

Article

Accuracy Assessment of GlobeLand30 2010 Land Cover over China Based on Geographically and Categorically Stratified Validation Sample Data

Yu Wang ¹ , Jingxiong Zhang ^{1,2,3,*}, Di Liu ², Wenjing Yang ¹ and Wangle Zhang ¹ 

¹ School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China; wangyuchn@whu.edu.cn (Y.W.); yangwj@whu.edu.cn (W.Y.); zhangwl@whu.edu.cn (W.Z.)

² School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China; alliu0815@whu.edu.cn

³ Collaborative Innovation Center of Geospatial Technology, Wuhan 430079, China

* Correspondence: jxzhang@whu.edu.cn

Received: 27 June 2018; Accepted: 31 July 2018; Published: 2 August 2018



Abstract: Land cover information is vital for research and applications concerning natural resources and environmental modeling. Accuracy assessment is an important dimension in use and production of land cover information. GlobeLand30 is a relatively new global land cover information product with a fine spatial resolution of 30 m and is potentially useful for many applications. This paper describes the methods for and results from the first country-wide and statistically based accuracy assessment of GlobeLand30 2010 land cover dataset over China. For this, a total of 8400 validation sample pixels were collected based on a sampling design featuring two levels of stratification (ten geographical regions, each with nine or eight land-cover classes). Validation sample data with reference class labels were acquired from visual interpretation based on Google Earth high-resolution satellite images. Error matrices for individual regions and entire China were estimated properly based on the sampling design adopted, with the former aggregated to get the latter through suitable weighting. Results were obtained, with agreement at a sample pixel defined both as a match between the map (class) label and either the primary or alternate reference label therein and, more strictly, as a match between the map label and the primary reference label only. Based on the former definition of agreement, the overall accuracy of GlobeLand30 2010 land cover for China was assessed to be 84.2%. User's accuracy and producer's accuracy were both greater than 80% for cultivated land, forest, permanent snow and ice, and bareland, with user's accuracy for water bodies estimated 94.2% (82.1% for wetland, 79.8% for artificial surface) and producer's accuracy for grassland estimated 89.0%. These indicate that GlobeLand30 2010 depicts land cover circa 2010 in China quite accurately, although estimates of accuracy indicators based on the latter definition of agreement were lower as expected with an estimated national overall accuracy of 81.0%. Regional and class variations in accuracy were revealed and examined in the light of their associations with land cover distributions and patterns. Implications for use and production of GlobeLand30 land cover information were discussed, so were commonality and lack of it between GlobeLand30 and other fine-resolution land cover products.

Keywords: GlobeLand30; accuracy assessment; stratified sampling; Google Earth; land cover; China

1. Introduction

As Earth observation technologies have evolved, huge amount of remote-sensing image data covering the whole globe has been acquired. Multi-scale land cover datasets are widely used in the Earth surface processes research, ecosystem assessments, environmental modeling,

and sustainable development planning [1–3]. Several global land cover products have been developed by national and international organizations during the past decades. These include, for example, the International Geosphere-Biosphere Program Data and Information System's land cover (IGBP DISCover) data product (1 km spatial resolution) [4], the University of Maryland (UMD) land cover maps (1 km resolution) [5], the Global Land Cover 2000 (GLC2000) maps from the European Commission's Joint Research Center (JRC) (1 km resolution) [6], the Moderate Resolution Imaging Spectroradiometer (MODIS) land cover maps (500/1000 m resolution) [7], the GlobCover land cover maps from European Space Agency (ESA), 2005–2006/2009 (300 m resolution) [8,9], the Climate Change Initiative Land Cover maps (CCI-LC) from ESA, 2000/2005/2010 (300m resolution) [10], and the Global Land Cover by National Mapping Organizations (GLCNMO) maps, 2003/2008/2013 (500/1000 m resolution) [11]. Recently, GlobeLand30 land cover datasets have been developed by the National Geomatics Center of China, based on integration of pixel- and object-based classification followed by knowledge-based interactive verification of classification results (shortly known as the POK-based approach) [12]. This product has been available worldwide since September 2014 [12]. In comparison with global land cover products previously developed, GlobeLand30 has the finest spatial resolution of 30 m and thus can provide information, in greater details, about land cover status and dynamics.

The importance of accuracy assessment is increasingly recognized. Accuracy assessment helps map users to evaluate the utility of land cover maps for their intended applications [13]. It also informs map productions so that map quality may be improved through adopting more robust classifiers, exploring more informative class-discriminant features, incorporating extra image data or ancillary data, or a combination of these.

To verify GlobeLand30's accuracy globally, a preliminary assessment was conducted based on a two-rank sampling strategy. In first-rank (first-stage) sampling, map sheets were selected globally, while in second-rank (second-stage) sampling sample data for each land cover type within each of the selected map sheets were collected [14]. Eighty sample map sheets were selected from a total of 847 map sheets in the first-rank sampling. A total number of 159,874 sample pixels were selected for the assessment of GlobeLand30 2010. The overall accuracy of the product was estimated about $80.3 \pm 0.2\%$ [15].

Regional and country-wide accuracy assessments for GlobeLand30 have also been carried out by researchers from many countries. Overall accuracies were estimated about 86% for Thessaly, Greece [16], 87% for Siberia [17], 90% for Kyiv Oblast, Ukraine [18], 78% for Iran [19], 59% for Kenya [20], 46% for Central Asia [21], 80% for Nepal [22], 92% for land surface water in Greece [23], and 80% for cropland in three continents with high risk of food insecurity [24]. Validation of GlobeLand30 was done also through comparisons with existing land cover products in Germany [25], Italy [26], Portugal [27] and East Africa [28], with rates of agreement ranging from 74% to 93%.

However, there have been few accuracy assessment efforts for GlobeLand30 in China. Below, we reviewed relevant results from country-wide, thematic (e.g., cropland), and regional assessments, respectively. To compare and assess seven global land cover datasets (including GlobeLand30) in China, Yang et al. [29] manually collected five sets of multi-scale validation sample units (VSUs), with spatial resolution ranging from $600 \text{ m} \times 600 \text{ m}$ to $2 \text{ km} \times 2 \text{ km}$. It was found that GlobeLand30 2010 was the most accurate datasets examined, with an overall accuracy of 82.4% reported based on 1063 validation sample units (with 600 m resolution, implying coarsening for GlobeLand30 and exclusion of sample units at heterogeneous locations), while the overall accuracies of the other datasets ranged from 33.9% to 67.2% [29].

Lu et al. [30] compared five global cropland datasets in China for the year 2010. Cropland census data at the provincial and regional scales were used to evaluate the cropland areas derived from these datasets, while 5704 validation sample units (including reference data at 2130 test sample units originally acquired for validating the Finer Resolution Observation and Monitoring of Global Land Cover (FROM-GLC) dataset by Gong et al. [31]) were utilized to verify locations of cropland parcels.

The results showed that GlobeLand30 datasets are the most accurate for cropland area estimation, with an overall accuracy of 79.6%.

Regional accuracy assessments of GlobeLand30 were also undertaken in China. For example, accuracy assessments of GlobeLand30 2010 in Shaanxi Province and Henan Province indicated overall accuracies of 80.0% and 81.5%, respectively [32,33].

Clearly, validation results in [30] and [32,33] need to be extended in terms of thematic and geographic coverages, respectively. For the country-wide assessment reported by Yang et al. [29], sampling was not probability-based, as admitted. Besides, the assessment was not done in its nominal (spatial) resolution of 30 m but in a coarsened resolution of 600 m. This country-wide assessment would be usefully strengthened with respect to statistical soundness and resolution refinement to enable inter-comparisons between classification accuracies of products with comparable resolution.

We should look for theoretical guidance and good practice examples regarding accuracy assessment from the literature. For example, the National Land Cover Database (NLCD) series products (also of 30 m resolution), which were developed by the MultiResolution Land Characteristics (MRLC) Consortium (www.mrlc.gov), provide consistent land cover information in the United States from decadal Landsat satellite imagery and other supplementary datasets [34]. Accuracy assessments were carried out for NLCD land cover following statistical framework and procedures well established as in [35–39]. To facilitate sustainable research and developments concerning use and production of GlobeLand30 datasets globally and in China, in particular, country-wide assessments of GlobeLand30 over China should be pursued based on well established procedures of sampling design, response design, and analyses. This is what this paper seeks to contribute to.

The remainder of the paper is organized as follows. Section 2 describes the methods for accuracy assessment of GlobeLand30 land cover in China. Section 3 reports accuracy assessment results obtained for different regions and for entire China. In Section 4, the main results are discussed, with patterns of misclassifications summarized, followed by conclusion in Section 5.

2. Methods

In GlobeLand30 maps, ten land-cover classes (as defined in Table 1) are depicted across the world at a nominal pixel resolution of 30 m, in accordance with a hierarchical classification method featuring pixel classification, object-based post-processing, and knowledge-based interactive verification [15]. In this paper, the accuracy assessment of GlobeLand30 2010 was performed by following the guidance suggested by leading scientists in the field, such as Stehman et al. [40] and Olofsson et al. [41]. According to these authors, three major components of the accuracy assessment are sampling design, response design, and analysis, which are described in this section.

In this paper, the spatial units for sampling and analysis concerning accuracy assessment are pixels. Although there is debate over the most appropriate sample unit for accuracy assessment, Stehman et al. [42] suggested that using pixels as sample units could avoid many of the complications that might arise when other spatial support units (e.g., blocks of pixels and polygons) were used. Use of pixels as units for sampling and analysis is a well-established practice in the validation of NLCD land cover information [35–39].

