Xinjiang, and were scattered across central China and the Tibetan Plateau. Using the annual window, the proportion of pixels with complete agreement decreased to 62.41%, whereas the proportions with partial agreement and no agreement increased to 35.24% and 2.35%, respectively. Compared with growing-season composites, areas of disagreement were concentrated mainly in northern Xinjiang, Inner Mongolia, and parts of northeastern China. These areas largely coincided with the distribution of Snow & Ice shown in Figure 7, indicating that annual composites were more sensitive to the choice of compositing method.

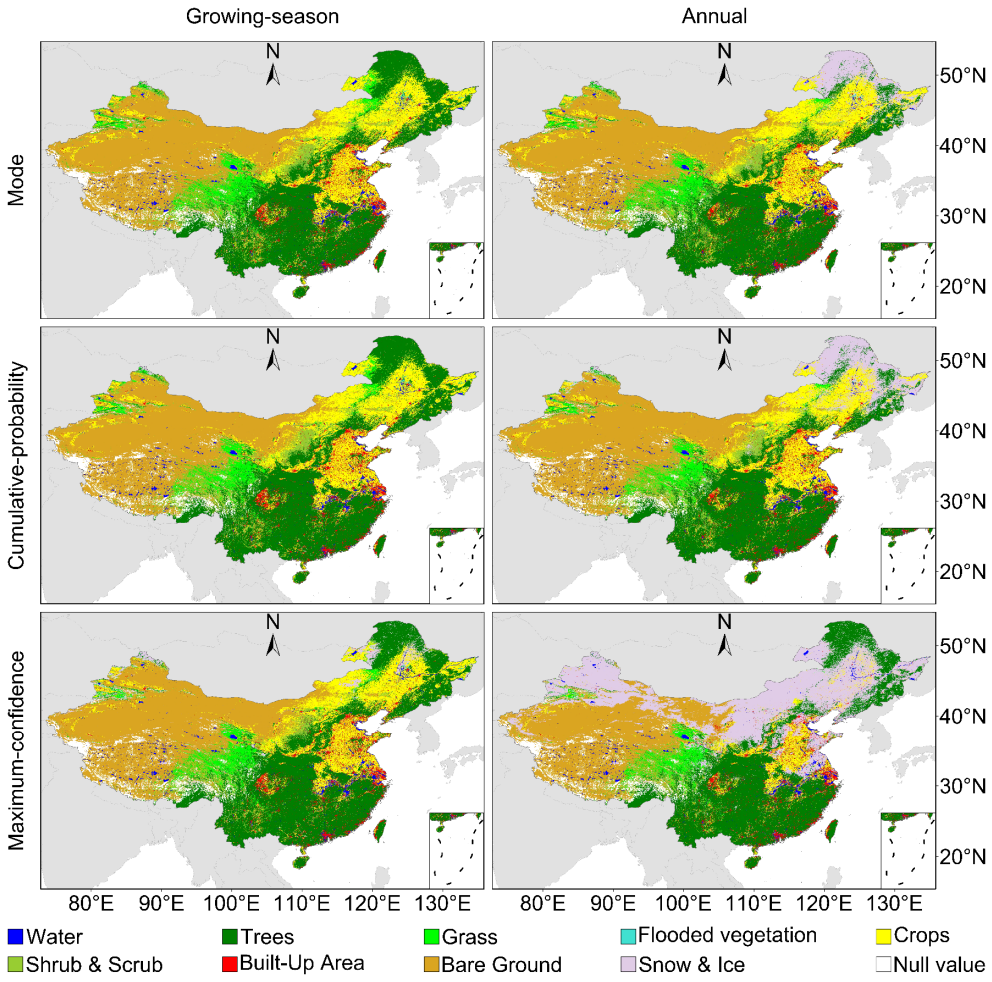


Figure 7. Growing-season and annual land-cover composites generated using the three compositing methods.

4.2.3. Compositing Performance Across Land-Cover Classes

Table 6 lists the valid and missing sample counts by land-cover class and temporal window. The accuracy assessments used 8488 samples for January, 8769 for July, 13,241 for the growing-season window, and 13,283 for the annual window. Longer windows reduced missing predictions, but only 65 of the 288 Snow & Ice samples had valid results in both the growing-season and annual windows. Samples with no classification results due to cloud masking and limited observations were excluded from the accuracy assessments.

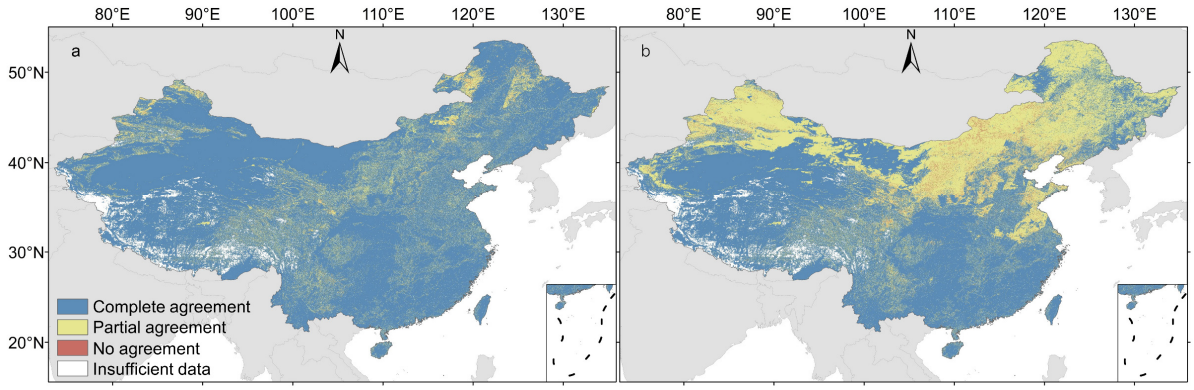


Figure 8. Spatial distribution of agreement among the land-cover composites generated using the three compositing methods: (**a**) growing-season window; (**b**) annual window.

