



Article Global Land Cover Assessment Using Spatial Uniformity Validation Dataset

Yoshie Ishii ¹, Koki Iwao ² and Tsuguki Kinoshita ^{3,*}

- ¹ United Graduate School of Agricultural Science, Tokyo University of Agricultural and Technology, 3-21-1 Chuuo, Ami 300-0393, Ibaraki, Japan; s208497w@st.go.tuat.ac.jp
- ² Geological Survey of Japan, National Institute of Advanced Industrial Science and Technology, Tsukuba Central 7, Higashi 1-1-1, Tsukuba 305-8567, Ibaraki, Japan; iwao.koki@aist.go.jp
- ³ College of Agriculture, Ibaraki University, 3-21-1 Chuo, Ami 300-0393, Ibaraki, Japan

Correspondence: tsuguki.kinoshita.00@vc.ibaraki.ac.jp

Abstract: The Degree Confluence Project (DCP) is a volunteer-based validation dataset that comprises useful information for global land cover map validation. However, there is a problem with using DCP points as validation data for the accuracy assessment of land cover maps. While resolutions of typical global land cover maps are several hundred meters to several kilometers, DCP points can only guarantee an area of several tens of meters that can be confirmed by ground photographs. So, the objective of this study is to create a land cover map validation dataset with added spatial uniformity information using satellite images and DCP points. For this, we devised a new method to semiautomatically guarantee the spatial uniformity of DCP validation data points at any resolution. This method can judge the validation data with guaranteed uniformity with a user's accuracy of 0.954. Furthermore, we conducted the accuracy assessment for the existing global land cover maps by the DCP validation data with guaranteed spatial uniformity and found that the trends differed by class and region.

check for updates

Citation: Ishii, Y.; Iwao, K.; Kinoshita, T. Global Land Cover Assessment Using Spatial Uniformity Validation Dataset. *Remote Sens.* **2021**, *13*, 2950. https://doi.org/10.3390/ rs13152950

Academic Editor: Fernando Camacho

Received: 30 June 2021 Accepted: 23 July 2021 Published: 27 July 2021

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



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). **Keywords:** Degree Confluence Project; global land cover map; land cover map validation; spatial uniformity; support vector machine

1. Introduction

In the land cover mapping process, accuracy assessment is an expensive yet essential step [1]. In particular, it is almost impossible to conduct a field survey to collect the required amount of validation data in a reasonable period of time to assess the accuracy of global land cover maps. One solution to these problems is the use of volunteer-based Geographic Information (VGI) [2]. VGI is provided by the project for use in a variety of applications, including land cover map validation. Projects that provide VGI for global coverage include Flickr (http://www.flickr.com/), OpenStreetMap (http://www.openstreetmap.org/), Panoramio (http://www.panoramio.com/), the Degree Confluence Project (DCP) (http: //confluence.org/), and the Geo-Wiki project (http://www.geo-wiki.org/). The DCP website provides ground-based photographs of points at integer latitudes and longitudes around the globe, and some studies have used these photographs to evaluate the accuracy of land cover maps. The first study that used DCP information as an accuracy evaluation of land cover maps was by Iwao et al. [3], who proposed a new land cover map validation method using DCP information. In that study, 749 confluences were used to evaluate the accuracy of land cover maps for the Eurasia region using Global Land Cover 2000 (GLC2000) [4], MODIS Land Cover (MOD12) [5], the University of Maryland's 1-km Global Land Cover products (UMD) [6], and the Global Land Cover Characteristics Data Base Version 2.0 (GLCC) [7]. In studies aimed at improving accuracy by integrating existing global land cover maps [8,9], a DCP-based land cover validation dataset was used for accuracy assessment. Foody and Boyd [10] also validated the availability of DCP

photographs by evaluating the accuracy of GlobCover for a region of tropical forests in West Africa. Soyama et al. [11] found that highly reliable reference data based on defined quality levels was produced for the IGBP land cover classification scheme using DCP information. Qian et al. [12] used DCP information as a land cover accuracy assessment in a study to consider the effect of uncertainty in map accuracy caused by references produced under the existing global land cover map matrix structure classification scheme.

As mentioned above, the potential use of DCP information in assessing the accuracy of land cover maps has already been shown in several studies. However, there is an issue regarding spatial representativeness between DCP points and global land cover maps. The resolution of a typical global land cover map is a few hundred meters to a few kilometers. On the other hand, DCP photos cannot confirm the land cover hundreds of meters or kilometers away from the DCP points. As a result, it may cause disagreement between the class that can be identified from the DCP photos and the true class of the land cover map at its particular resolution. If such validation data lacking spatial representativeness is used to evaluate the accuracy of the land cover map, it will be underestimated and cannot serve as validation data. Therefore, when ground validation data such as DCP are used to evaluate the accuracy of land cover maps, it is necessary to show that the spatial representativeness of the ground validation data is guaranteed. However, there has been no quantitative discussion of spatial representativeness for DCP validation data.

The objective of this study is to create a ground validation dataset that included the information of spatial uniformity using satellite images. For this, we devised a new method to semiautomatically guarantee the spatial uniformity of DCP validation data points at any resolution. The method constructs an SVM model using some DCP points as training data, and then determines whether each DCP point has spatial uniformity at a certain resolution or not. In addition, we conducted accuracy assessments for existing global land cover maps using a DCP validation dataset with guaranteed spatial uniformity.

2. Materials and Methods

In this study, we create a validation dataset with spatial uniformity information and evaluate the accuracy of the global land cover map using said validation dataset. The method for adding spatial uniformity information to the validation data is presented in Figure 1. First, we obtain a list of DCP points and download the satellite images for the corresponding latitude and longitude of a list of DCP. Then, a visual interpretation of the class is conducted for all DCP points, and that of spatial uniformity is performed for some DCP points. After that, the SVM model for automatic determination of spatial uniformity is constructed and the determination of spatial uniformity is performed for all DCP points. The data and methods required for these processes are described below.