2.1. Sampling Design

It may not be feasible to compare GlobeLand30 maps with reference data of complete coverages even for a county let alone whole China (with 34 provincial administrative regions) pixel by pixel. Sampling design determines both the cost and statistical rigor of accuracy assessment [43]. By sampling design, we select certain number of locations at which reference data will be collected. In this paper, two-level stratified random sampling is adopted. The first level of stratification is based on partitioning Chinese administrative regions (including the regions of Hong Kong, Macao, and Taiwan, but with small islands and the seas excluded) into 10 geographical regions (Figure 1). The regional stratification helps to reduce standard errors in accuracy estimates, facilitate regional reporting of

accuracy, and provide an indication of how accuracy varies spatially across China. The second level of stratification is by GlobeLand30 map classes (Table 1 and Figure 1).

Table 1. Classification, codes, and definition of each land cover type of GlobeLand30.

Code	Type	Definition
10	Cultivated land	Land used for agriculture, horticulture and gardens, including paddy fields, irrigated and dry farmland, vegetable and fruit gardens, etc.
20	Forest	Land covered by trees, vegetation covers over 30%, including deciduous and coniferous forests, and sparse woodland with cover 10–30%, etc.
30	Grassland	Land covered by natural grass with cover over 10%, etc.
40	Shrubland	Land covered by shrubs with cover over 30%, including deciduous and evergreen shrubs, and desert steppe with cover over 10%, etc.
50	Wetland	Land covered by wetland plants and water bodies, including inland marsh, lake marsh, river floodplain wetland, forest/shrub wetland, peat bogs, mangrove and salt marsh, etc.
60	Water bodies	Water bodies in land area, including river, lake, reservoir, fish pond, etc.
70	Tundra	Land covered by lichen, moss, hardy perennial herb and shrubs in the polar regions, including shrub tundra, herbaceous tundra, wet tundra, and barren tundra, etc.
80	Artificial Surfaces	Land modified by human activities, including all kinds of habitation, industrial and mining area, transportation facilities, and interior urban green zones and water bodies, etc.
90	Bareland	Land with vegetation cover lower than 10%, including desert, sandy fields, Gobi, bare rocks, saline and alkaline land, etc.
100	Permanent snow and ice	Lands covered by permanent snow, glacier and icecap.

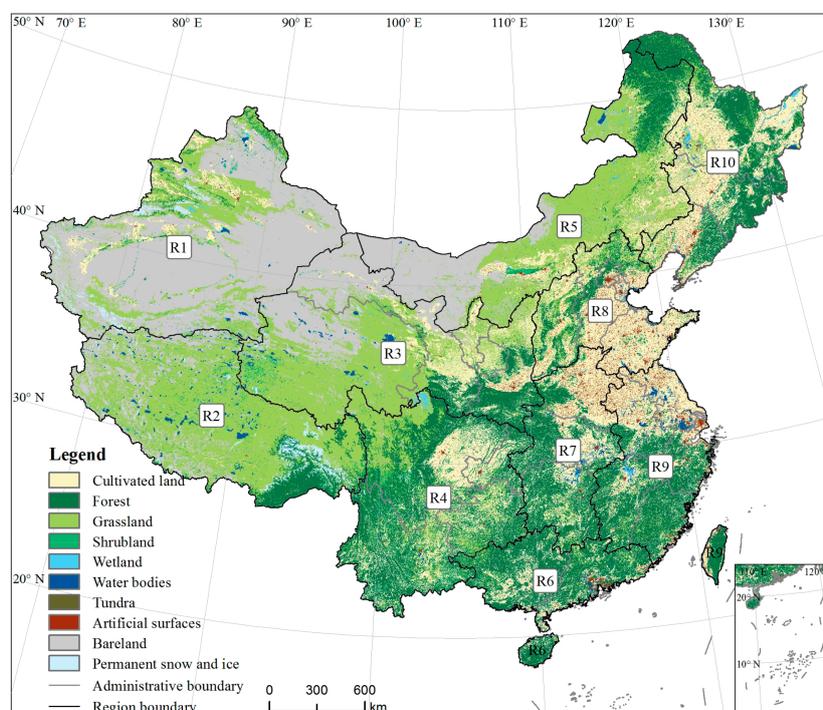


Figure 1. Regional stratification for GlobeLand30 2010 accuracy assessment over China. The boundaries of 10 geographic strata are shown in black. The labels “R1–R10” identify the regions used to geographically stratify sample data (e.g., R1 = Region 1).

In terms of sample size, each class in each region is allocated 100 sample pixels. The statistical explanation is that 100 sample pixels per stratum results in an expected standard error of 0.05 for simple random sampling within a stratum, if the true user's accuracy is 50% with a confidence level of 95% [35]. It should be noted that the tundra area in China is almost zero and it was not included in the validation. In addition, permanent snow and ice is only distributed in R1–R4. Thus, in total, 8400 sample pixels were collected for the accuracy assessment, as shown in Table 2.

Table 2. The regional distributions of the mapped land-cover type (percent of area) for GlobeLand30 2010 over China. The sample size is 100 for all land-cover types in all regions with the exception that permanent snow and ice is not sampled in R5–R10.

Type	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10
Cultivated land	5.42	0.36	13.09	28.51	12.01	27.19	45.01	57.69	39.97	43.71
Forest	1.15	8.45	10.46	45.65	11.55	61.22	42.27	16.81	44.63	37.62
Grassland	21.95	64.28	45.96	20.82	47.75	5.79	4.95	15.52	4.42	12.03
Shrubland	0.83	1.95	1.06	2.87	0.83	0.95	0.06	0.07	0.11	0.00
Wetland	0.29	0.18	0.40	0.39	0.52	0.02	0.33	0.31	0.45	1.36
Water bodies	0.64	2.59	1.34	0.60	0.46	2.32	2.79	1.73	4.18	1.51
Artificial Surfaces	0.29	0.03	0.55	0.59	0.77	2.49	4.56	7.81	6.18	3.09
Bareland	67.0	17.75	26.61	0.33	26.17	0.01	0.02	0.06	0.07	0.67
Permanent snow and ice	2.43	4.4	0.50	0.24	0.0	0.0	0.0	0.0	0.0	0.0
Country-wide proportion	17.2	12.68	14.56	11.89	12.08	4.71	5.94	5.55	7.08	8.31

2.2. Response Design

Reference land-cover labels for sample pixels are usually obtained from images of finer resolution. Online access to reference images is greatly enhanced thanks to Google Earth. GlobeLand30 2010 datasets for China are available in raster format with 54 tiles; the datasets are provided in WGS84 (World Geodetic System 1984) reference system and UTM (Universal Transverse Mercator) projection. Before selecting sample locations, the datasets need to be re-projected into the Web Mercator projection used for Google Earth. Other available datasets (e.g., Bing maps, Yahoo maps and color composites of Landsat images) were used when Google historical images of good quality were not available.

Three experienced interpreters carried out reference data acquisition. All reference data for each of individual regions were collected by a single interpreter. Interpreters had no a priori knowledge of map land-cover labels to avoid interpreter bias [36]. In addition to the primary reference land-cover labels, alternate land-cover labels should also be recorded for sample locations which it would be more appropriately labeled with both reference labels than any one alone. A nominal level of confidence was assigned to reference land-cover labels, namely “confident”, “somewhat confident”, and “not confident”, as in Wickham et al. [36]. The acquisition dates of Google historical images (used for determining the reference land-cover classes) and satellite images (used for map production, acquired by Landsat satellites and the Chinese Environmental and Disaster satellite) were also recorded, since the satellite images used for GlobeLand30 2010 were not restricted to the year 2010 alone.

After the first round of visual image interpretation, verification of “somewhat confident” and “not confident” reference class labels was performed by the project manager, who was typically the most experienced photo-interpreter in the team, to minimize errors in reference data. Consistency in reference label assignments within the team is also important, but precedence is given to the project manager's assignments when label assignments disagree. Especially, great attention should be paid to “not confident” labels, which are allowed to be re-assigned to “somewhat confident” or “confident” labels. In this step, ancillary data, such as DEM data, local photos, and Landsat images could be used to assist image interpretation.

With consideration for positional uncertainty and thematic ambiguity between land-cover map and reference data, agreement is registered for a sample pixel if the map class matches either the primary or alternate reference label therein [39]. To examine the effects of surrounding pixels on

accuracy and to increase the reliability of interpretation, reference land-cover labels in a 3×3 -pixel neighborhood centered on each sample pixel were also collected.

2.3. Analysis

By cross-tabulation between map and reference classes at sample pixels, error matrix can be constructed, allowing for estimating accuracy indicators, such as overall accuracy, user’s accuracy, and producer’s accuracy [44]. An example error matrix, as shown in Table 3, is constructed from sample counts, where an element n_{ij} represents the number of sample pixels classified as map category i ($i = 1, 2, \dots, k$; k being the total number of candidate classes) but actually verified to be reference category j ($j = 1, 2, \dots, k$). In Table 3 (and Table 4 below), the rows indicate map classes while the columns represent reference classes.

Table 3. Error matrix of sample counts, n_{ij} .

Class	1	2	...	k	Total
1	n_{11}	n_{12}	...	n_{1k}	n_1
2	n_{21}	n_{22}	...	n_{2k}	n_2
...
k	n_{k1}	n_{k2}	...	n_{kk}	n_k
Total	$n_{.1}$	$n_{.2}$...	$n_{.k}$	n

Olofsson et al. [45] described a more informative presentation of the error matrix. It is in terms of unbiased estimators of area proportions in cells (i, j) corresponding to map-reference class label pairs i and j :

$$\hat{P}_{ij} = W_i \frac{n_{ij}}{n_i} \tag{1}$$

where W_i represents the proportion of the area mapped as category i (in the region under study) and is calculated as $W_i = A_{m,i} \div A_{tot}$. A_{tot} is the total area of the region under study and $A_{m,i}$ (subscript m denotes “mapped”) is the mapped area of category i . An example error matrix populated with estimated area proportions is shown in Table 4, from which accuracy measures can be computed (see Equations (2)-(4)).

Table 4. Error matrix of estimated area proportions, \hat{P}_{ij} (Equation (1)).