Table 6. Numbers of samples with valid and missing predictions by land-cover class and temporal window (valid/missing).

Land-Cover Class	January	July	Growing Season	Annual
Water	184/192	257/119	373/3	373/3
Trees	2060/1045	1403/1702	3044/61	3077/28
Grass	744/948	1016/676	1617/75	1623/69
Flooded Vegetation	165/46	164/47	211/0	211/0
Crops	2050/363	1752/661	2413/0	2413/0
Shrub & Scrub	865/471	917/419	1310/26	1313/23
Built-Up Area	497/98	371/224	595/0	595/0
Bare Ground	1922/2104	2870/1156	3613/413	3613/413
Snow & Ice	1/287	19/269	65/223	65/223
Total	8488/5554	8769/5273	13,241/801	13,283/759

Figure 9 shows the monthly F1 scores for each land-cover class in the composites produced by the three methods. Overall, Trees, Built-Up Area, and Bare Ground maintained relatively high F1 scores in most months, followed by Crops and Water. The F1 score for Grass was low in winter but increased markedly during the growing season. Shrub & Scrub and Flooded Vegetation generally had low F1 scores, whereas Snow & Ice showed the greatest seasonal variation. The three compositing methods exhibited similar monthly trends for most classes, but no single method performed best across all classes and months.

Table 7 presents the class-specific accuracies of the three compositing methods using the growing-season window. Mode compositing and cumulative-probability compositing performed well overall, with macro-F1 scores of 64.96% and 64.18%, respectively, both higher than the 57.00% obtained by maximum-confidence compositing. Across the land-cover classes, Trees, Built-Up Area, and Bare Ground maintained relatively high F1 scores under all three methods, followed by Water and Crops, whereas Shrub & Scrub and Flooded Vegetation generally had low F1 scores. Cumulative-probability compositing achieved the highest F1 scores for Water, Trees, Flooded Vegetation, Crops, Shrub & Scrub, Built-Up Area, and Bare Ground, while mode compositing performed better for Grass and Snow & Ice. The differences between UA and PA provide further insight into class-specific errors. For Water, PA was at least 97.86% under all three methods, whereas UA ranged from only 62.99% to 64.56%, indicating that most reference water pixels were correctly identified but that pixels from other classes were also misclassified as Water. Grass showed the opposite pattern, with UA ranging from 71.29% to 73.79% and PA from only 32.71% to 40.07%, indicating that omission error was the main problem. The PA of Flooded Vegetation ranged from only 4.27% to 27.49%, indicating particularly severe omission. Using maximum-confidence

compositing, Snow & Ice had a PA of 86.15% but a UA of only 13.63%, indicating that many pixels classified as Snow & Ice actually belonged to other classes.

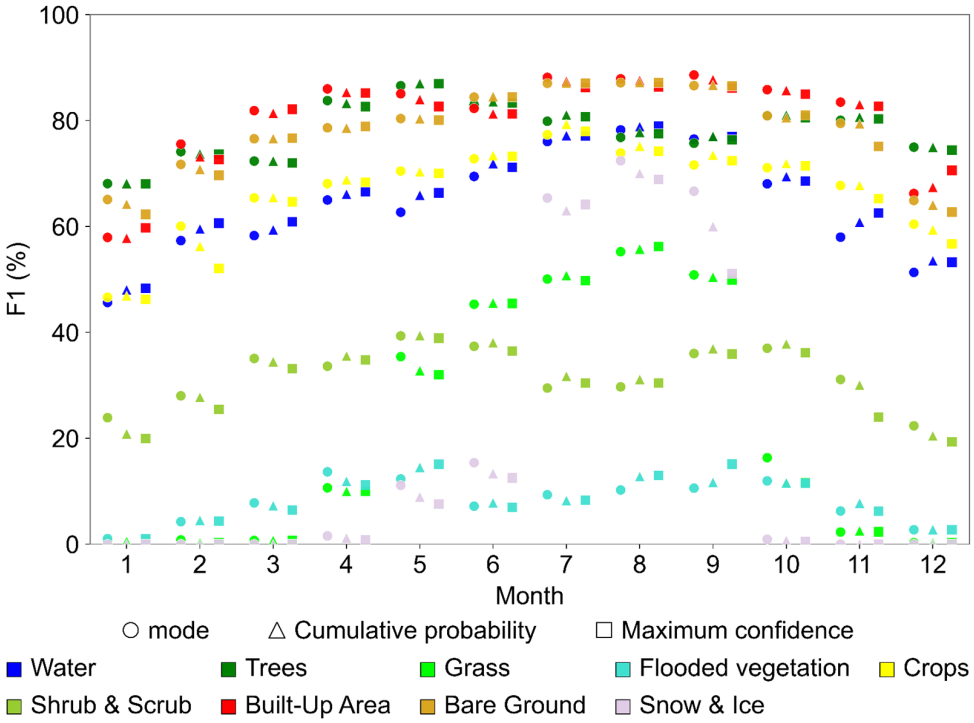


Figure 9. Monthly variation in class-specific F1 scores for the composites produced by the three methods.

Table 7. Class-specific accuracy of the three compositing methods using the growing-season window.