2.1. Materials

2.1.1. The Degree Confluence Project (DCP)

DCP information was used to generate validation data for evaluating the accuracy of global land cover maps. The DCP is a volunteer-based ground validation dataset that provides field photographs and descriptions of sites at integer latitudes and longitudes around the globe, including the date of the survey and quality information. In this study, we used DCP points from 2003 to 2007. The reason that we chose this period is that we can assume that there will be little land cover change for five years, and the amount of DCP data in the five years centered on 2005 is the largest among 2000, 2005, 2010, and 2015.

2.1.2. Satellite Images

In this study, the semiautomatic determination of spatial uniformity was performed using information from satellite images. The satellite images used in this study were ASTER L1T Radiance (hereafter referred to as ASTER) from 2003 to 2007, Landsat Global Land Survey 2005 Landsat 5 + 7 scenes (hereafter referred to as GLS2005) whose Landsat images were acquired from 2003 to 2008, and Global PALSAR-2/PALSAR Yearly Mosaic (hereafter referred to as PALSAR) in 2007. The downloaded images are listed in Table 1. These satellite images were downloaded from Google Earth Engine Code (GEE) in the range of 300 m \times 300 m and 990 m \times 990 m centered on integer latitude and longitude. This is because these resolutions are commonly used in global land cover maps. These clipped satellite images are referred to as patches.



Figure 1. Flow chart of the method for adding spatial uniformity information to the validation data.

Table 1. List of satellite data downloaded from GEE

ASTER [13]	GLS2005 [14] (Combination of Landsat 5 and 7)	PALSAR [15]
B01 (band 1)	10 (band 1)	HH
B02 (band 2)	20 (band 1)	HV
B03 (band 3)	30 (band 3)	
	40 (band 4)	
	50 (band 5)	
	70 (band 7)	
	NDVI	
	NDBI	

The following preprocessing was performed on the downloaded satellite images: ASTER was downloaded from GEE with 0% cloud cover conditions. The processing for ASTER was to convert radiance to reflectance and to create NDVI from Band 2 and Band 3. The common preprocessing for ASTER and GLS2005 was to select the band with the highest mean NDVI value when there were multiple overlapping satellite images at the same latitude and longitude.

The rationale for choosing the time of maximum NDVI is that the larger the NDVI, the more likely it is that the area contains vegetation and the easier it is to determine whether it is uniform or non-uniform. As a common preprocessing step for ASTER, GLS2005, and PALSAR, if even one pixel in a patch contained "Nodata," the patch was discarded.

2.1.3. World Cities Database

The World Cities Database (https://simplemaps.com/data/world-cities) is a simple database of cities and towns in the world. In this research, when creating the training data for determining uniformity or non-uniformity by SVM, we created uniform training data for each class. Then, we referred to the World Cities Database to create additional training data since the Urban class lacks DCP points. Of these databases, we used the Basic database.

2.1.4. Google Earth

Google Earth was used in the visual interpretation of classes for the DCP validation data and to create the training data for the spatial uniformity judgment.

2.2. Method for Creating Validation Data Sets with Guaranteed Spatial Uniformity

This section describes a method for creating a validation dataset to evaluate the accuracy of global land cover maps with guaranteed spatial uniformity.

2.2.1. Visual Interpretation for Classes

Between 2003 and 2007, a total of 6243 points were visually inspected, including multiple visits to a single point. For the visual interpretation of the classes, we adopted the classification system of the IGBP [16] (MCD12 definitions) and further consolidated them into the seven classes defined in this study. The aggregation from the IGBP classes to the classes defined in this study was automated. DCP photographs and aerial photographs from Google Earth were used as references for visual interpretation. This work was carried out by two people, and only their results that were in agreement were considered valid.

2.2.2. Visual Interpretation for Uniformity/Non-Uniformity

In this study, in order to automate the judgment of uniformity and non-uniformity by SVM, we created the training data for SVM model constructing with some of the DCP and World Cities Database data. In Section 2.2.1, we used seven class definitions, but wetland and the mosaic of vegetation and cropland included in "Other" are not included in this study because they can be considered as non-uniform. Therefore, the six classes that were judged to be uniform or non-uniform were Forest, Grass/Shrub, Cropland, Urban, Barren, and Water. Stratified random sampling was used to obtain uniform training data with 100 points per class. Since some data that would be considered non-uniform were yielded, a class called "non-uniform" was also created for 100 points. These uniform and non-uniform training data were also created by two people who performed a visual interpretation of the data, and only the results that were in agreement were adopted. The criterion for judging uniformity was whether a class representing a 30 m \times 30 m area centered on an integer latitude and longitude contained 90% or more of the classes when viewed in a 300 m \times 300 m area.

2.2.3. Building the SVM Model

For each patch of the preprocessed satellite images, the mean and the sum of squared deviations (SSD) to the center pixel were calculated. The sum of deviation square *SSD* of the pixels contained in the patch, which centered a certain integer latitude and longitude, is defined as follows.

$$SSD = \frac{1}{n} \sum_{i=1}^{n} (p_i - p_0)^2$$
(1)

Here, p_i is the value of the *i*-th pixel, p_0 is the value of the pixel that is the center of the area, and *n* is the number of pixels contained in the patch. Since the mean and *SSD* were calculated for 14 bands, a total of 28 bands were used as input data for SVM.

In this study, two cost parameters were used for SVM. The intended use of the two cost parameters is to retain the "certainly uniform data" and screen out the "possibly non-uniform data" when determining spatial uniformity by SVM. In this study, the two

cost parameters were adjusted using LIBSVM version 3.25 software. The RBF (radial basis function) kernel was used in this SVM model.

In this study, nested cross-validation [17], which consists of outer cross-validation (OuterCV) and inner cross-validation (InnerCV), was used. The flow of nested cross-validation is shown in Figure 2. In nested cross validation, all data are first divided into test data and outer training data. In InnerCV, these outer training data are divided into validation data and inner training data, and five-fold cross-validation [18] is applied to determine the optimal hyper parameters. In OuterCV, five-fold cross-validation [18] is applied to the test data and outer training data to evaluate the model performance. The procedure for building the SVM model using nested cross-validation is shown below.



Figure 2. Flow of nested cross-validation.