Class	1	2	...	k	Total
1	\hat{P}_{11}	\hat{P}_{12}	...	\hat{P}_{1k}	\hat{P}_1
2	\hat{P}_{21}	\hat{P}_{22}	...	\hat{P}_{2k}	\hat{P}_2
...
k	\hat{P}_{k1}	\hat{P}_{k2}	...	\hat{P}_{kk}	\hat{P}_k
Total	$\hat{P}_{.1}$	$\hat{P}_{.2}$...	$\hat{P}_{.k}$	1

Estimating the proportion of area in each cell of the error matrix using Equation (1) takes into account the inclusion probabilities of the stratified design. An inclusion probability is defined as the probability that a particular pixel is included in the sample. Unlike in simple random and systematic sampling where inclusion probability of each selected pixel is the same (so that accuracy measures may be computed directly based on Table 3), in stratified random sampling, sample units from different strata usually have different weights, as sample units from different strata likely have different inclusion probabilities. Thus, area proportions of the map classes (W_i) must be incorporated in the stratified estimators of overall and producer’s accuracies to account for different sampling intensities in different strata.

Once the area proportions are estimated (Equation (1)) as in Table 4, user's accuracy (\hat{U}_i) and producer's accuracy (\hat{P}_j) for any category and overall map accuracy (\hat{O}) are estimated directly from area proportions (Table 4). The estimators are:

$$\hat{U}_i = \frac{\hat{P}_{ii}}{\hat{P}_{i.}} \quad (2)$$

$$\hat{P}_j = \frac{\hat{P}_{jj}}{\hat{P}_{.j}} \quad (3)$$

$$\hat{O} = \sum_{j=1}^k \hat{P}_{jj} \quad (4)$$

Variance estimators are also necessary for accuracy assessment, which could be applied to calculate confidence intervals [43]. For stratified random sampling, the estimated variance of overall accuracy is

$$V(\hat{O}) = \sum_{i=1}^k W_i^2 \hat{U}_i (1 - \hat{U}_i) / (n_{i.} - 1) \quad (5)$$

As for the estimated variance of user's and producer's accuracy, the method described by Wickham et al. [38] is usually applied. Further detail about variance estimation is provided in Appendix A.

Equations (1)–(5) are applied for estimating accuracies on the assumption that stratified random sample data are used. For the study in this paper, a two-level stratified sampling design was adopted (Section 2.1). Thus, Equations (1)–(5) are not directly applicable for estimating national accuracies, although they are perfectly suitable for estimating accuracies in individual regions where map-class-stratified random sampling was applied independently.

To evaluate nationwide accuracies, regional error matrices need to be aggregated to a national error matrix at first. This is done by summing up corresponding cell values at (i, j) (i.e., area proportions for map-reference class pair (i, j)) in regional error matrices via proper weights that are individual regions' areal proportions in the entire study area. Consider the study here as an example. The ten regional error matrices were estimated using Equations (1)–(4), as listed in Appendix B. The weights (W_h in Equation (6)) were the ten regions' areal proportion relative to the whole study area (Table 2, the last row, Row "Country-wide proportion"). In terms of a formula, the aggregated proportion for a particular class pair (i, j) ($\hat{P}_{i,j}$) is computed as a weighted sum of corresponding proportions in regional error matrices:

$$\hat{P}_{i,j} = \sum_{h=1}^H W_h \hat{P}_{i,j|h} \quad (6)$$

where H is the total number of regional strata in the study area ($H = 10$ for this paper), W_h is region h 's areal proportion in the whole study area ($W_h = N_h/N$, with N_h being the population size of region h , and N being the population size for the whole study area), and $\hat{P}_{i,j|h}$ represents cell (i, j) in region h 's error matrix. These properly calculated cell values ($\hat{P}_{i,j}$) constitute the national error matrix. Based on it, national accuracy indicators are computed using Equations (2)–(4) by inserting properly estimated $\hat{P}_{i,j}$ values (Equation (6)), as was implemented in this study.

Alternatively, country-wide accuracy estimates can be computed by following the method described in Wickham et al. [37,38]. It is the so-called combined ratio estimator based on suitably formulated indicator functions. This method is highly recommendable for computing accuracy estimates in situations where stratifications are more than one level as in this case study here and where strata are different from land cover classes for which accuracy indicators need to be estimated properly [37,46].

In Appendix A, we provide further detail about how national accuracy indicators are estimated based on aggregation of regional error matrices and the combined ratio estimator. We also show the equivalence of these two kinds of methods with explanations and using data from the study reported in this paper.

3. Results

In this section, we mainly report results of accuracy assessment based on the less strict definition of agreement (at a sample pixel) as a match between the map label and either the primary or alternate reference label. Assessment results obtained using the stricter definition of agreement (as a match between the map label and the primary reference label only) are reported briefly at the end section.

The ten regional error matrices were estimated (using the less strict definition of agreement unless stated otherwise in the remainder of this section), as shown in Appendix B (Tables A1–A10). Overall accuracies based on estimated population class area proportions (Equation (4)) are shown in Figure 2 (also shown in Table 5, the last row, Row “Overall”).

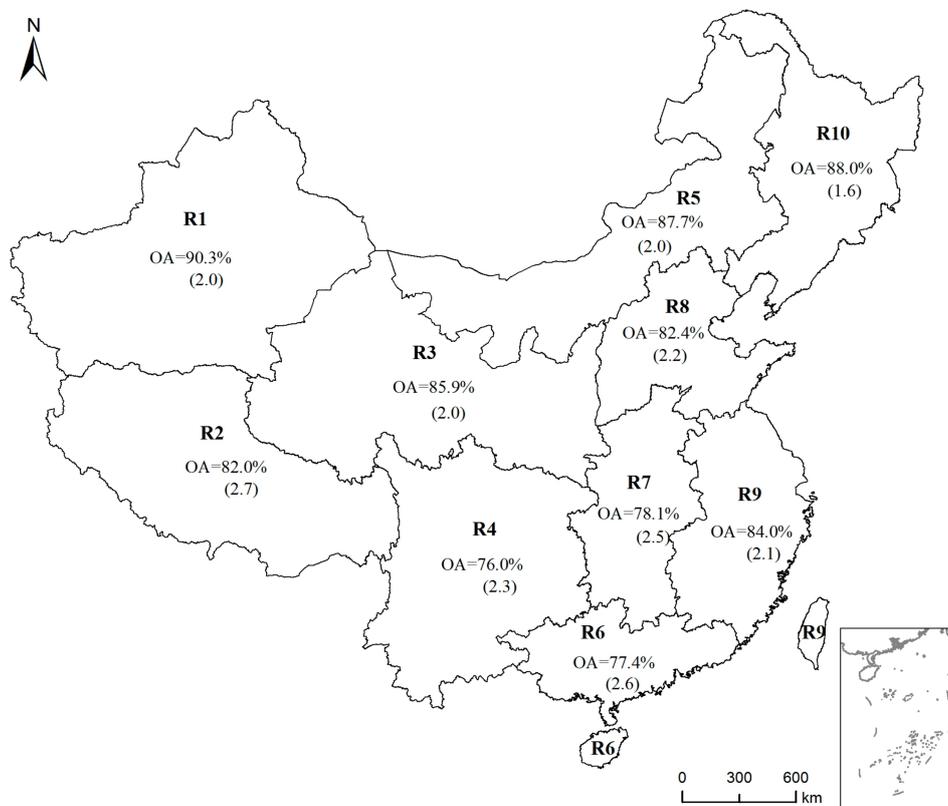


Figure 2. Regional overall accuracies for GlobeLand30 2010 over China based on the definition of agreement as a match between the map class and either the primary or alternate reference class. Standard errors for the overall accuracies are in parentheses.

The overall accuracies of ten regions range from 76.0% to 90.3%, while standard errors for estimated overall accuracy range between 1.6% and 2.6%, as shown in Figure 2. Most regions have overall accuracies over 80%, except for three regions (R4, R6 and R7). Overall accuracies in these three regions are about 8.6% lower than that of the other seven regions. Geographically, overall accuracies of northern regions (i.e., R1, R3, R5, and R10) rank the top four, with the three southern regions (R4, R6, and R7) ranking bottom three, while the rest (R2, R8, and R9) are in the middle. These differences are related to various factors, such as class composition, classification system, and landscape heterogeneity. In terms of land cover compositions, land cover classes are unequally distributed across China,

Tables 5 and 6 reveal an evident geographic pattern in classification errors related to class abundance. User's and producer's accuracies of grassland and bareland tend to decrease from west to east. These classes are abundant in the west (R1–R3 and R5) but generally rare in the east (R4 and R6–R10). Conversely, accuracy for forest tends to decrease from east to west, and this is correlated with the proportions of forest in the regions concerned. Permanent snow and ice of R1–R4 is well classified since it has unique spectral signature.

To evaluate the nationwide thematic accuracy, a country-wide error matrix of estimated area proportions was constructed by aggregating the ten regional error matrices using methods described in Section 2.3, with results shown in Table 7. At the country level, the overall accuracy for GlobeLand30 2010 over China is 84.2%, as shown in Table 7. This estimate of overall accuracy is slightly higher than the results reported in [29]. Below, we report results of accuracy evaluation obtained following the procedures above but using the stricter definition of agreement at a sample pixel (as a match between the map label and the primary reference label only). The results are shown in Appendix C. Specifically, regional overall accuracies and user's accuracies are shown in Table A11 (Appendix C), while regional producer's accuracies in Table A12. The national error matrix is shown in Table A13. Tables A11–A13 are similar in format to Tables 5–7, respectively, for convenience of comparisons between them.