Land-Cover Class	Mode Compositing			Cumulative-Probability Compositing			Maximum-Confidence Compositing		
	UA (%)	PA (%)	F1 (%)	UA (%)	PA (%)	F1 (%)	UA (%)	PA (%)	F1 (%)
Water	63.81	97.86	77.25	64.56	98.66	78.05	62.99	98.12	76.73
Trees	77.16	96.55	85.77	78.37	96.78	86.61	74.24	96.65	83.97
Grass	72.08	40.07	51.51	73.79	37.79	49.98	71.29	32.71	44.85
Flooded Vegetation	23.08	4.27	7.20	67.44	27.49	39.06	38.71	5.69	9.92
Crops	78.25	77.08	77.66	79.06	77.91	78.48	75.42	72.86	74.11
Shrub & Scrub	43.14	32.90	37.33	42.34	33.97	37.70	45.61	22.98	30.56
Built-Up Area	80.41	93.11	86.29	80.17	93.78	86.44	74.83	93.95	83.31
Bare Ground	84.63	88.38	86.46	88.64	87.46	88.05	84.21	87.99	86.06
Snow & Ice	66.67	86.15	75.17	20.13	95.38	33.24	13.63	86.15	23.53

Table 8 presents the class-specific accuracies of the three compositing methods using the annual window. The mean F1 scores for mode compositing, cumulative-probability compositing, and maximum-confidence compositing were 56.97%, 56.30%, and 49.85%, respectively, all lower than those obtained using the growing-season window. Water continued to exhibit a high PA and a relatively low UA. The UA for Grass remained between 73.42% and 76.20%, whereas its PA declined to 25.69–32.29%, indicating more severe omission. The PA for Flooded Vegetation was only 1.42–3.79%, indicating that this class remained poorly identified. Snow & Ice showed the most pronounced change: its PA was 86.15% under all three methods, but its UA decreased to 1.76–7.46%, resulting in F1 scores of only 3.45–13.73%. Inspection of the confusion matrices revealed that only 65 valid sample points were available for Snow & Ice, possibly because a large proportion of samples in plateau regions had been removed as clouds during cloud masking. This indicates that many pixels belonging to other land-cover classes were misclassified as Snow & Ice when using the annual window.

Table 8. Class-specific accuracy of the three compositing methods using the annual window.

Land-Cover Class	Mode Compositing			Cumulative-Probability Compositing			Maximum-Confidence Compositing		
	UA (%)	PA (%)	F1 (%)	UA (%)	PA (%)	F1 (%)	UA (%)	PA (%)	F1 (%)
Water	64.86	95.98	77.41	65.21	95.98	77.66	64.06	96.51	77.01
Trees	79.41	87.23	83.13	79.79	88.04	83.71	78.26	95.81	86.15
Grass	74.64	32.29	45.08	76.20	30.38	43.44	73.42	25.69	38.06
Flooded Vegetation	42.11	3.79	6.96	45.45	2.37	4.50	18.75	1.42	2.64
Crops	76.77	75.18	75.96	78.65	72.69	75.55	86.26	48.11	61.77
Shrub & Scrub	41.86	33.89	37.46	41.94	32.67	36.73	40.75	13.25	20.00
Built-Up Area	82.23	91.76	86.74	82.41	92.94	87.36	77.99	91.09	84.03
Bare Ground	84.27	88.38	86.27	84.14	89.01	86.51	87.94	66.21	75.54
Snow & Ice	7.46	86.15	13.73	5.99	86.15	11.20	1.76	86.15	3.45

4.3. Comparison of Composite Results in Representative Areas

In order to further analyze spatial distribution differences in the compositing results using different time windows, six representative small areas were selected and compared with Google Earth high-resolution satellite images. Figure 10 shows, from left to right, high-resolution Google Earth imagery and the January, July, growing-season, and annual composite results obtained using the cumulative-probability compositing method. Area I is dominated by cropland, while Area II is dominated by forest and rivers. Area III contains a mix of forest, cropland, and built-up land. Area IV is dominated by cropland and artificial canals. Area V is dominated by natural vegetation and water bodies. Area VI mainly consists of bare mountains, with some snow and natural vegetation.

Overall, Water and Built-Up Area were identified more consistently in the July, growing-season, and annual windows, whereas vegetation classes—including Crops, Trees, Grass, and Flooded Vegetation—varied substantially. Specifically, the annual composite for Area I retained the main distribution of Crops, whereas the July and growing-season composites contained more scattered vegetation patches. In Area II, much of the main area was classified as Flooded Vegetation in the July composite, while the extent of Trees increased markedly in the growing-season and annual composites. In the lower part of Area III, the distribution of Built-Up Area closely matched the reference imagery, whereas the classification of cropland and natural vegetation varied substantially in the upper part. Both the July and growing-season composites for Area IV captured Crops and artificial canals, with more continuous cropland patches in the growing-season composite; however, large areas of cropland were classified as Snow & Ice in both the January and annual composites. In Area V, the growing-season composite identified the main waterways and surrounding Grass, whereas the annual composite classified much of the area as Snow & Ice, and the July composite contained more Trees patches. In addition, the January composites for all areas (except for Area VI) exhibited varying degrees of misclassification as Snow & Ice and differed markedly from the land cover shown in the reference imagery.

Data gaps associated with short temporal windows were most evident in Area VI. Both the January and July composites consisted entirely of missing values and therefore provided no valid land-cover information. The growing-season composite was dominated by Bare Ground and captured the overall distribution of exposed surfaces; however, a large gap remained on the left side. Annual compositing filled more of these gaps and produced a more complete spatial distribution of Bare Ground and surrounding vegetation but still failed to identify local snow cover accurately. These results indicate that the growing-season window can better represent seasonally sensitive land-cover types, such as cropland and natural vegetation, while annual compositing may produce extensive Snow & Ice. When data gaps remain in the growing-season composite, extending the compositing window to the full year could improve spatial completeness; however, an increase in valid pixels does not necessarily mean that all land-cover classes are accurately identified. When winter

observations are included, particular attention should be paid to misclassification caused by seasonal changes in surface conditions, such as snow cover.