Initially, all the training data with uniform/non-uniform visual interpretation were divided into five sets, one of which was used as the test data and the rest as outer training data. Then, the outer training data were further divided into five sets, one of which was used for validation data and the rest for inner training data. The inner training data were used to construct a two-class SVM model with each uniform and non-uniform class. The optimum parameters were searched by grid-search in the range of 2^{2i-1} (i = -2, -1, ..., 6) for C1 and C2, and 2^{2i-1} (i = -8, -7, ..., 2) for gamma. In general, the parameters are adjusted to maximize the mean OA (Overall Accuracy) of InnerCV, but in this study, the following original decision method was established:

Step 1: Create confusion matrix from InnerCV validation data.

Step 2: Find the parameter whose uniform UA (User Accuracy) $\left(=\frac{m_{11}}{m_{11}+m_{12}}\right)$ of the confusion matrix is closest to 100%.

Step 3: If there are multiple parameters that satisfy the condition of UA chosen in Step 2, choose the one with the highest OA.

Step 4: If there are multiple parameters selected in Step 3, select the one with the smallest cost parameter C1.

Maximizing OA in Step 3 is equivalent to maximizing m_{11} in Table 2 when UA is the same. The reason for maximizing m_{11} is that it is better to have a large number of valid DCP points with guaranteed uniformity. Step 4 is required as overfitting is generally more likely to occur when the cost parameter of SVM is large.

		Valid	UА	
		Uniformity	Non- Uniformity	
Classified	Uniformity	m_{11}	<i>m</i> ₁₂	$\frac{m_{11}}{m_{11}+m_{12}}$
Classifieu	Non-uniformity	<i>m</i> ₂₁	<i>m</i> ₂₂	$\frac{m_{22}}{m_{21}+m_{22}}$
РА		$\frac{m_{11}}{m_{11}+m_{21}}$	$\frac{m_{22}}{m_{12}+m_{22}}$	$OA = rac{\sum_{i=1}^2 m_{ii}}{\sum_{i=1}^2 \sum_{j=1}^2 m_{ij}}$

Table 2. Confusion matrix for hyperparameter setting.

In this study, SVM models were created for each class, and the final uniform/nonuniform decision was made by integrating the classification results by each model. If we denote by M_k the SVM model that discriminates a uniform class k (k = 1, 2, ..., n) from non-uniform when there are n classes, the result of the decision by the SVM model for any sample whose attributes are denoted by $x = (x_1, x_2, ..., x_{28})$ can be expressed as follows:

$$y_k = M_k(x), \ k = 1, 2, \dots, n.$$
 (2)

Here, $y_k = \{0, 1\}$, where 1—uniform and 0—non-uniform. Since the results of each model identify "certainly uniform data", if any SVM model judges a point to be uniform, that point is considered uniform. Therefore, the final judgment of uniformity or non-uniformity can be obtained as follows:

$$y_{\text{final}} = \bigvee_{k} y_k, \ k = 1, 2, \dots, n.$$
(3)

In this way, a certain sample is classified uniformly when $y_{\text{final}} = 1$ and non-uniformly when $y_{\text{final}} = 0$. Instead of determining uniformity or non-uniformity for all DCP points from the beginning, we perform a two-class classification of uniformity of a class versus non-uniformity for each class and, finally, integrate them. This is to account for the fact that the magnitude of the mean and the SSD from the center pixel when each class is uniform is different.

The means of uniform UA and that of OA obtained by the InnerCV, which is divided into five parts, are five ways because the OuterCV is also divided into five parts. Therefore, the mean of uniform UA and that of OA for the OuterCV are also obtained. In other words, a double mean is taken. As a result of this double averaging, the optimal parameter is the one that satisfies the conditions of Step 2 to Step 4.

After the hyperparameters were determined, the SVM model was reconstructed using the determined parameters C1, C2, and gamma for each class using the outer training data. The performance of the SVM model was evaluated by classifying the test data with the SVM model trained by the outer training data and obtaining the mean accuracy of the five datasets. Finally, we reconstructed the final SVM model using all the training data.

The number of training, validation, and test data used in each stage of the nested cross-validation process are shown in Table 3.

Table 3. Number of training, validation, and test data for nested cross-validation.

Class –		InnerCV		OuterCV		Final Model
		Training	Validation	Training	Test	Training
	Forest	64	16	80	20	100
	Grass/Shrub	64	16	80	20	100
T. T	Cropland	64	16	80	20	100
Uniformity	Urban	64	16	80	20	100
	Barren	64	16	80	20	100
	Water	64	16	80	20	100
Non-uniformity		64	16	80	20	100

When building the SVM model, we used the mean and SSD from the center pixel for a 300 m \times 300 m patch for each band. As long as the mean and SSD from the center pixel can be calculated, the same SVM model can be applied even if the target patch size is different. Therefore, in the evaluation of the accuracy of the existing global land cover map, we considered not only 300 m \times 300 m, but also 990 m \times 990 m as the patch size to judge the uniformity or non-uniformity of DCP.

2.2.4. Accuracy Assessment of Existing Global Land Cover Maps

The accuracy of existing global land cover maps was evaluated using the DCP validation data with and without the spatial uniformity information generated in this study. The target global land cover maps were The MODIS Land Cover Type Product (MCD12Q1; https://lpdaac.usgs.gov/products/mcd12q1v006/) v006, GLCNMO version2 (https://globalmaps.github.io/glcnmo.html), Global Land Cover 2000 (CLC2000; https://forobs.jrc.ec.europa.eu/products/glc2000/glc2000.php), and GlobCover Land Cover Maps (GlobCover; http://due.esrin.esa.int/page_globcover.php) (Table 4).

	MCD12 [19]	GLCNMO [20]	GLC2000 [4]	GlobCover [21]
Used data year Resolution	2005 500 m	2008 500 m	2003 1 km	2005–2006 300 m
Classification system	IGBP (17 classes)	FAO LCCS (20 classes)	FAO LCCS (22 classes)	FAO LCCS (22 classes)

Table 4. Information of global land cover maps.

2.2.5. Land Cover Class Definition

The class definitions for each global land cover map used in this study are shown in Table 5. A common class definition was also created for accuracy comparison. "Other" in the common class includes the class definition of each global land cover map, and those that cannot be evaluated for spatial uniformity were consolidated into one class. The definition of the common classes was performed by referring to the methods described in [22,23].