Consider national accuracy assessment results. As expected, with the stricter definition of agreement, country-wide overall accuracy is estimated lower at 81.0% (Table A13), reduced by about 3% as opposed to the estimated overall accuracy of 84.2% (Table 7) based on the less strict definition of agreement. User's and producer's accuracies are decreased to differing extents depending on the specific classes concerned, although there are no differences between either user's accuracies or producer's accuracies for bareland and permanent snow and ice. The most obvious decreases in accuracies are observed for the class of artificial surface, which registers a user's accuracy of 70% (Table A13) (as opposed to 80%, Table 7) and a producer's accuracy of 53% (Table A13) (as opposed to 62%, Table 7). Further detail about decreases in accuracies (using the stricter definition of agreement as opposed to the less strict definition) and their variations in regional, categorical, and national terms is shown in Tables A11–A13 in Appendix C.

Table 7. The country-wide error matrix, cell entries are expressed as percent of area (see Table 5 for meanings of CuL, ArS, and PSI). Agreement is defined as a match between the map class and either the primary or alternate reference class. User's accuracy (UA) and producer's accuracy (PA) are reported with standard errors (SE) in parentheses. Overall accuracy is 84.2% (0.7%).

	CuL	Forest	Grassland	Shrubland	Wetland	Water	ArS	Bareland	PSI	Total	UA
CuL	18	1	0.5	0.6	0.2	0.2	0.7	0.1	0	21.3	84(1)
Forest	0.3	19.3	0.6	1.9	0	0.1	0	0.1	0	22.2	87(1)
Grassland	0.9	1	23.2	2.8	0.5	0.1	0.1	1.1	0	29.6	78(2)
Shrubland	0	0.1	0.2	0.7	0	0	0	0	0	1	70(2)
Wetland	0	0	0	0	0.4	0	0	0	0	0.4	82(2)
Water	0	0	0	0	0.1	1.5	0	0	0	1.6	94(1)
ArS	0.2	0.1	0	0	0	0	1.4	0	0	1.8	80(2)
Bareland	0	0	1.6	0.4	0.1	0	0	18.8	0	20.9	90(2)
PSI	0	0	0.1	0	0	0	0	0	0.9	1.1	88(2)
Total	19.5	21.5	26	6.5	1.2	1.9	2.3	20.2	1	100	
PA	93(1)	90(1)	89(1)	11(1)	30(5)	79(5)	62(4)	93(1)	93(4)		

4. Discussion

In this section, discussion is based on assessment results obtained with the less strict definition of agreement (as a match between the map label and either the primary or the alternate reference label at a sample pixel). It is, however, sensible to be aware of the implications of using a stricter definition vs. a less strict definition of agreement for accuracy assessment and to appreciate the relevance of the latter definition of agreement for accuracy validation in a landscape of complexity.

4.1. Patterns of Misclassification Errors

Error patterns for GlobeLand30 2010 over China are summarized below (in the remainder of the paper, without causing ambiguity, reference to China will not be made unless necessary): (1) Some dominant classes are overestimated where there are inclusions, such as the forest and cultivated land included in artificial surfaces, artificial surfaces and grassland in cultivated land, shrubland and grassland in forest, grassland and shrubland in bareland. (2) Grassland, shrubland and forest are difficult to distinguish as their spectral characteristics are often similar. These three classes are usually mixed in natural environment, making it “look like its surroundings”. (3) Fragmented patches, such as scattered villages and small blocks of cultivated land in hilly/mountainous areas, are likely to be omitted and classified as dominant classes in neighborhoods. (4) “Salt and pepper” noise is common in GlobeLand30 maps. GlobeLand30 has been prescribed minimum mapping unit (MMU) for each land-cover class and allowable minimum error of omission or commission per scene for each class. However, there is no restriction to small blocks with size smaller than specified MMU. There are still many small blocks of shrubland, grassland, bareland, and forest in GlobeLand30 maps. When sample pixels are located in these small blocks, they are likely to be classified as dominant classes in neighborhoods. (5) Time lags between map image acquisition dates and the dates for image interpretation also have effects on reported accuracies, because of the possibility of land-cover change. The dates of Landsat images used for GlobeLand30 2010 over China range from January to December (2010), while 89.4% of sample pixels were based on fine resolution images flown from May to November. Although Land-cover change within the year is relatively rare, misclassifications caused by time lags are observed (e.g., bareland in summer is misclassified as permanent snow and ice, cultivated land in non-growing season is misclassified as bareland). (6) Map heterogeneity has the expected effect of reducing reported agreements between map and reference classes. Land-cover heterogeneity is defined as the number of land-cover classes occurring in a 3×3 -pixel window centered on the sample pixel [47]. A heterogeneity value equal to one is defined as homogeneous (interior pixels), otherwise heterogeneous (edge pixels) in this paper. Nationwide, 72.6% of the sample pixels are interior, while the probability of disagreements for edge pixels is 2.3 times that of interior pixels.

4.2. Information for User Community and Product Improvement

Accuracy assessment is a standard component of GlobeLand30 mapping protocol. The overall accuracy for GlobeLand30 2010 dataset was estimated to be 84.2%, while user’s accuracies for individual classes (except for shrubland) exceeded 78% (Table 7). This indicates that the GlobeLand30 2010 dataset depicts spatial distributions of different land cover types in China with relatively great accuracy, given its fine resolution. Regional accuracy assessment results showed variations in accuracies due to regional differences in landscape patterns, as summarized at the end of Section 3.

Accuracy assessment can also inform and guide future map production. The results of GlobeLand30 2010 accuracy assessment show greater than 79% producer’s accuracies for all classes except for shrubland (only 11%), wetland, and artificial surface (Table 7). This suggests that mapping protocols need to be further developed and refined to better distinguish scattered grassland, shrubland, and small mixed grass-shrub patches, which are usually misclassified as other classes nearby. In China, especially southwestern China, spectrally mixed pixels are common in complicated and fragmented landscape. Although image segmentation and object-based classification techniques developed in GlobeLand30 data production can suppress “pepper-and-salt” effects in resultant classifications to some extent, determination of suitable segmentation parameters that are globally applicable is extremely difficult. Therefore, omission errors (of shrubland) are inevitable, especially in areas with a complex landscape. As shown in [48,49], scattered classes of small extents are likely misclassified as dominant classes in neighborhoods. This was similarly observed for wetland and artificial surface (e.g., human settlements in rural areas dominated by cultivated land). Further work is required to solve these problems in future global land cover mapping at fine spatial resolution. Regions and

classes prone to misclassifications (in particular, shrubland in all regions, wetland in R2, R6 and R7, and artificial surface in R2, R4 and R7; see Table 6) should be given special attention.

4.3. Comparison of GlobeLand30 with Other Related Data Products of 30 m Resolution

In this subsection, we compare GlobeLand30 with two fine-resolution land cover products, whose Level I thematic resolution (i.e., number of classes) is similar to that of GlobeLand30. One is a global product, known as FROM-GLC [31], the other is a US national product, NLCD [50], as mentioned in Section 1.

FROM-GLC contains two levels of land cover classes: 10 Level I classes (i.e., cropland, forest, grassland, shrubland, water bodies, impervious areas, bare lands, snow and ice, clouds and unclassified) and 29 Level II classes. It utilized thousands of Landsat images flown from 1981 to 2011 and was generated from fully automated image classification. The overall accuracies (Level I) range from 54% (maximum likelihood classifier) to 65% (support vector machine, SVM), while its best overall accuracy (Level II) is 53%. Yu et al. [51] improved FROM-GLC classification results using MODIS time series image and other auxiliary data, achieving an overall classification accuracy of 67%. In comparison, FROM-GLC is of lower accuracy than GlobeLand30 2010 as the former resulted from fully automated classification while the latter was enhanced with object-based classification and knowledge utilization.

As mentioned in Section 1, US NLCD time series land cover products are also of 30 m spatial resolution and based on Landsat TM images. NLCD 1992 is based primarily on unsupervised image classification [50], while NLCD 2001, NLCD 2006, and NLCD 2011 are based on more advanced decision-tree classifiers [52–54]. According to Wickham et al. [35–38] and Stehman et al. [39], overall accuracies for NLCD 1992, 2001, 2006, and 2011 land cover (Level I, 8 classes, see below) were 80% (unweighted averages of 10 US regional estimates), 85%, 84%, and 88%, respectively, while those for level II were 55% (unweighted averages of 10 US regional estimates), 79%, 78%, and 82%, respectively. In general, user's accuracies for water, forest, shrubland, agriculture, and high intensity developed exceeded 80% for most NLCD products. For NLCD 2011 (which is temporally close to GlobeLand30 2010), in particular, producer's accuracies for all 8 Level I classes (i.e., water, developed, barren, forest, shrubland, grassland, agriculture, and wetland) were estimated to be greater than 80%, while user's accuracies for all Level I classes (except for barren and wetland) were estimated to be greater than 81% [38]. Misclassifications usually occurred in regions with heterogeneous landscape, as also observed with GlobeLand30 2010.

Clearly, in terms of overall accuracy, GlobeLand30 2010 is slightly inferior to NLCD level I products (except for NLCD 1992 and NLCD 2006). We take the liberty in assuming approximate semantic equivalence between the following GlobeLand30 and NLCD classes: "cultivated land" = "agriculture", "bareland" = "barren", and "artificial surface" = "developed". As for user's accuracies, shrubland, grassland, and artificial surface are less accurately classified in GlobeLand30 than in NLCD 2011, while barren and wetland are less accurately classified in NLCD 2011 than in GlobeLand30. With respect to producer's accuracies, NLCD 2011 is apparently superior to GlobeLand30, as there were large omission errors for shrubland, wetland, and artificial surface in the latter.

However, it should be noted that comparisons with FROM-GLC and NLCD were made for Level I classes above. A total of 15 land cover classes will be mapped in GlobeLand30 2015 which is under development [55]. Class ambiguity tends to increase as thematic detail of classification increases (e.g., from 10 classes to 15 classes), and this increasing ambiguity may have a negative impact on accuracy. For instance, overall accuracies decrease from 84% to 78% for NLCD 2006, when thematic resolution increases from Level I (8 classes) to Level II (16 classes) [37]. Users and producers should be aware of the implications of increased thematic resolution for GlobeLand30 2015.