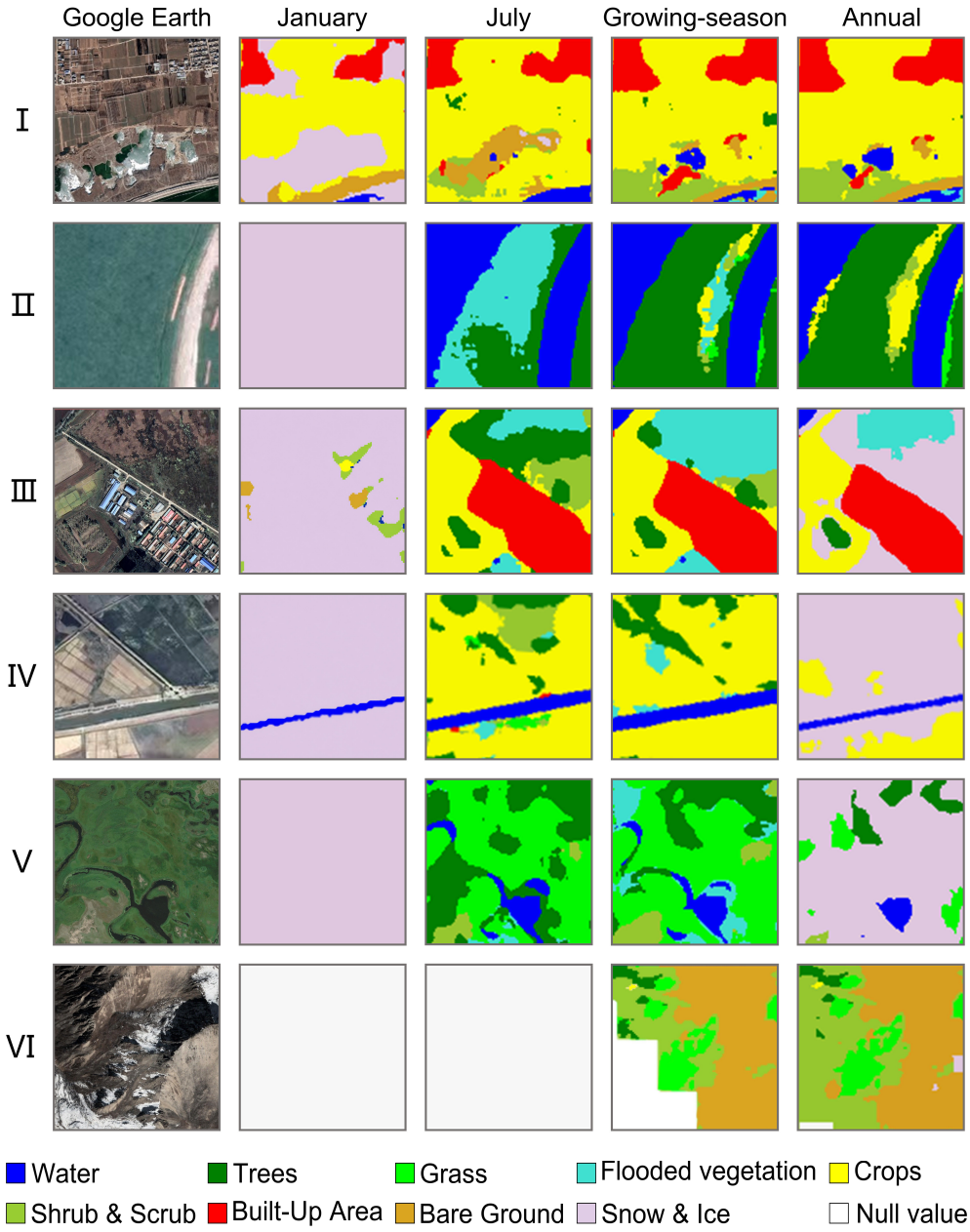


Figure 10. Comparison of composites in representative areas.

5. Discussion

The high-frequency update of Dynamic World provides important data for dynamic monitoring of regional land cover, but there is still a lack of systematic research on how to transform the scene-by-scene results into complete, continuous, and reliable land-cover products for specific tasks. Previous studies have mainly used time-composited Dynamic World products for land-cover mapping, product comparison, or multi-source data integration. However, there are relatively few studies on the joint evaluation of the effects of different time windows and compositing methods on Dynamic World classification [32,33]. By systematically comparing performance across temporal windows and compositing methods, this study assesses their effects on spatial completeness and classification accuracy, and provides guidance for multi-temporal Dynamic World compositing.

The choice of time window and compositing method should take into account spatial integrity and classification accuracy. Cumulative-probability compositing using the growing-season window achieved the highest overall accuracy (76.19%) while the spatial coverage reached 98.52%. Increasing the window to a year increases the number of observations, but the spatial coverage gain is limited, and the classification accuracy was significantly reduced, which may be related to the fact that the annual window includes surface conditions such as snow and vegetation in different seasons. Within the growing-season window, cumulative-probability compositing achieved the highest OA, whereas mode compositing achieved the highest macro-F1. Analysis of the common sample locations supported the advantage of cumulative-probability compositing. When the task needs to be included in the annual observation, the performance of mode compositing is better than that of the two other methods.