Table 5. Class definition.

Class No.	Common Close	MCD12 [1(]	
Class No.	Common Class	MCD12 [16]	GLUNMO [20]
1	Forest	Evergreen Needleleaf Forests Evergreen Broadleaf Forests Deciduous Needleleaf Forests Deciduous Broadleaf Forests Mixed Forests	Broadleaf Evergreen Forest Broadleaf Deciduous Forest Needleleaf Evergreen Forest Needleleaf Deciduous Forest Mixed Forest Tree Open
2	Grass/Shrub	Closed Shrublands Open Shrublands Woody Savannas Savannas Grasslands	Shrub Herbaceous Herbaceous with Sparse Tree/Shrub Mangrove
3	Cropland	Croplands	Cropland Paddy field
4	Urban	Urban and Built-up Lands	Urban
5	Barren	Barren	Bare area, consolidated (gravel, rock) Bare area, unconsolidated (sand) Sparse vegetation
6	Water	Permanent Snow and Ice Water Bodies	Snow/Ice Water bodies
7	Other	Permanent Wetlands Cropland/Natural Vegetation Mosaics	Cropland/Other Vegetation Mosaic

Class No.	Common Class GLC2000 [24]		GlobCover [25]		
1	Forest	Tree Cover, broadleaved, evergreen Tree Cover, broadleaved, deciduous, closed Tree Cover, broadleaved, deciduous, open Tree Cover, needle-leaved, evergreen Tree Cover, needle-leaved, deciduous Tree Cover, nixed leaf type Tree Cover, regularly flooded, fresh water Tree Cover, regularly flooded, saline water	Closed (>40%) needleleaved evergreen forest (>5 m) Open (15–40%) needleleaved deciduous or evergreen forest (>5 m) Closed to open (>15%) broadleaved evergreen or semideciduous forest (>5 m) Closed (>40%) broadleaved deciduous forest (>5 m) Open (15–40%) broadleaved deciduous forest/woodland (>5 m) Closed to open (>15%) mixed broadleaved and needleleaved forest (>5 m) Closed to open (>15%) broadleaved forest regularly flooded (semipermanently or temporarily—Fresh or brackish water		
2	Grass/Shrub	Shrub Cover, closed–open, evergreenShrub Cover, closed–open, deciduousHerbaceous Cover, closed–openRegularly flooded shrub and/or herbaceous cover	Mosaic forest or shrubland (50–70%)/grassland (20–50%) Mosaic grassland (50–70%)/forest or shrubland (20–50%) Closed to open (>15%) (broadleaved or needleleaved, evergreen or deciduous) shrubland (<5 m) Closed to open (>15%) herbaceous vegetation (grassland, savannas or lichens/mosses) Closed to open (>15%) grassland or woody vegetation on regularly flooded or waterlogged soil—fresh, brackish or saline water		
3	Cropland	Cultivated and managed areas	Postflooding or irrigated croplands (or aquatic) Rainfed croplands		
4	Urban	Artificial surfaces and associated areas	Artificial surfaces and associated areas (Urban areas > 50%)		
5	Barren	Sparse herbaceous or sparse shrub coverBare Areas	Sparse (<15%) vegetation Bare areas		
6	Water	Water Bodies Snow and Ice	Water bodies Permanent snow and ice		
7	Other	Mosaic: Tree Cover/Other natural vegetation Mosaic: Cropland/Tree Cover/Other natural vegetation Mosaic: Cropland/Shrub and/or grass cover	Mosaic vegetation (grassland/shrubland/forest; 50–70%)/cropland (20–50%) Closed (>40%) broadleaved forest or shrubland permanently flooded—saline or brackish water		

Table 5. Cont.

3. Results and Discussion

3.1. Results of Visual Interpretation of DCP

Between 2003 and 2007, a total of 6243 points, including multiple surveys of a single site, were visually interpreted for class. Of these, the number of DCP points that were the target six classes and in agreement by the visual interpretation of classes by two people was 4721 points. The remaining 1522 points (about 24.4%) were either "disagreement", "Other" class, or "impossible to determine" by a visual interpretation by two people. Therefore, the DCP validation data removed here are not included in the determination of uniformity/non-uniformity by the SVM model and in the global land cover accuracy assessment.

3.2. Adjusting SVM Parameters

The hyperparameters of the finally adopted model and the results of the performance evaluation are shown in Table 6. The mean of uniform UA of InnerCV in Table 6 is the highest in each class, and furthermore, the mean of OA is the highest under the highest uniform UA. Similarly, for the OutorCV, the highest mean of the uniform UA and the highest OA under the highest uniform UA for each class are written in Table 6. For the OuterCV, we also calculated the accuracy of the final model after integrating the SVM models of each class. In other words, the UA of that model when uniform DCP validation data are secured using the final SVM model constructed in this study was 0.954. In addition, the variance of uniform UA for the final model in the OuterCV was 8.90×10^{-4} .

	RBF Kernel SVM Parameter			Inne	erCV	OuterCV	
Class	C1	C2	Gamma	Mean of Uniform UA	Mean of OA	Mean of Uniform UA	Mean of OA
Forest	32	128	0.03125	0.934	0.796	0.942	0.790
Grass/Shrub	128	512	0.03125	0.891	0.785	0.894	0.790
Cropland	0.5	2	8	0.805	0.630	0.853	0.630
Urban	0.125	128	2	1.000	0.689	0.985	0.731
Barren	0.125	0.5	8	0.980	0.828	0.980	0.840
Water	0.125	2	2	1.000	0.938	1.000	0.950
Final model						0.954	0.746

Table 6. Results of adjusting SVM hyperparameters and model accuracy.

3.3. Guaranteed Spatial Uniformity

In this study, DCP validation data with guaranteed spatial uniformity in the 300 m and 990 m scale were generated. Tables 7 and 8 show the number of uniform and nonuniform DCP validation data determined by the constructed SVM model. Note that, of the 4721 points after visual interpretation, 1958 points at 300 m scale and 2114 points at 990 m scale were reduced to one if they overlapped at the same latitude and longitude and we removed the points for which no satellite image existed. So, the total number of DCP validation data is 2763 and 2607 points here, at 300 m and 990 m scale, respectively.