5. Conclusions

Accurate assessment in land cover information is of great importance to research and applications concerning natural resources and the environment. This paper provides detailed information about regional vs. national accuracies (overall, user's, and producer's) of GlobeLand30 2010 land cover over China by adopting a two-level stratified random sampling design and furnishing suitable methods for aggregating regional error matrices to a national one. The national overall accuracy for GlobeLand30 2010 was estimated as 84.2% (with agreement at a sample pixel defined as a match between the map label and either the primary or the alternate reference label), indicating GlobeLand30's relatively high accuracy. The national overall accuracy was estimated 81.0% when defining agreement at a sample pixel more strictly as a match between the map label and the primary reference label only. However, areas with heterogeneous landscapes and scattered small patches, in particular, need to be mapped with improved classification methods in future endeavors, as such areas tend to be labeled with low accuracies, as revealed in this study. GlobeLand30 production and validation teams will benefit from the research reported in this paper.

Author Contributions: Y.W. and J.Z. conceived and designed the experiments; Y.W., D.L. and W.Y. undertook data validation; W.Z. contributed to the data analysis; and Y.W. and J.Z. wrote the paper.

Funding: This research was funded by National Natural Science Foundation of China, grant number 41471375 and the APC was also funded by National Natural Science Foundation of China (No. 41471375).

Acknowledgments: Research reported in this paper was supported by the National Natural Science Foundation of China (grant No. 41471375). Constructive comments and suggestions from anonymous reviewers were received with thanks. Michael F. Goodchild (Emeritus Professor of Geography at the University of California, Santa Barbara) and Jun Chen (The National Geomatics Center of China (NGCC)) have provided long-term advice about spatial uncertainty for Jingxiong Zhang, the corresponding author.

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

Appendix A. Some Detail about the Methods for Estimating National Accuracy Indicators and Their Variance

In this section, error matrices are assumed to be populated with estimated area proportions unless stated otherwise. As described in Section 2.3, estimates of country-wide accuracy indicators are computed using Equations (2)–(4) after creating a nationally aggregated error matrix based on proper weighting using Equation (6). We may call this method a regional error matrix aggregation method. The combined ratio estimator method described in Wickham et al. [37] is also suitable for computing estimates of user's accuracies and producer's accuracies. Below, both methods are described with examples, with their equivalence explained and illustrated afterwards. Lastly, formulas for variance of estimated national accuracy indicators including user's accuracy (UA), producer's accuracy (PA), and overall accuracy (OA) are re-visited.

For the error-matrix-aggregation-based method, the key is to have proper weightings applied to proportions ($\hat{P}_{i,j|h}$) in regional error matrices so that values of re-weighted $\hat{P}_{i,j|h}$ become the ratio of the area mapped as category i in region h to the national total area (N) instead of region h 's total area (N_h). This can be seen to be the case for what Equation (6) sums:

$$\hat{P}_{ij|h-reweighted} = W_h \hat{P}_{ij|h} = W_h W_{i|h} \frac{n_{ij}}{n_i} = W_{i,h|H} \frac{n_{ij}}{n_i} \quad (A1)$$

where W_h is areal proportion of region h in the whole study area (N_h/N), $W_{i|h}$ is actually W_i in Equation (1) representing the proportion of the area mapped as category i in region h , n_{ij} and n_i are the same as in Equation (1) (assuming the error matrix of sample counts for region h is similar to Table 3), and $W_{i,h|H}$ is the combined weight representing the ratio of the area mapped as category i in region h over the whole study area (as indicated by H).

Thus, all proportions $\hat{P}_{i,j|h}$ (for map-reference class pairs (i, j)) in regional error matrices are re-weighted nationally as desired (Equation (A1)). These re-weighted regional error matrices were

aggregated with their corresponding proportions $\hat{P}_{ij|h-reweighted}$ summed up, giving rise to the national error matrix with correct (national) proportions $\hat{P}_{i,j}$ (Equation (6)).

For example, $\hat{P}_{2,2}$ for forest-forest class pair in the national error matrix (shown in Table 7) is computed as:

$$\hat{P}_{2,2} = (17.2 \times 0.8 + 12.68 \times 7.5 + 14.56 \times 8.7 + 11.89 \times 37.9 + 12.08 \times 11.1 + 4.71 \times 51.4 + 5.94 \times 35.9 + 5.55 \times 13.8 + 7.08 \times 40.2 + 8.31 \times 35.4) / (100 \times 100) = 0.193.$$

After creation of the national error matrix, UA, PA, and OA can be easily computed using Equations (2)–(4), respectively. For example, based on the national error matrix in Table 7, UA and PA for forest are:

$$\text{UA}(\text{forest}) = \hat{P}_{2,2} / \hat{P}_{2\cdot} = (19.3/100) / (22.2/100) = 0.868, \text{ and}$$

$$\text{PA}(\text{forest}) = \hat{P}_{\cdot 2} / \hat{P}_{2\cdot} = (19.3/100) / (21.5/100) = 0.899,$$

where $\hat{P}_{i\cdot}$ and $\hat{P}_{\cdot j}$ represent row i and column j 's totals in the national error matrix, respectively ($i = j = 2$ for forest). The estimate for OA is:

$$\text{OA} = (18 + 19.3 + 23.2 + 0.7 + 0.4 + 1.5 + 1.4 + 18.8 + 0.9) / 100 = 0.842.$$

By the combined ratio estimator method, on the other hand, UA and PA are estimated as a ratio $R = Y/X$, where Y is the population total of y_u and X is the population total of x_u (u being a pixel in the population). y_u and x_u are indicator functions for pixel u on condition A and condition B , respectively. For UA of a particular class, say “forest”, condition A is that the map and reference labels are both forest, while condition B is that map label is forest. For PA of forest, condition A remains the same, but condition B is that the reference label is forest, as also explained in [38].

The combined ratio estimator for UA or PA is:

$$\hat{R} = \frac{\hat{Y}}{\hat{X}} = \frac{\sum_{h=1}^H N_h \bar{y}_h}{\sum_{h=1}^H N_h \bar{x}_h} \quad (\text{A2})$$

where \bar{x}_h is the sample mean of x_u in stratum h , \bar{y}_h is the sample mean of y_u in stratum h , N_h is the population size in stratum h , and H is the number of strata in the study area [37]. This ratio estimator is very general as it can handle sample data with double stratifications (as in this paper) and situations where there is no one-to-one correspondence between strata and classes for which accuracy indicators need to be estimated.

In this study, the sample data were collected following a two-level stratification (ten regions, each with nine or eight land cover classes), as described in Section 2.1. We can treat the sample data as consisting of 84 strata and calculate \hat{R} with Equation (A2) ($H = 84$).

However, given the congruence among regional error matrices and one-one correspondence of strata and mapped classes in individual regions, simplified use of Equation (A2) is possible on the basis of regional error matrices. In other words, it is more sensible to view regions in the study as strata to work with when applying Equation (A2) (i.e., $H = 10$). Then, for a particular region h , N_h refers to the region's areal proportion (i.e., Table 2, bottom row). In addition, given regional error matrices, we can easily get sample statistics required in Equation (A2). Specifically, \bar{x}_h is class i 's sample proportion in the region (e.g., row or column i 's totals in the error matrix for region h , $\hat{P}_{i\cdot|h}$ and $\hat{P}_{\cdot i|h}$ depending on whether UA or PA is concerned), while \bar{y}_h is the proportion of sample pixels of reference class i classified correctly as class i in the region (e.g., $\hat{P}_{i,i|h}$ in the error matrix for region h).

For example, using regional error matrices (Tables A1–A10, Appendix B), UA and PA for forest are computed as:

$$\begin{aligned} \text{UA}(\text{forest}) &= \frac{\sum_{h=1}^{10} W_h \hat{P}_{2,2|h}}{\sum_{h=1}^{10} W_h \hat{P}_{\cdot|2|h}} = [(17.2 \times 0.8 + 12.68 \times 7.5 + 14.56 \times 8.7 \\ &+ 11.89 \times 37.9 + 12.08 \times 11.1 + 4.71 \times 51.4 + 5.94 \times 35.9 + 5.55 \times 13.8 + 7.08 \times 40.2 \\ &+ 8.31 \times 35.4) / (100 \times 100)] / [(17.2 \times 1.2 + 12.68 \times 8.5 + 14.56 \times 10.5 + 11.89 \times 45.7 \\ &+ 12.08 \times 11.5 + 4.71 \times 61.2 + 5.94 \times 42.3 + 5.55 \times 16.8 + 7.08 \times 44.6 \\ &+ 8.31 \times 37.6) / (100 \times 100)] = 0.868, \end{aligned}$$

$$\begin{aligned} \text{PA}(\text{forest}) &= \frac{\sum_{h=1}^{10} W_h \hat{P}_{2,2|h}}{\sum_{h=1}^{10} W_h \hat{P}_{2|\cdot|h}} = [(17.2 \times 0.8 + 12.68 \times 7.5 + 14.56 \times 8.7 \\ &+ 11.89 \times 37.9 + 12.08 \times 11.1 + 4.71 \times 51.4 + 5.94 \times 35.9 + 5.55 \times 13.8 + 7.08 \times 40.2 \\ &+ 8.31 \times 35.4) / (100 \times 100)] / [(17.2 \times 1.3 + 12.68 \times 8.5 + 14.56 \times 9.6 + 11.89 \times 42.7 \\ &+ 12.08 \times 11.6 + 4.71 \times 56.7 + 5.94 \times 38.8 + 5.55 \times 16.5 + 7.08 \times 44.4 \\ &+ 8.31 \times 39.1) / (100 \times 100)] = 0.899. \end{aligned}$$

when using Equation (A2) to compute UA and PA above, proportionality between W_h and N_h (non-zero ones) is implicitly applied.