Although growing-season cumulative-probability compositing and mode compositing achieved better overall performance, there are differences in the temporal characteristics of the land-cover types and the conditions affecting their identification. Therefore, the time window and compositing method should be selected according to the target class. Vegetation classes such as Crops, Grass, and Trees exhibit pronounced phenological variation. The growing-season window captures much of the active growing period in many regions, but different vegetation classes respond differently to the compositing methods. Within this window, cumulative-probability compositing performed better for Trees and Crops, whereas mode compositing performed better for Grass. Maximum-confidence compositing relies on the highest confidence from a single observation and is therefore susceptible to anomalously high-confidence predictions. All three methods still showed substantial omission errors for Grass and Flooded Vegetation. For Snow & Ice, the low UA of the annual composites indicates that many pixels belonging to other reference classes were misclassified as Snow & Ice. These class-specific differences show that overall accuracy alone is insufficient; the compositing window and method should be adjusted according to the seasonal characteristics and commission and omission errors when selecting a strategy for a target class.

This study still has some limitations regarding the reference samples, spatial and time ranges, and compositing method design. Firstly, due to limitations of visual interpretation costs and data acquisition conditions, although 14,042 sample points were selected in this study, samples of low-proportion land types such as wetlands and Snow & Ice are still relatively sparse. The limited temporal availability of Google Earth high-resolution imagery made it difficult to establish reference labels representing actual surface conditions for each temporal window. Therefore, we used the same reference samples to assess how well the composites represent stable annual land-cover types. However, temporary seasonal changes may reflect real surface conditions, so disagreement with the annual reference labels should not automatically be interpreted as scene-level classification errors. In addition, heterogeneous areas may contain multiple land-cover types at the 10 m scale, making it difficult to assign a single reference class through visual interpretation. We therefore prioritized relatively homogeneous areas and excluded samples with inconsistently interpreted classes across dates to obtain reliable and temporally stable reference samples. However, this screening also limited the representativeness of the reference data. The reported accuracies therefore mainly reflect the performance of the compositing strategies over relatively homogeneous and stable surfaces and may not fully represent their performance in complex or changing areas. Secondly, this study covered China in 2020 and used April–October as a fixed growing-season window. This window does not fully account for regional differences in phenology and cropping systems. Our findings therefore apply to the study area, year, and temporal windows evaluated here. They do not establish the optimal growing-season

compositing strategy for all regions and years. In the future, the cross-regional and cross-year applicability of the compositing strategy can be tested by combining multi-year data and regional phenological information. Finally, the three methods used do not fully consider the differences in observation quality at different time phases. In the future, quality information such as effective observation times, cloud and cloud shadow residues, and imaging conditions can be integrated to explore a quality-weighted compositing method and evaluate whether it improves the identification of easily confused land types. In addition, future work could also combine OA and macro-F1 through weighting to consider both overall accuracy and performance across land-cover classes.

6. Conclusions

Using China as the study area, this research examined the impact of different temporal windows and compositing methods on the multi-temporal compositing of Dynamic World, providing guidance for producing stable annual land-cover maps of China. The results indicate that effective observations in Dynamic World exhibit significant spatiotemporal differences. Multi-temporal compositing can significantly improve spatial completeness, with monthly average spatial coverage reaching 68.78%, and coverage during the growing season and across the entire year reaching 98.52% and 98.57%, respectively. Expanding the window from the growing season to the full year only increased coverage by 0.05%, but overall accuracy decreased. Using the growing-season compositing window, mode and cumulative-probability compositing achieved the highest macro-F1 (64.96%) and OA (76.19%), respectively. Both provided high spatial completeness, with each method showing advantages for different land-cover types. This conclusion is based on the national assessment for China in 2020 using the April–October window. However, classification performance varied among land-cover classes, with substantial omission or commission errors for Grass, Flooded Vegetation, and Snow & Ice. Therefore, in practical applications, the compositing strategy should be selected according to the target land-cover class. The research findings provide a reference for selecting multi-temporal compositing strategies for Dynamic World at large regional scales.

Author Contributions: Conceptualization, Y.X. (Yan Xu) and Y.S.; methodology, Y.X. (Yan Xu) and Y.S.; software, Y.X. (Yan Xu); validation, Y.X. (Yan Xu) and Y.S.; formal analysis, Y.X. (Yan Xu) and Y.S.; investigation, Y.X. (Yan Xu) and Y.S.; data curation, Y.X. (Yan Xu) and Y.S.; resources, Y.W. and Y.X. (Yongming Xu); visualization, Y.X. (Yan Xu); writing—original draft preparation, Y.X. (Yan Xu) and Y.S.; writing—review and editing, Y.X. (Yan Xu), Y.X. (Yongming Xu) and S.Z.; supervision, Y.X. (Yongming Xu) and S.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Jiangsu Provincial Department of Education through the national-level NUIST Students' Platform for Innovation Training Program, grant number 202510300094; and the National Natural Science Foundation of China, grant numbers 42271351, 42671484, and 42571424.

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

Acknowledgments: The authors acknowledge Google Earth Engine (GEE) for providing access to the land-cover product data. We also acknowledge Google Earth Pro for providing a platform and high-resolution imagery for sample selection and visual interpretation.

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

Appendix A

Table A1. Overall accuracy, Kappa coefficients, and macro-F1 scores of the three compositing methods across temporal windows, evaluated using common sample locations.