Table 7. Spatial uniformity of 300 m scale for DCP validation data.

Class	Uniformity	Non-Uniformity	Total	
Forest	327	272	599	
Grass/Shrub	567	423	990	
Cropland	220	369	589	
Urban	10	35	45	
Barren	245	51	296	
Water	225	19	244	
Total	1594	1169	2763	

Table 8. Spatial uniformity of 990 m scale for DCP validation data.

Class	Uniformity	Non-Uniformity	Total
Forest	251	307	558
Grass/Shrub	486	455	941
Cropland	134	420	554
Urban	4	38	42
Barren	223	59	282
Water	196	34	230
Total	1294	1313	2607

Figure 3 shows the proportion of DCP points with guaranteed spatial uniformity among all DCP points in each class. More than half of the DCP points of Forest, Grass/Shrub, Cropland, and Urban are judged to have a high possibility of non-uniformity at both 300 m and 990 m scales. On the other hand, only about 10–20% of Barren and Water were judged as having non-uniformity potential. These results mean the optimal resolution differs depending on the class. For example, more than half of Forest, Grass/Shrub, Cropland, and Urban are under-resolved at 300 m resolution, which is likely to have been underestimated in the DCP validation data before spatial uniformity was guaranteed. On the other hand, since Barren and Water have nearly not decreased from the originally created DCP, the results of accuracy assessment should be almost the same between the validation data before and after guaranteeing spatial uniformity at the 300 m and 990 m scales. The relationship between the screened amount of non-uniform validation data and the difference in the results in the accuracy assessment will be discussed in detail later in the accuracy assessment of existing global land cover maps.



Figure 3. The proportion of DCP points with guaranteed spatial uniformity among all DCP points in each class.

Figure 4 shows the distribution of (a) all DCP points and (b) DCP points with guaranteed uniformity at the 300 m scale. Note that the distribution of non-uniformity is not guaranteed to be non-uniform, only highly likely. In the same way, Figure 5 shows the distribution of DCP points at the 990 m scale.

Next, we compared the proportion of DCP points with guaranteed spatial uniformity among all DCP points by continent for North America, Asia, South America, Europe, Oceania, and Africa. Figures 6 and 7 show the class on the horizontal axis and the proportion of uniformity DCP points in all DCP points on the vertical axis. Note that "Water" focuses on the interior of the continent, and the ocean is not included. These results show that the smaller the proportion of Uniformity DCP/All DCP is, the more the validation data do not meet the spatial uniformity at the 300 m or 990 m scales. For example, Water in Africa has a proportion of 1.0, indicating that the exact same validation data can be used as when spatial uniformity is not taken into account. However, note that the distribution of DCP points is not necessarily representative of the entire continent, as Figures 4 and 5 show. For example, in Figure 6, there is a possibility that many forests in South America are smaller than the 300 m area. Intuitively, the mean resolution of forests in South America seems to be larger than the 300 m scale, but the discrepancy between the intuition and the results is due to the spatial bias of the DCP points as shown in Figure 4. From these results, we can obtain a trend of how much each class of DCP has spatial uniformity at the 300 m or 990 m scales for the regions where DCP points are distributed. Such information may be useful in



for discussing the optimal resolution for each class when mapping each region where DCP points exist.



(b) Uniform DCP points (1594 points)

Figure 4. Distribution of (a) all DCP points and (b) DCP points with guaranteed uniformity at the 300 m scale.



(a) All DCP points (2607 points)



(b) Uniform DCP points (1294 points)

Figure 5. Distribution of (a) all DCP points and (b) DCP points with guaranteed uniformity at the 990 m scale.

The method for guaranteeing spatial uniformity using information from satellite images in this research was created with 300 m \times 300 m as the training data, but it is possible to guarantee spatial uniformity at any resolution as long as the mean value and the sum of squared deviations from the center pixel are obtained from the satellite image in a certain area. We used this method to determine the uniformity or non-uniformity at the 990 m scale. The results of screening DCP validation data using this method showed that the number of DCP validation data was reduced by half for Forest, Grass/Shrub, Cropland, and Urban, while most of the data for Water and Barren remained at both the 300 m and 990 m scale. This means that the optimal resolution for each class of land cover classification may be obtained by looking at the variation in the amount of screening of non-uniformity for each class using the proposed method at, for example, 10 m, 100 m, 1 km, and 10 km.



Figure 6. Proportion of DCP points with guaranteed spatial uniformity among all DCP points per continent in the 300 m scale.



Figure 7. Proportion of DCP points with guaranteed spatial uniformity among all DCP points per continent in the 990 m scale.

3.4. Accuracy Assessment of Existing Global Land Cover Maps

In this study, we created a new DCP validation dataset with guaranteed spatial uniformity. Here, we evaluate the accuracy of the global land cover maps using the DCP validation dataset with guaranteed spatial uniformity and the validation dataset including all DCP points. As both validation datasets have deficiencies, the following points should be noted when the accuracy assessment of global land cover maps using these validation datasets. The validation dataset with guaranteed spatial uniformity has the deficiency of screening the validation data and leaves only spatially uniform points, where the bias of the population of the validation data points is relatively large. In this case, the accuracy assessment, DCP data do not guarantee spatial representativeness at a certain map resolution. In such cases, the accuracy assessment may be underestimated.

With these points in mind, we compared the accuracy of GlobCover using the DCP validation data with guaranteed spatial uniformity at the 300 m scale and with all DCP validation data. Table 9 shows the confusion matrices with and without spatial uniformity

guaranteed. The class numbers correspond to the following: 1—Forest, 2—Grass/Shrub, 3—Cropland, 4—Urban, 5—Barren, 6—Water, 7—Other. Note that "Other" is not included in the validation data in this study because it may contain mixtures.