For estimating national OA, we can apply Equation (A2) but only the numerator part, with summation over all map classes. Thus, based on regional error matrices (Tables A1–A10, Appendix B) again, we compute national OA as:

$$\begin{aligned} \text{OA} &= \frac{\sum_{h=1}^{10} W_h \sum_{i=1}^9 \hat{P}_{ii|h}}{\sum_{h=1}^{10} W_h \sum_{i=1}^9 \hat{P}_{i|i|h}} = [17.2 \times (4.9 + 0.8 + 19.3 + 0.5 + 0.3 + 0.6 + 0.2 + 61.6 + 2.1) \\ &+ 12.68 \times (0.2 + 7.5 + 52.1 + 1.3 + 0.2 + 2.5 + 14.4 + 3.8) + 14.56 \times (11.3 + 8.7 + 40 + 0.7 \\ &+ 0.4 + 1.3 + 0.4 + 22.6 + 0.5) + 11.89 \times (22.8 + 37.9 + 11.2 + 2.3 + 0.4 + 0.6 + 0.4 + 0.2 \\ &+ 0.2) + 12.08 \times (11.2 + 11.1 + 38.7 + 0.5 + 0.5 + 0.4 + 0.6 + 24.8) + 4.71 \times (19 + 51.4 \\ &+ 2 + 0.7 + 2 + 2.2) + 5.94 \times (34.2 + 35.9 + 2 + 0.3 + 2.6 + 3.1) + 5.55 \times (50.8 + 13.8 \\ &+ 9.6 + 0.3 + 1.6 + 6.3) + 7.08 \times (32.8 + 40.2 + 1.5 + 0.1 + 0.3 + 4 + 5.2) \\ &+ 8.31 \times (40.6 + 35.4 + 6.5 + 1 + 1.4 + 2.6 + 0.5)] / (100 \times 100) = 0.842. \end{aligned}$$

Clearly, estimates for national UA and PA computed from the two methods are identical. The equivalence between the two methods' results is not only for forest but for any class i , as it is established by:

$$\text{UA}(i) = \frac{\hat{P}_{i,i}}{\hat{P}_{\cdot|i}} = \frac{\sum_{h=1}^{10} W_h \hat{P}_{i,i|h}}{\sum_{h=1}^{10} W_h \hat{P}_{\cdot|i|h}} \quad (\text{A3})$$

$$\text{PA}(i) = \frac{\hat{P}_{i,i}}{\hat{P}_{i|\cdot}} = \frac{\sum_{h=1}^{10} W_h \hat{P}_{i,i|h}}{\sum_{h=1}^{10} W_h \hat{P}_{i|\cdot|h}} \quad (\text{A4})$$

where Equation (6) is applied for numerators and denominators, separately.

As for the estimated variance of UA and PA, the method described in [38] is applicable. Specifically, the estimated variance of the combined ratio estimator is computed as:

$$\hat{V}(\hat{R}) = \left(\frac{1}{\hat{X}^2} \right) \left[\sum_{h=1}^H N_h^2 \left(1 - \frac{n_h}{N_h} \right) (s_{y_h}^2 + \hat{R}^2 s_{x_h}^2 - 2\hat{R} s_{xy_h}) / n_h \right] \quad (\text{A5})$$

where n_h is sample size for stratum h (N_h being population size for stratum h , as previously in Equation (A2)), H is the number of strata (84 for the study in this paper), $s_{y_h}^2$ and $s_{x_h}^2$ are the sample variances of y_u and x_u for stratum h , and s_{xy_h} is the sample covariance of x_u and y_u in stratum h .

The estimated variance of national OA can be calculated using Equation (5). Adaptation is, however, required by viewing the country-wide population as consisting of 84 strata (region-class combinations), for which weights $W_{i, h|H}$ need to be calculated properly (see Equation (A1) and the

example for computing OA three lines above in Equation (A3)). In addition, we can compute variance of national OA using Equation (A6):

$$\hat{V}(\hat{O}) = \frac{1}{N^2} \sum_{h=1}^H N_h(N_h - n_h) \frac{s_h^2}{n_h} = \sum_{h=1}^H \frac{W_{i,h|H}^2 s_h^2}{n_h} - \sum_{h=1}^H \frac{W_{i,h|H}^2 s_h^2}{N} \tag{A6}$$

where s_h^2 is the sample variances of stratum h , N is the population size (total number of pixels) of the study area, and H is the number of strata to run the summation (84 for the study in this paper). We employed Equation (A6) for computing variance of estimated national OA, although we tested both methods (Equations (5) and (A6)) and obtained identical results.

Appendix B. Regional Error Matrices of Estimated Area Proportions when Defining Agreement at a Sample Pixel as a Match between the Map Label and Either the Primary or Alternate Reference Label

Tables A1–A10 show error matrices of estimated area proportions for GlobeLand30 2010 in all regions (R1–R10). In Tables A1–A10, cultivated land is abbreviated as CuL, artificial surfaces as ArS, and permanent snow and ice as PSI, – means there is no permanent snow and ice in a given region, as in Tables 5–7. User’s accuracy (UA) and producer’s accuracy (PA) are reported with standard errors (SE) in parentheses. These results are based on defining agreement as a match between the map label and the primary or alternate reference label at sample pixels.

Table A1. The error matrix of estimated area proportions for R1: Overall accuracy is 90.3% (2.0%).

	CuL	Forest	Grassland	Shrubland	Wetland	Water	ArS	Bareland	PSI	Total	UA
CuL	4.9	0	0.2	0.1	0	0	0.1	0.1	0	5.4	91(3)
Forest	0	0.8	0.1	0.2	0	0	0	0	0	1.2	68(5)
Grassland	0	0.4	19.3	1.3	0.2	0	0	0.4	0.2	21.9	88(3)
Shrubland	0	0	0.2	0.5	0	0	0	0.1	0	0.8	57(5)
Wetland	0	0	0	0	0.3	0	0	0	0	0.3	86(3)
Water	0	0	0	0	0	0.6	0	0	0	0.6	92(3)
ArS	0	0	0	0	0	0	0.2	0	0	0.3	76(4)
Bareland	0	0	4	1.3	0	0	0	61.6	0	67	92(3)
PSI	0	0	0.1	0	0	0	0	0.1	2.1	2.4	87(3)
Total	5	1.3	24.1	3.5	0.6	0.6	0.3	62.3	2.4	100	
PA	99(0)	63(16)	80(5)	14(4)	44(17)	100(0)	67(15)	99(1)	90(8)		

Table A2. The error matrix of estimated area proportions for R2: Overall accuracy is 82.0% (2.7%).

	CuL	Forest	Grassland	Shrubland	Wetland	Water	ArS	Bareland	PSI	Total	UA
CuL	0.2	0	0	0	0	0	0	0	0	0.4	64(5)
Forest	0	7.5	0.5	0.4	0	0	0	0	0	8.5	89(3)
Grassland	0	0.6	52.1	7.1	1.9	0	0	2.6	0	64.3	81(4)
Shrubland	0	0.1	0.4	1.3	0.1	0	0	0	0	2	68(5)
Wetland	0	0	0	0	0.2	0	0	0	0	0.2	85(4)
Water	0	0	0	0	0.1	2.5	0	0	0	2.6	97(2)
ArS	0	0	0	0	0	0	0	0	0	0	59(5)
Bareland	0	0.2	2.7	0.2	0.2	0	0	14.4	0.2	17.8	81(4)
PSI	0	0.1	0.3	0.1	0	0	0	0.1	3.8	4.4	87(3)
Total	0.2	8.5	55.9	9.2	2.4	2.5	0	17.1	4	100	
PA	99(0)	88(7)	93(1)	14(3)	6(3)	99(0)	46(11)	84(6)	95(4)		

Table A3. The error matrix of estimated area proportions for R3: Overall accuracy is 85.9% (2.0%).

	CuL	Forest	Grassland	Shrubland	Wetland	Water	ArS	Bareland	PSI	Total	UA
CuL	11.3	0.4	0.7	0.4	0	0	0.4	0	0	13.1	86(3)
Forest	0.1	8.7	0.6	1	0	0	0	0	0	10.5	83(4)
Grassland	0	0.5	40	2.8	0.9	0	0	1.8	0	46	87(3)
Shrubland	0	0	0.2	0.7	0	0	0	0	0	1.1	70(5)
Wetland	0	0	0	0	0.4	0	0	0	0	0.4	90(3)
Water	0	0	0	0	0	1.3	0	0	0	1.3	97(2)
ArS	0.1	0	0	0	0	0	0.4	0	0	0.6	78(4)
Bareland	0	0	2.9	0.5	0.5	0	0	22.6	0	26.6	85(4)
PSI	0	0	0	0	0	0	0	0	0.5	0.5	98(1)
Total	11.5	9.6	44.4	5.5	1.9	1.3	0.8	24.5	0.5	100	
PA	98(1)	91(5)	90(2)	14(3)	20(8)	100(0)	52(14)	92(3)	96(3)		

Table A4. The error matrix of estimated area proportions for R4: Overall accuracy is 76.0% (2.3%).

	CuL	Forest	Grassland	Shrubland	Wetland	Water	ArS	Bareland	PSI	Total	UA
CuL	22.8	2.6	0	2	0.3	0.3	0.6	0	0	28.5	80(4)
Forest	0.5	37.9	0.9	5.5	0	0	0	0.9	0	45.7	83(4)
Grassland	1.2	1.9	11.2	5	0.4	0	0.2	0.8	0	20.8	54(5)
Shrubland	0.1	0.3	0.1	2.3	0	0	0	0	0	2.9	80(4)
Wetland	0	0	0	0	0.4	0	0	0	0	0.4	90(3)
Water	0	0	0	0	0	0.6	0	0	0	0.6	95(2)
ArS	0.1	0	0	0	0	0	0.4	0	0	0.6	75(4)
Bareland	0	0	0	0	0	0	0	0.2	0	0.3	67(5)
PSI	0	0	0	0	0	0	0	0	0.2	0.2	81(4)
Total	24.7	42.7	12.4	14.8	1.1	0.9	1.3	2	0.2	100	
PA	92(3)	89(2)	91(5)	15(2)	33(13)	64(21)	35(13)	11(4)	92(3)		

Table A5. The error matrix of estimated area proportions for R5: Overall accuracy is 87.7% (2.0%).