Temporal Window	Mode Compositing			Cumulative-Probability Compositing			Maximum-Confidence Compositing		
	OA (%)	Kappa	Macro-F1 (%)	OA (%)	Kappa	Macro-F1 (%)	OA (%)	Kappa	Macro-F1 (%)
January	34.03	0.24	25.42	32.19	0.23	24.56	29.94	0.21	24.72
February	44.80	0.34	33.06	41.86	0.31	32.09	38.62	0.29	30.88
March	57.76	0.46	39.86	56.41	0.45	39.08	54.00	0.42	38.95
April	66.55	0.55	46.78	66.90	0.56	46.69	65.26	0.54	45.76
May	70.97	0.62	51.47	71.00	0.62	51.70	70.83	0.62	51.52
June	73.50	0.65	52.54	74.07	0.66	53.32	73.85	0.66	52.73
July	74.95	0.67	53.42	75.89	0.69	54.01	75.01	0.67	52.74
August	73.14	0.65	54.34	73.94	0.66	55.18	73.36	0.66	54.88
September	71.90	0.64	52.91	72.95	0.65	53.79	72.07	0.64	52.89
October	70.20	0.61	49.70	69.79	0.61	48.95	68.77	0.59	47.73
November	62.62	0.52	42.17	60.81	0.50	42.15	52.40	0.43	40.18
December	41.91	0.32	31.05	40.51	0.30	31.13	37.82	0.29	30.46
Growing season	76.24	0.69	54.59	77.01	0.70	58.64	71.88	0.63	49.84
Annual	70.45	0.62	50.20	69.90	0.61	49.82	47.02	0.39	41.67

Table A2. The numbers of initially selected, supplementary, rejected, and retained samples by class.

Land-Cover Class	Initially Allocated Samples	Final Samples	Removed or Added Samples
Water	314	376	62
Trees	3906	3105	−801
Grass	1854	1692	−162
Flooded Vegetation	27	211	184
Crops	2424	2413	−11
Shrub & Scrub	1560	1336	−224
Built-Up Area	599	595	−4
Bare Ground	4218	4026	−192
Snow & Ice	99	288	189
Total	15,001	14,042	−959

Table A3. Confusion matrix for mode compositing using the growing-season window.

		Composite-Result Class								
	Class	Water	Trees	Grass	Flooded Vegetation	Crops	Shrub & Scrub	Built-Up Area	Bare Ground	Snow & Ice
Reference class	Water	365	0	0	0	3	1	0	3	1
	Trees	4	2939	17	0	17	35	16	13	3
	Grass	10	220	648	12	232	300	2	186	7
	Flooded Vegetation	102	17	22	9	23	7	3	28	0
	Crops	68	194	37	14	1860	98	109	33	0
	Shrub & Scrub	4	388	104	2	85	431	4	290	2
	Built-Up Area	1	2	5	0	9	3	554	21	0
	Bare Ground	15	49	66	2	148	124	1	3193	15
	Snow & Ice	3	0	0	0	0	0	0	6	56

Table A4. Confusion matrix for cumulative-probability compositing using the growing-season window.

		Composite-Result Class								
	Class	Water	Trees	Grass	Flooded Vegetation	Crops	Shrub & Scrub	Built-Up Area	Bare Ground	Snow & Ice
Reference class	Water	368	0	0	1	1	0	0	0	3
	Trees	3	2946	14	4	15	33	15	0	14
	Grass	11	217	611	12	248	318	3	71	126
	Flooded Vegetation	99	0	19	58	22	9	2	0	2
	Crops	67	174	28	7	1880	110	114	33	0
	Shrub & Scrub	4	373	96	2	83	445	3	301	3
	Built-Up Area	1	2	5	0	6	3	558	0	20
	Bare Ground	14	47	55	2	123	133	1	3160	78
	Snow & Ice	3	0	0	0	0	0	0	0	62

Table A5. Confusion matrix for maximum-confidence compositing using the growing-season window.

		Composite-Result Class								
	Class	**Water**	**Trees**	**Grass**	**Flooded Vegetation**	**Crops**	**Shrub & Scrub**	**Built-Up Area**	**Bare Ground**	**Snow & Ice**
Reference class	Water	366	0	0	1	1	0	0	1	4
	Trees	4	2942	15	0	14	26	14	9	20
	Grass	15	251	529	8	260	181	4	171	198
	Flooded Vegetation	97	15	12	12	23	6	7	26	13
	Crops	70	272	25	6	1758	57	146	51	28
	Shrub & Scrub	6	423	93	2	131	301	14	313	27
	Built-Up Area	1	3	4	0	3	4	559	19	2
	Bare Ground	19	57	64	2	141	85	3	3179	63
	Snow & Ice	3	0	0	0	0	0	0	6	56

Table A6. Confusion matrix for mode compositing using the annual window.

		Composite-Result Class								
	Class	**Water**	**Trees**	**Grass**	**Flooded Vegetation**	**Crops**	**Shrub & Scrub**	**Built-Up Area**	**Bare Ground**	**Snow & Ice**
Reference class	Water	358	0	0	0	1	0	0	1	13
	Trees	9	2684	11	0	13	42	9	20	289
	Grass	7	178	524	3	248	310	2	183	168
	Flooded Vegetation	88	6	14	8	20	11	2	26	36
	Crops	67	154	17	5	1814	126	101	37	92
	Shrub & Scrub	4	325	75	1	111	445	3	296	53
	Built-Up Area	1	3	4	0	6	6	546	27	2
	Bare Ground	15	30	57	2	150	123	1	3193	42
	Snow & Ice	3	0	0	0	0	0	0	6	56

Table A7. Confusion matrix for cumulative-probability compositing using the annual window.

		Composite-Result Class								
	Class	**Water**	**Trees**	**Grass**	**Flooded Vegetation**	**Crops**	**Shrub & Scrub**	**Built-Up Area**	**Bare Ground**	**Snow & Ice**
Reference class	Water	358	0	0	0	1	0	0	0	14
	Trees	6	2709	10	0	10	34	10	12	286
	Grass	8	170	493	2	227	309	2	192	220
	Flooded Vegetation	90	6	12	5	19	8	2	23	46
	Crops	67	147	14	2	1754	116	101	39	173
	Shrub & Scrub	4	325	69	0	97	429	2	310	77
	Built-Up Area	1	2	4	0	4	7	553	24	0
	Bare Ground	12	36	45	2	118	120	1	3216	63
	Snow & Ice	3	0	0	0	0	0	0	6	56

Table A8. Confusion matrix for maximum-confidence compositing using the annual window.