	Uniform Validation/All Validation							
	Class No.	1	2	3	4	5	6	UA
	1	286/442	14/72	3/30	4/13	1/1	2/4	0.923/0.786
	2	26/75	249/419	46/94	1/11	6/10	2/4	0.755/0.684
	3	4/39	39/115	112/327	3/15	1/2	0/3	0.704/0.653
Map	4	1/1	0/1	0/1	2/6	0/0	0/0	0.667/0.667
	5	5/13	220/283	14/38	0/0	231/273	0/1	0.491/0.449
	6	0/1	1/4	0/2	0/0	6/10	220/290	0.969/0.931
	7	5/28	44/96	45/97	0/0	0/0	1/2	-
	РА	0.875/0.738	0.439/0.423	0.509/0.555	0.200/0.133	0.943/0.922	0.978/0.943	Total = 1594/2763
OA = 0.690/0.614.								

Table 9. Confusion matrix for GlobCover in 300 m scale.

Comparing the overall accuracy (OA), OA was 0.076 higher when evaluated with DCP validation data that guaranteed spatial uniformity than when evaluated with all DCP validation data. Looking at the user accuracy (UA) of each class, the UA of Forest, Grass/Shrub, and Cropland are 0.137, 0.071, and 0.051 higher, respectively, when evaluated with the DCP validation data that guarantee spatial uniformity than when evaluated with all DCP validation data. On the other hand, the difference was relatively low for Barren at 0.042 and Water at 0.038, and there was no change in Urban. As shown in Figure 3, the classes with larger amounts of screened data, which have a higher possibility of non-uniformity, showed a larger change in UA, while the classes with smaller amounts of screened data with a high possibility of non-uniformity is large and the accuracy improves, accuracy could be considered underestimated because the spatial representativeness of the screened validation data was not ensured.

Table 10 shows the results of accuracy assessment for GlobCover resampled to 1 km resolution. The OA was 0.09 higher when 990 m uniformity was guaranteed than when all DCP points were used. Looking at the breakdown of the improvement in accuracy, Forest, Grass/Shrub, and Urban improved their accuracy by about 10%, while Cropland, Barren, and Water improved relatively little.

Table 10. Confusion matrix for GlobCover in 990 m scale.

	Uniform Validation/All Validation							
	Class No.	1	2	3	4	5	6	UA
	1	223/414	14/65	4/30	3/12	0/1	2/4	0.907/0.695
	2	18/71	218/403	25/89	0/11	3/10	1/3	0.823/0676
	3	4/34	26/108	65/305	0/13	1/2	0/3	0.677/0.624
Map	4	0/1	0/1	0/1	1/6	0/0	0/0	1.000/0.500
	5	4/12	195/271	11/35	0/0	216/265	0/1	0.507/0.446
	6	0/1	1/3	0/2	0/0	3/4	193/217	0.980/0.934
	7	2/25	32/90	29/91	0/0	0/0	0/2	-
	PA	0.835/0.742	0.449/0.428	0.485/0.551	0.250/0.143	0.969/0.940	0.985/0.943	Total = 1294/2607

```
OA = 0.708/0.618.
```

Tables 11–13 show the results of the accuracy assessment of MCD12, GLCNMO, and GLC2000 when the spatial uniformity of the 990 m scale is guaranteed and when all DCP points are used, as in Table 10. In all maps, the accuracy of Forest and Grass/Shrub is much higher when the spatial uniformity is guaranteed than when all DCP points are used. On the other hand, for Barren and Water, the improvement in accuracy is relatively small depending on whether spatial uniformity is guaranteed or not for any map. However, Water in MCD12 shows an exceptionally large improvement in accuracy. This trend is consistent with the discussion in Figure 3, which shows that when the amount of screening of validation data with a high possibility of non-uniformity is large, the accuracy improves; when the amount of screening is small, there is almost no improvement in accuracy. However, despite the fact that Cropland has a large amount of screened validation data with a high possibility of non-uniformity, shown in Figure 3, the improvement in accuracy is small or may be lower in Tables 10-13. This is thought to be due to the fact that the mean of OA of Cropland's OuterCV by SVM is the lowest, as can be seen from Table 6. This low mean of OA means that the amount of erroneously screened uniform data included when screening non-uniform data is higher than other classes. In fact, checking Tables 10–13, only about 20% of the data in Cropland (row 3 column 3 elements in confusion matrix) is left when uniformity is ensured. On the other hand, in MCD12 and GLC2000, where the UA of Cropland is reduced, the accuracy of Cropland is considered to be reduced because the other classes have relatively more data remaining as uniform than Cropland.

Table 11. Confusion matrix for MCD12 in 990 m scale.

	Uniform Validation/All Validation							TTA
	Class No.	1	2	3	4	5	6	UA
	1	113/185	7/20	8/28	0/3	0/1	13/19	0.801/0.723
	2	99/269	387/737	58/198	2/16	18/28	18/27	0.665/0.578
	3	20/57	11/55	63/275	0/11	0/4	6/10	0.630/0.667
Map	4	3/9	3/6	0/13	1/7	0/0	6/6	0.077/0.171
-	5	2/2	69/94	1/5	0/0	203/246	3/5	0.780/0.699
	6	11/26	8/24	4/16	1/2	2/3	147/158	0.850/0.690
	7	3/10	1/5	0/19	0/3	0/0	3/5	-
	РА	0.450/0.332	0.796/0.783	0.470/0.496	0.250/0.167	0.910/0.872	0.750/0.687	Total = 1294/2607

OA = 0.706/0.617.

Table 12. (Confusion	matrix for	GLCNMO	in 990 m	scale.
-------------	-----------	------------	--------	----------	--------

	Uniform Validation/All Validation							T T A
	Class No.	1	2	3	4	5	6	UA
	1	237/450	28/121	12/67	1/6	0/1	1/8	0.849/0.689
	2	10/48	262/459	28/96	0/4	6/11	0/3	0.856/0.739
	3	1/31	24/66	70/288	0/15	0/3	0/4	0.737/0.708
Map	4	0/2	0/1	0/1	2/16	0/1	0/1	1.000/0.727
-	5	1/1	161/230	8/20	0/0	213/262	0/3	0.556/0.508
	6	2/3	0/0	0/0	0/0	3/3	195/210	0.975/0.972
	7	0/23	11/64	16/82	1/1	1/1	0/1	-
	РА	0.944/0.806	0.539/0.488	0.522/0.520	0.500/0.381	0.955/0.929	0.995/0.913	Total = 1294/2607

OA = 0.756/0.646.