	CuL	Forest	Grassland	Shrubland	Wetland	Water	ArS	Bareland	PSI	Total	UA
CuL	11.2	0	0.2	0.2	0	0	0.2	0.1	–	12	93(3)
Forest	0	11.1	0.1	0.3	0	0	0	0	–	11.5	96(2)
Grassland	2.9	0.5	38.7	2.4	0	0.5	0	2.9	–	47.8	81(4)
Shrubland	0.1	0	0.2	0.5	0	0	0	0	–	0.8	59(5)
Wetland	0	0	0	0	0.5	0	0	0	–	0.5	90(3)
Water	0	0	0	0	0	0.4	0	0	–	0.5	93(3)
ArS	0.1	0	0	0	0	0	0.6	0	–	0.8	80(4)
Bareland	0	0	0.8	0.5	0	0	0	24.8	–	26.1	95(2)
PSI	–	–	–	–	–	–	–	–	–	–	–
Total	14.2	11.6	40.1	4	0.5	0.9	0.9	27.9	–	100	
PA	79(6)	96(4)	97(1)	12(4)	97(2)	47(25)	72(14)	89(4)	–		

Table A6. The error matrix of estimated area proportions for R6: Overall accuracy is 77.4% (2.6%).

	CuL	Forest	Grassland	Shrubland	Wetland	Water	ArS	Bareland	PSI	Total	UA
CuL	19	3.8	0.5	1.4	0.8	0.5	0.5	0.5	–	27.2	70(5)
Forest	1.2	51.4	1.2	6.1	0	0.6	0.6	0	–	61.2	84(4)
Grassland	0.5	1.3	2	1.6	0	0	0.3	0	–	5.8	35(5)
Shrubland	0	0.1	0	0.7	0	0	0	0	–	1	78(4)
Wetland	0	0	0	0	0	0	0	0	–	0	67(5)
Water	0	0	0.1	0	0.1	2	0	0	–	2.3	87(3)
ArS	0.1	0.1	0	0.1	0	0	2.2	0	–	2.5	87(3)
Bareland	0	0	0	0	0	0	0	0	–	0	56(5)
PSI	–	–	–	–	–	–	–	–	–	–	–
Total	21	56.7	3.9	9.9	0.9	3.2	3.7	0.5	–	100	
PA	91(4)	91(2)	52(13)	7(2)	2(1)	63(14)	58(12)	1(1)	–		

Table A7. The error matrix of estimated area proportions for R7: Overall accuracy is 78.1% (2.5%).

	CuL	Forest	Grassland	Shrubland	Wetland	Water	ArS	Bareland	PSI	Total	UA
CuL	34.2	1.8	0	2.3	1.4	0.9	4.5	0	–	45	76(4)
Forest	1.3	35.9	1.7	3.4	0	0	0	0	–	42.3	85(4)
Grassland	0.7	0.8	2	1	0	0	0.2	0	–	5	41(5)
Shrubland	0	0	0	0	0	0	0	0	–	0.1	78(4)
Wetland	0	0	0	0	0.3	0.1	0	0	–	0.3	81(4)
Water	0.1	0	0	0	0.1	2.6	0	0	–	2.8	92(3)
ArS	1.2	0.2	0	0	0	0	3.1	0	–	4.6	67(5)
Bareland	0	0	0	0	0	0	0	0	–	0	54(5)
PSI	–	–	–	–	–	–	–	–	–	–	–
Total	37.6	38.8	3.7	6.7	1.7	3.6	7.8	0.1	–	100	
PA	91(2)	93(2)	54(13)	1(0)	15(7)	72(13)	39(7)	21(17)	–		

Table A8. The error matrix of estimated area proportions for R8: Overall accuracy is 82.4% (2.2%).

	CuL	Forest	Grassland	Shrubland	Wetland	Water	ArS	Bareland	PSI	Total	UA
CuL	50.8	1.7	2.3	0	0	0.6	1.7	0.6	–	57.7	88(3)
Forest	0.7	13.8	0.7	1.7	0	0	0	0	–	16.8	82(4)
Grassland	1.7	0.6	9.6	3.4	0	0	0.2	0	–	15.5	62(5)
Shrubland	0	0	0	0	0	0	0	0	–	0.1	67(5)
Wetland	0	0	0	0	0.3	0	0	0	–	0.3	85(4)
Water	0	0	0	0	0.1	1.6	0	0	–	1.7	92(3)
ArS	0.5	0.4	0.2	0.2	0.1	0	6.3	0	–	7.8	81(4)
Bareland	0	0	0	0	0	0	0	0	–	0.1	37(5)
PSI	–	–	–	–	–	–	–	–	–	–	–
Total	53.7	16.5	12.8	5.4	0.4	2.2	8.2	0.6	–	100	
PA	95(1)	83(5)	75(7)	1(0)	60(12)	73(19)	77(9)	4(4)	–		

Table A9. The error matrix of estimated area proportions for R9: Overall accuracy is 84.0% (2.1%).

	CuL	Forest	Grassland	Shrubland	Wetland	Water	ArS	Bareland	PSI	Total	UA
CuL	32.8	3.2	1.2	0.4	0	1.2	1.2	0	–	40	82(4)
Forest	0.4	40.2	0.4	3.1	0	0.4	0	0	–	44.6	90(3)
Grassland	0.8	0.8	1.5	0.7	0.1	0.2	0.2	0	–	4.4	35(5)
Shrubland	0	0	0	0.1	0	0	0	0	–	0.1	53(5)
Wetland	0	0	0	0	0.3	0.1	0	0	–	0.5	63(5)
Water	0.1	0	0	0	0.1	4	0	0	–	4.2	95(2)
ArS	0.7	0.2	0	0	0	0.1	5.2	0.1	–	6.2	84(4)
Bareland	0	0	0	0	0	0	0	0	–	0.1	42(5)
PSI	–	–	–	–	–	–	–	–	–	–	–
Total	34.9	44.4	3.2	4.3	0.5	6	6.6	0.1	–	100	
PA	94(1)	90(2)	48(13)	1(0)	56(11)	66(9)	78(8)	31(22)	–		

Table A10. The error matrix of estimated area proportions for R10: Overall accuracy is 88.0% (1.6%).

	CuL	Forest	Grassland	Shrubland	Wetland	Water	ArS	Bareland	PSI	Total	UA
CuL	40.6	0.4	0.9	0.9	0	0	0.9	0	–	43.7	93(3)
Forest	0	35.4	0.4	1.9	0	0	0	0	–	37.6	94(2)
Grassland	1.9	3.2	6.5	0	0	0	0.4	0	–	12	54(5)
Shrubland	0	0	0	0	0	0	0	0	–	0	69(5)
Wetland	0.3	0	0	0	1	0.1	0	0	–	1.4	73(4)
Water	0	0	0	0	0.1	1.4	0	0	–	1.5	94(2)
ArS	0.4	0.1	0	0	0	0	2.6	0	–	3.1	84(4)
Bareland	0	0	0.1	0	0	0	0	0.5	–	0.7	72(5)
PSI	–	–	–	–	–	–	–	–	–	–	–
Total	43.3	39.1	7.9	2.8	1.1	1.5	3.9	0.5	–	100	
PA	94(1)	90(2)	82(8)	0(0)	91(3)	93(3)	67(11)	100(0)	–		

Appendix C. Regional and National Accuracy Estimates Obtained when Defining Agreement as a Match between the Map Label and the Primary Reference Label Only

On the other hand, Tables A11 and A12 show regional user's and producer's accuracies for nine land cover classes with standard errors (SE) in parentheses when agreement is defined as a match of map class labels with the primary reference class label only. In Tables A11 and A12, Column "National" represents countrywide user's or producer's accuracies as in Tables 5 and 6, respectively. National accuracy indicators were computed based on the nationwide error matrix obtained by properly aggregating the ten regional error matrices. As shown in Table A13, country-wide overall accuracy is 81.0%, reduced by about 3.2% compared to that obtained with a more relaxed definition of agreement as a match of map labels with the primary or alternate reference label.

Table A11. Regional user's accuracies for nine land cover classes with standard errors (SE) in parentheses when agreement is defined as a match with the primary label.

	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	National
CuL	86(3)	58(5)	77(4)	69(5)	90(3)	63(5)	70(5)	87(3)	72(5)	91(3)	78(1)
Forest	65(5)	86(3)	76(4)	77(4)	94(2)	81(4)	82(4)	78(4)	86(3)	91(3)	83(1)
Grassland	88(3)	80(4)	86(3)	51(5)	77(4)	25(4)	32(5)	55(5)	29(5)	45(5)	76(2)
Shrubland	56(5)	64(5)	67(5)	69(5)	54(5)	69(5)	70(5)	57(5)	43(5)	57(5)	64(2)
Wetland	85(4)	83(4)	86(3)	90(3)	88(3)	58(5)	80(4)	79(4)	58(5)	70(5)	79(2)
Water bodies	90(3)	97(2)	92(3)	87(3)	89(3)	86(3)	88(3)	91(3)	93(3)	91(3)	92(1)
ArS	63(5)	50(5)	63(5)	67(5)	67(5)	75(4)	58(5)	78(4)	73(4)	71(5)	70(2)
Bareland	92(3)	81(4)	84(4)	59(5)	95(2)	50(5)	53(5)	29(5)	41(5)	67(5)	90(2)
PSI	87(3)	87(3)	98(1)	80(4)	–	–	–	–	–	–	88(2)
Overall	90(2)	81(3)	83(2)	69(3)	85(2)	73(3)	73(3)	80(2)	77(2)	84(2)	

Table A12. Regional producer's accuracies for nine land cover classes with standard errors (SE) in parentheses when agreement is defined as a match with the primary label.