		Composite-Result Class								
	Class	**Water**	**Trees**	**Grass**	**Flooded Vegetation**	**Crops**	**Shrub & Scrub**	**Built-Up Area**	**Bare Ground**	**Snow & Ice**
Reference class	Water	360	0	0	0	1	0	0	0	12
	Trees	4	2948	10	0	10	26	12	6	61
	Grass	14	207	417	6	65	130	0	110	674
	Flooded Vegetation	90	11	6	3	14	1	4	14	68
	Crops	66	205	14	5	1161	29	117	27	789
	Shrub & Scrub	6	347	68	1	55	174	11	157	494
	Built-Up Area	1	5	3	0	4	3	542	8	29
	Bare Ground	18	44	50	1	36	64	9	2392	999
	Snow & Ice	3	0	0	0	0	0	0	6	56

References

1. Pflugmacher, D.; Rabe, A.; Peters, M.; Hostert, P. Mapping Pan-European Land Cover Using Landsat Spectral-Temporal Metrics and the European LUCAS Survey. *Remote Sens. Environ.* **2019**, *221*, 583–595. [CrossRef]
2. Franklin, J.; Serra-Diaz, J.M.; Syphard, A.D.; Regan, H.M. Big Data for Forecasting the Impacts of Global Change on Plant Communities. *Glob. Ecol. Biogeogr.* **2017**, *26*, 6–17. [CrossRef]

3. Friedl, M.A.; Sulla-Menashe, D.; Tan, B.; Schneider, A.; Ramankutty, N.; Sibley, A.; Huang, X. MODIS Collection 5 Global Land Cover: Algorithm Refinements and Characterization of New Datasets. *Remote Sens. Environ.* **2010**, *114*, 168–182. [CrossRef]
4. Guo, Z.T.; Wang, C.M.; Liu, X.; Yang, J.; Zhang, H.Y. Accuracy Assessment of FROM-GLC30 Dataset Based on Small Watershed Sampling Units in China. *J. Geo-Inf. Sci.* **2021**, *23*, 524–535. [CrossRef]
5. Legesse, D.; Ayenew, T. Effect of Improper Water and Land Resource Utilization on the Central Main Ethiopian Rift Lakes. *Quat. Int.* **2006**, *148*, 8–18. [CrossRef]
6. Chauhan, V.; Gupta, L.; Dixit, J. Machine Learning-Based Prediction of Land Use and Land Cover, and Estimation of Surface Runoff Using the GIS-Based NRCS-CN Model. *Environ. Sustain. Indic.* **2026**, *31*, 101269. [CrossRef]
7. Sterling, S.; Ducharne, A. Comprehensive Data Set of Global Land Cover Change for Land Surface Model Applications. *Glob. Biogeochem. Cycles* **2008**, *22*, GB3017. [CrossRef]
8. Brown, C.F.; Brumby, S.P.; Guzder-Williams, B.; Birch, T.; Hyde, S.B.; Mazzariello, J.; Czerwinski, W.; Pasquarella, V.J.; Haertel, R.; Ilyushchenko, S.; et al. Dynamic World, Near Real-Time Global 10 m Land Use Land Cover Mapping. *Sci. Data* **2022**, *9*, 251. [CrossRef]
9. Gong, P.; Liu, H.; Zhang, M.; Li, C.; Wang, J.; Huang, H.; Clinton, N.; Ji, L.; Li, W.; Bai, Y.; et al. Stable Classification with Limited Sample: Transferring a 30-m Resolution Sample Set Collected in 2015 to Mapping 10-m Resolution Global Land Cover in 2017. *Sci. Bull.* **2019**, *64*, 370–373. [CrossRef] [PubMed]
10. Karra, K.; Kontgis, C.; Statman-Weil, Z.; Mazzariello, J.C.; Mathis, M.; Brumby, S.P. Global Land Use/Land Cover with Sentinel-2 and Deep Learning. In *Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*; IEEE: New York, NY, USA, 2021; pp. 4704–4707. [CrossRef]
11. Dutta, R.; Das, A.; Aryal, J. Big Data Integration Shows Australian Bush-Fire Frequency Is Increasing Significantly. *R. Soc. Open Sci.* **2016**, *3*, 150241. [CrossRef] [PubMed]
12. Paul, O.; Nayem, A.B.S.; Sarker, A.; Rajib, R.H.; Jit, T.H.; Alam, A.; Amin, M.A.; Ali, A.A.; Rahman, A.M. BOLM High Resolution Land Use and Land Cover Dataset and Benchmark Results for the Rapidly Developing City of Dhaka Bangladesh. *Sci. Rep.* **2025**, *15*, 37689. [CrossRef] [PubMed]
13. Turner, W.; Spector, S.; Gardiner, N.; Fladeland, M.; Sterling, E.; Steininger, M. Remote Sensing for Biodiversity Science and Conservation. *Trends Ecol. Evol.* **2003**, *18*, 306–314. [CrossRef]
14. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the Radiometric and Biophysical Performance of the MODIS Vegetation Indices. *Remote Sens. Environ.* **2002**, *83*, 195–213. [CrossRef]
15. Wang, Y.; Xu, Y.; Xu, X.; Jiang, X.; Mo, Y.; Cui, H.; Zhu, S.; Wu, H. Evaluation of Six Global High-Resolution Global Land Cover Products over China. *Int. J. Digit. Earth* **2024**, *17*, 2301673. [CrossRef]
16. Tsendbazar, N.; Herold, M.; Li, L.; Tarko, A.; de Bruin, S.; Masiliunas, D.; Lesiv, M.; Fritz, S.; Buchhorn, M.; Smets, B.; et al. Towards Operational Validation of Annual Global Land Cover Maps. *Remote Sens. Environ.* **2021**, *266*, 112686. [CrossRef]
17. Feng, D.; Zhao, Y.; Yu, L.; Li, C.; Wang, J.; Clinton, N.; Bai, Y.; Belward, A.; Zhu, Z.; Gong, P. Circa 2014 African Land-Cover Maps Compatible with FROM-GLC and GLC2000 Classification Schemes Based on Multi-Seasonal Landsat Data. *Int. J. Remote Sens.* **2016**, *37*, 4648–4664. [CrossRef]
18. Venter, Z.S.; Barton, D.N.; Chakraborty, T.; Simensen, T.; Singh, G. Global 10 m Land Use Land Cover Datasets: A Comparison of Dynamic World, WorldCover and Esri Land Cover. *Remote Sens.* **2022**, *14*, 4101. [CrossRef]
19. Wang, Y.; Sun, Y.; Cao, X.; Wang, Y.; Zhang, W.; Cheng, X. A Review of Regional and Global Scale Land Use/Land Cover (LULC) Mapping Products Generated from Satellite Remote Sensing. *ISPRS J. Photogramm. Remote Sens.* **2023**, *206*, 311–334. [CrossRef]
20. Loveland, T.R.; Reed, B.C.; Brown, J.F.; Ohlen, D.O.; Zhu, Z.; Yang, L.; Merchant, J.W. Development of a Global Land Cover Characteristics Database and IGBP DISCover from 1 km AVHRR Data. *Int. J. Remote Sens.* **2000**, *21*, 1303–1330. [CrossRef]
21. 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]
22. Gemitzi, A.; Kopsidas, O.; Stefani, F.; Polymeros, A.; Bellos, V. A constantly updated flood hazard assessment tool using satellite-based high-resolution land cover dataset within Google Earth Engine. *Land* **2024**, *13*, 1929. [CrossRef]
23. He, K.; Wang, L.; Liu, Z. Global decarbonization corresponding with unseasonal land cover change. *Nat. Commun.* **2025**, *16*, 7884. [CrossRef] [PubMed]
24. Kerner, H.; Nakalembe, C.; Yang, A.; Zvonkov, I.; McWeeny, R.; Tseng, G.; Becker-Reshef, I. How accurate are existing land cover maps for agriculture in Sub-Saharan Africa? *Sci. Data* **2024**, *11*, 486. [CrossRef] [PubMed]
25. Lasko, K.; Pan, C.G.; Griffin, S.P. Winter tree and terrain shadow correction reveals land cover mapping errors in Sentinel-2 satellite imagery. *Ecol. Inform.* **2026**, *96*, 103809. [CrossRef]
26. Xu, P.; Tsendbazar, N.-E.; Herold, M.; de Bruin, S.; Koopmans, M.; Birch, T.; Carter, S.; Fritz, S.; Lesiv, M.; Mazur, E.; et al. Comparative validation of recent 10 m-resolution global land cover maps. *Remote Sens. Environ.* **2024**, *311*, 114316. [CrossRef]