	Uniform Validation/All Validation							
	Class No.	1	2	3	4	5	6	UA
	1	208/385	24/102	7/48	2/7	0/2	1/10	0.849/0.695
	2	13/68	216/395	27/93	0/9	5/12	4/7	0.856/0676
	3	13/55	34/112	69/327	0/17	0/2	4/11	0.737/0.624
Map	4	0/3	0/3	1/2	2/8	0/0	0/0	1.000/0.500
1	5	1/8	205/290	11/24	0/1	216/262	0/3	0.556/0.446
	6	3/6	0/4	0/2	0/0	1/2	187/199	0.975/0.934
	7	11/31	7/34	19/58	0/0	1/2	0/0	-
	РА	0.835/0.692	0.444/0.420	0.515/0.590	0.500/0.190	0.969/0.929	0.994/0.865	Total = 1292/2604
				OA = 0.694/(0.605			

Table 13. Confusion matrix for GLC2000 in 990 m scale.

Additionally, the "Other" (numbering 7 in confusion matrix) of each confusion matrix is actually a mixture of Cropland and Grass/Shrub, or a mixture of water bodies and vegetation such as wetlands. By calculating the following equation, we can check how much of the DCP determined to be non-uniform by SVM was actually a class of mixture (non-uniformity).

$$UA_{i=7(Other)} = \frac{\sum_{j=1}^{6} m_{ij}^{All \ validation} - \sum_{j=1}^{6} m_{ij}^{Uniform \ validation}}{\sum_{j=1}^{6} m_{ij}^{All \ validation}}, \ j = 1, 2, \dots, 6$$
(4)

Here, m_{ij} denotes the element of the confusion matrix with class number i = 7, j = 1, ..., 6. The calculation results of this equation for each map are shown in Table 14. In this study, the SVM model was constructed with the aim of increasing the uniform accuracy as much as possible at the expense of accuracy on the non-uniform side, but it can be confirmed that the non-uniform data can be separated with high accuracy in some maps.

Table 14. Accuracy of non-uniformity judgment.

Мар	GlobCover (300 m)	GlobCover (990 m)	MCD12	GLCNMO	GLC2000
$UA_{i=7(Other)}$	0.574	0.698	0.913	0.830	0.696

These results indicate that the use of DCP to evaluate the accuracy of land cover maps may have resulted in underestimation, especially for Forest and Grass/Shrub, because the data without spatial representativeness were used. On the other hand, it was clarified that the accuracy assessment for Barren and Water by DCP was reasonable in terms of spatial representativeness if the resolution was 990 m or higher. In the case of Cropland, the screening of potentially non-uniform DCP points had the effect of improving accuracy by allowing only uniform DCP validation data for accuracy assessment, which offset the decrease in accuracy caused by the screening of a large amount of validation data compared with other classes; as a result, the accuracy was almost the same before and after screening by SVM. The reason why the separation of uniform and non-uniform Cropland is more difficult than other classes is thought to be that Cropland may be exposed to soil, covered with vegetation, or both, depending on the crop calendar, and it is difficult to distinguish uniform and non-uniform features from a satellite image. In addition, the number of data in the original DCP for Urban was small, and screening of non-uniform data at 300 m and 990 m scales left almost no data. Furthermore, the discriminative performance of uniform Urban was as high as 0.985 from the OuterCV results (Table 6). From these results, the resolution at which spatial representativeness of Urban can be ensured is lower than 300 m. Considering the above discussion on spatial uniformity, we will discuss the accuracy assessment of the global land cover map using the DCP validation data that ensure uniformity. In GlobCover, GLCNMO, and GLC2000, half of Grass/Shrub is classified as Barren, while fewer data are classified as Barren in MCD12 (5 rows and 2 columns in confusion matrix). In MCD12, many of the data that are originally classified as Forest or Cropland are classified as Grass/Shrub. This is more pronounced than in other global land cover maps. This difference is thought to be due to the difference in the class definition. As shown in Table 4, the difference between GlobCover, GLCNMO, and GLC2000 is based on the definition of LCCS, while MCD12 is based on the definition of IGBP, which can be seen in this result.

4. Conclusions

In this study, we created DCP validation data with added information of spatial uniformity at 300 m and 990 m scales and compared the accuracy of the existing global land cover maps with the accuracy assessment by all DCP validation data. The judgment of uniformity and non-uniformity was made semiautomatically by the SVM classifier of the RBF kernel with two cost parameters, using some DCP points as training data. It was confirmed that the constructed SVM model was able to identify uniform DCP validation data with a high accuracy, UA = 0.954.

As a result of judging the DCP validation data as uniform or non-uniform using this classifier, Forest, Grass/Shrub, Cropland, and Urban were screened as non-uniform in more than half of the cases. This means that the typical spatial scale of these classes is smaller than 300 m. Therefore, it is indicated that the data used in these classes did not ensure spatial representativeness, and thus may have been underestimated when accuracy was assessed at larger than 300 m scale in previous studies. On the other hand, most of the DCP validation data for Barren and Water were judged to be uniform at 300 m and 990 m scales. Therefore, for Barren and Water, accuracy assessment by DCP is spatially representative if the resolution is 990 m or higher. For example, the results of the accuracy assessment at 1 km are reasonable from the viewpoint of spatial representativeness even if the validation data do not guarantee spatial representativeness so far. Due to the nature of the Cropland class, it is relatively difficult to determine uniformity and non-uniformity from satellite images, and the amount of data is extremely reduced by screening.

Author Contributions: Conceptualization, T.K., K.I. and Y.I.; methodology, T.K., K.I. and Y.I.; formal analysis, Y.I. and T.K.; investigation, Y.I.; data curation, Y.I.; writing—original draft preparation, Y.I.; writing—review and editing, T.K. and K.I. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Japan Society for the Promotion of Science (JSPS) KAKENHI, Grant Number 19K06307 and JSPS KAKENHI Grant Number JP 21J15348.