27. Myers, D.T.; Jones, D.; Oviedo-Vargas, D.; Schmit, J.P.; Ficklin, D.L.; Zhang, X. Seasonal Variation in Land Cover Estimates Reveals Sensitivities and Opportunities for Environmental Models. *Hydrol. Earth Syst. Sci.* **2024**, *28*, 5295–5310. [CrossRef]
28. Rangel, A.; Terven, J.; Córdova-Esparza, D.-M.; Romero-González, J.-A.; Ramírez-Pedraza, A.; Chávez-Urbiola, E.A.; Willars-Rodríguez, F.J.; Alfonso-Francia, G. Tracking U.S. Land Cover Changes: A Dataset of Sentinel-2 Imagery and Dynamic World Labels (2016–2024). *Data* **2025**, *10*, 67. [CrossRef]
29. Li, Z.; Ota, T.; Mizoue, N. Attribution of Forest Disturbance Types Based on the Dynamic World Class Probability Data: A Case Study of Myanmar. *Int. J. Appl. Earth Obs. Geoinf.* **2024**, *134*, 104216. [CrossRef]
30. Rashid, M.; Ullah, S.; Farnaz; Farooq, S.; Haider, S.; Liso, I.S.; Parise, M. Integrated Data-Driven Multi-Criteria Analysis and Machine Learning Approaches for Assessment of Flood Susceptibility Mapping. *Water* **2026**, *18*, 844. [CrossRef]
31. Liu, L.; Zhang, X.; Gao, Y.; Chen, X.; Shuai, X.; Mi, J. Finer-Resolution Mapping of Global Land Cover: Recent Developments, Consistency Analysis, and Prospects. *J. Remote Sens.* **2021**, *2021*, 5289697. [CrossRef]
32. Arévalo, P.; Olofsson, P.; Woodcock, C.E. Continuous Monitoring of Land Change Activities and Post-Disturbance Dynamics from Landsat Time Series: A Test Methodology for REDD+ Reporting. *Remote Sens. Environ.* **2020**, *238*, 111051. [CrossRef]
33. Yu, H.; Li, G.; Liu, H.; Zhu, S.; Xu, J.; Dong, W.; Li, C.; Shi, J. Synergistic Fusion of Sentinel-1 and Sentinel-2 for Global LULC Mapping: The Multimodal Network LULC-Former and Dynamic World+ Dataset. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2026**, *19*, 2511–2524. [CrossRef]

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