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

References

- Strahler, A.; Boschetti, L.; Foody, G.; Friedl, M.; Hansen, M.; Herold, M.; Mayaux, P.; Morisette, J.; Stehman, S.; Woodcock, C. Global Land Cover Validation: Recommendations for Evaluation and Accuracy Assessment of Global Land Cover Maps; Office for Official Publications of the European Communities: Luxembourg, 2006; pp. 1–60.
- Fonte, C.C.; Bastin, L.; See, L.; Foody, G.; Lupia, F. Usability of VGI for validation of land cover maps. Int. J. Geogr. Inf. Sci. 2015, 29, 1269–1291. [CrossRef]
- 3. Iwao, K.; Nishida, K.; Kinoshita, T.; Yamagata, Y. Validating land cover maps with Degree Confluence Project information. *Geophys. Res. Lett.* **2006**, *33*, 1–5. [CrossRef]
- 4. Joint Research Centre. Global Land Cover 2000—Products. Available online: https://bit.ly/2Vh1WdF (accessed on 29 June 2021).
- 5. NASA. Status for: Land Cover/Dynamics (MCD12). Available online: https://go.nasa.gov/3iQ9IZQ (accessed on 2 June 2021).
- 6. Hansen, M.C.; Defries, R.S.; Townshend, J.R.G.; Sohlberg, R. Global Land Cover Classification at 1 km Spatial Resolution Using a Classification Tree Approach. *Int. J. Remote Sens.* **2010**, *21*, 1331–1364. [CrossRef]
- GLCC; USGS. Global Land Cover Characteristics Data Base Version 2.0. Available online: https://on.doi.gov/3rE9gfv (accessed on 2 June 2021).

- 8. Iwao, K.; Nasahara, K.N.; Kinoshita, T.; Yamagata, Y.; Patton, D.; Tsuchida, S. Creation of New Global Land Cover Map with Map Integration. J. Geogr. Inf. Syst. 2011, 3, 160–165. [CrossRef]
- 9. Kinoshita, T.; Iwao, K.; Yamagata, Y. Creation of a global land cover and a probability map through a new map integration method. *Int. J. Appl. Earth Obs. Geoinf.* 2014, 28, 70–77. [CrossRef]
- 10. Foody, G.M.; Boyd, D.S. Using volunteered data in land cover map validation: Mapping west African forests. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 1305–1312. [CrossRef]
- Soyama, N.; Sasai, T.; Muramatsu, K.; Daigo, M.; Ochiai, F.; Nasahara, K. Production Method of Reference Data Focusing on Quality Assessment for Global Land Cover Map Validation using the Degree Confluence Project Information. *Trans. Jpn. Soc. Aeronaut. Sp. Sci. Aerosp. Technol. Jpn.* 2016, 14, Pn_1–Pn_6. [CrossRef]
- 12. Qian, T.; Kinoshita, T.; Fujii, M.; Bao, Y. Analyzing the uncertainty of degree confluence project for validating global land-cover maps using reference data-based classification schemes. *Remote Sens.* **2020**, *12*, 2589. [CrossRef]
- 13. NASA. LP DAAC ASTER Level 1 Precision Terrain Corrected Registered At-Sensor Radiance V003. Available online: https://doi.org/10.5067/ASTER/AST_L1T.003 (accessed on 30 June 2021).
- 14. U.S. Geological Survey (USGS); National Aeronautics and Space Administration (NASA). USGS EROS Archive—Landsat Legacy—Global Land Survey (GLS). Available online: https://www.usgs.gov/centers/eros/science/usgs-eros-archive-landsat-legacy-global-land-survey-gls (accessed on 30 June 2021).
- 15. Shimada, M.; Itoh, T.; Motooka, T.; Watanabe, M.; Shiraishi, T.; Thapa, R.; Lucas, R. New global forest/non-forest maps from ALOS PALSAR data (2007–2010). *Remote Sens. Environ.* **2014**, 155, 13–31. [CrossRef]
- Sulla-Menashe, D.; Friedl, M.A. User Guide to Collection 6 MODIS Land Cover (MCD12Q1 and MCD12C1) Product. Available online: https://bit.ly/2TDjLmZ (accessed on 2 June 2021).
- 17. Dora, L.; Agrawal, S.; Panda, R.; Abraham, A. Nested cross-validation based adaptive sparse representation algorithm and its application to pathological brain classification. *Expert Syst. Appl.* **2018**, *114*, 313–321. [CrossRef]
- 18. Marcot, B.G.; Hanea, A.M. What is an optimal value of k in k-fold cross-validation in discrete Bayesian network analysis? *Comput. Stat.* **2020**, 1–23. [CrossRef]
- Friedl, M.; Sulla-Menashe, D. MCD12Q1 MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 500 m SIN Grid V006. Available online: https://doi.org/10.5067/MODIS/MCD12Q1.006 (accessed on 30 June 2021).
- 20. The Global Land Cover by National Mapping Organizations. Available online: https://globalmaps.github.io/glcnmo.html (accessed on 29 June 2021).
- 21. MEDIAS-France/POSTEL. ESA and the ESA Globcover Project. Available online: http://due.esrin.esa.int/page_globcover.php (accessed on 29 June 2021).
- 22. Hiroshi, S.; Ryutaro, T. Global Land Use, Land Cover and Vegetation Classification Systems: A Review. GSI J. 2001, 96, 69–99.
- 23. Hua, T.; Zhao, W.; Liu, Y.; Wang, S.; Yang, S. Spatial consistency assessments for global land-cover datasets: A comparison among GLC2000, CCI LC, MCD12, GLOBCOVER and GLCNMO. *Remote Sens.* **2018**, *10*, 1846. [CrossRef]
- 24. Fritz, S.; Bartholome, E.; Belward, A.; Hartley, A.J.; Stibig, H.J.; Eva, H.; Mayaux, P.; Bartalev, S.; Latifovic, R.; Kolmert, S.; et al. The Global Land Cover for the Year 2000. Available online: https://bit.ly/3x1IT4o (accessed on 2 June 2021).
- Bicheron, P.; Defourny, P.; Brockmann, C.; Schouten, L.; Vancutsem, C.; Huc, M.; Bontemps, S.; Leroy, M.; Achard, F.; Herold, M.; et al. GLOBCOVER Description and Validation Report. Available online: https://bit.ly/2USSaPd (accessed on 2 June 2021).