	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	National
CuL	98(0)	98(0)	96(1)	86(4)	75(6)	89(4)	87(2)	94(1)	91(2)	91(2)	89(1)
Forest	52(12)	87(7)	88(5)	83(3)	87(6)	89(2)	90(3)	79(6)	88(2)	87(2)	86(1)
Grassland	80(5)	93(1)	88(2)	83(6)	95(1)	31(9)	48(12)	73(8)	33(9)	79(9)	87(1)
Shrubland	13(4)	13(3)	11(2)	13(2)	10(3)	6(1)	1(0)	1(0)	1(0)	0(0)	9(1)
Wetland	42(16)	6(3)	18(7)	32(12)	93(2)	2(1)	15(6)	58(12)	45(9)	89(3)	28(5)
Water bodies	99(1)	77(15)	98(1)	62(21)	46(24)	63(14)	63(12)	72(19)	58(8)	72(14)	70(5)
ArS	62(16)	38(10)	31(9)	19(6)	68(15)	46(10)	34(6)	76(10)	71(9)	64(12)	53(4)
Bareland	99(1)	84(6)	92(3)	9(3)	89(4)	1(1)	30(21)	3(3)	5(4)	100(0)	93(1)
PSI	90(8)	95(4)	96(3)	84(4)	–	–	–	–	–	–	93(4)

Table A13. The country-wide error matrix; cell entries are expressed as percent of area (see Table 5 for meanings of CuL, ArS, and PSI). Agreement is defined as a match between the map label and the primary reference label. User's accuracy (UA) and producer's accuracy (PA) are reported with standard errors (SE) in parentheses. Overall accuracy is 81.0% (0.8%).

	CuL	Forest	Grassland	Shrubland	Wetland	Water	ArS	Bareland	PSI	Total	UA
CuL	16.7	1.5	0.6	1	0.2	0.3	1	0.1	0	21.3	78(1)
Forest	0.6	18.4	0.8	2.2	0	0.1	0	0.1	0	22.2	83(1)
Grassland	1	1.2	22.5	3	0.5	0.2	0.1	1.1	0	29.6	76(2)
Shrubland	0	0.1	0.2	0.7	0	0	0	0	0	1	64(2)
Wetland	0	0	0	0	0.3	0	0	0	0	0.4	79(2)
Water	0	0	0	0	0.1	1.4	0	0	0	1.6	92(1)
ArS	0.3	0.1	0	0	0	0	1.3	0	0	1.8	70(2)
Bareland	0	0	1.6	0.4	0.1	0	0	18.7	0	20.9	90(2)
PSI	0	0	0.1	0	0	0	0	0	0.9	1.1	88(2)
Total	18.7	21.3	25.8	7.3	1.2	2	2.4	20.2	1	100	
PA	89(1)	86(1)	87(1)	9(1)	28(5)	70(5)	53(4)	93(1)	93(4)		

References

1. Feddema, J.J.; Oleson, K.W.; Bonan, G.B.; Mearns, L.O.; Buja, L.E.; Meehl, G.A.; Washington, W.M. The importance of land-cover change in simulating future climates. *Science* **2005**, *310*, 1674–1678. [[CrossRef](#)] [[PubMed](#)]
2. Herold, M.; See, L.; Tsendbazar, N.E.; Fritz, S. Towards an Integrated Global Land Cover Monitoring and Mapping System. *Remote Sens.* **2016**, *8*, 1036. [[CrossRef](#)]
3. Anderson, K.; Ryan, B.; Sonntag, W.; Kavvada, A.; Friedl, L. Earth observation in service of the 2030 agenda for sustainable development. *Geo-Spat. Inf. Sci.* **2017**, *20*, 77–96. [[CrossRef](#)]
4. Loveland, T.R.; Belward, A. The IGBP-DIS global 1km land cover data set, DISCover: First results. *Int. J. Remote Sens.* **1997**, *18*, 3289–3295. [[CrossRef](#)]
5. 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.* **2000**, *21*, 1331–1364. [[CrossRef](#)]
6. Bartholomé, E.; Belward, A.S. GLC2000: A new approach to global land cover mapping from Earth observation data. *Int. J. Remote Sens.* **2005**, *26*, 1959–1977. [[CrossRef](#)]
7. 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](#)]
8. Arino, O.; Bicheron, P.; Achard, F.; Latham, J.; Witt, R.; Weber, J.-L. GLOBCOVER: The most detailed portrait of Earth. *Eur. Space Agency Bull.* **2008**, *136*, 24–31.
9. Bontemps, S.; Defourny, P.; Bogaert, E.V.; Arino, O.; Kalogirou, V.; Perez, J.R. GLOBCOVER 2009-Products Description and Validation Report, Version 2.2, February 2011. Available online: http://due.esrin.esa.int/files/GLOBCOVER2009_Validation_Report_2.2.pdf (accessed on 2 August 2018).
10. Plummer, S.; Lecomte, P.; Doherty, M. The ESA Climate Change Initiative (CCI): A European contribution to the generation of the Global Climate Observing System. *Remote Sens. Environ.* **2017**, *203*, 2–8. [[CrossRef](#)]
11. Kobayashi, T.; Tateishi, R.; Alsaaidh, B.; Sharma, R.C.; Wakaizumi, T.; Miyamoto, D.; Bai, X.; Long, B.D.; Gegentana, G.; Maitiniyazi, A. Production of Global Land Cover Data–GLCNMO2013. *J. Geogr. Geol.* **2017**, *9*, 1. [[CrossRef](#)]
12. Chen, J.; Ban, Y.; Li, S. China: Open access to Earth land-cover map. *Natural* **2014**, *514*, 434. [[CrossRef](#)]
13. Stehman, S.V.; Czaplewski, R.L. Design and Analysis for Thematic Map Accuracy Assessment: Fundamental Principles. *Remote Sens. Environ.* **1998**, *64*, 331–344. [[CrossRef](#)]
14. Tong, X.; Wang, Z.; Xie, H.; Liang, D.; Jiang, Z.; Li, J.; Li, J. Designing a two-rank acceptance sampling plan for quality inspection of geospatial data products. *Comput. Geosci.* **2011**, *37*, 1570–1583. [[CrossRef](#)]
15. Chen, J.; Chen, J.; Liao, A.; Cao, X.; Chen, L.; Chen, X.; He, C.; Han, G.; Peng, S.; Lu, M. Global land cover mapping at 30 m resolution: A POK-based operational approach. *ISPRS J. Photogramm. Remote Sens.* **2015**, *103*, 7–27. [[CrossRef](#)]
16. Manakos, I.; Karakizi, C.; Gkinis, I.; Karantzalos, K. Validation and inter-comparison of spaceborne derived global and continental land cover products for the Mediterranean region: The case of Thessaly. *Land* **2017**, *6*, 34. [[CrossRef](#)]
17. Zhang, Y.; Chen, J.; Chen, L.; Li, R.; Zhang, W.; Lu, N.; Liu, J. Characteristics of land cover change in Siberia based on GlobeLand30, 2000–2010 (in Chinese). *Prog. Geogr.* **2015**, *34*, 1324–1333. [[CrossRef](#)]
18. Kussul, N.; Shelestov, A.; Basarab, R.; Skakun, S.; Kussul, O.; Lavreniuk, M. Geospatial intelligence and data fusion techniques for sustainable development problems. In Proceedings of the 11th International Conference on ICT in Education, Research and Industrial Applications, Lviv, Ukraine, 14–16 May 2015.
19. Jokar Arsanjani, J.; Tayyebi, A.; Vaz, E. GlobeLand30 as an alternative fine-scale global land cover map: Challenges, possibilities, and implications for developing countries. *Habitat Int.* **2016**, *55*, 25–31. [[CrossRef](#)]
20. See, L.; Laso Bayas, J.; Schepaschenko, D.; Perger, C.; Dresel, C.; Maus, V.; Salk, C.; Weichselbaum, J.; Lesiv, M.; McCallum, I. LACO-Wiki: A New Online Land Cover Validation Tool Demonstrated Using GlobeLand30 for Kenya. *Remote Sens.* **2017**, *9*, 754. [[CrossRef](#)]
21. Sun, B.; Chen, X.; Zhou, Q. Uncertainty Assessment of GLOBELAND30 Land Cover Data Set Over Central Asia. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *XLI-B8*, 1313–1317. [[CrossRef](#)]

22. Cao, X.; Li, A.; Lei, G.; Tan, J.; Zhang, Z.; Yan, D.; Xie, H.; Zhang, S.; Yang, Y.; Mingjiang, S. Land Cover Mapping and Spatial Pattern Analysis with Remote Sensing in Nepal. *J. Geo-Inf. Sci.* **2016**, *18*, 1384–1398. (In Chinese) [[CrossRef](#)]
23. Manakos, I.; Petrou, Z.I.; Filchev, L.; Apostolakis, A. Globalland30 Mapping Capacity of Land Surface Water in Thessaly, Greece. *Land* **2014**, *4*, 1–18. [[CrossRef](#)]
24. Pérezhoyos, A.; Rembold, F.; Kerdiles, H.; Gallego, J. Comparison of Global Land Cover Datasets for Cropland Monitoring. *Remote Sens.* **2017**, *9*. [[CrossRef](#)]
25. Arsanjani, J.J.; See, L.; Tayyebi, A. Assessing the suitability of GlobeLand30 for mapping land cover in Germany. *Int. J. Digital Earth* **2016**, *9*, 873–891. [[CrossRef](#)]
26. Brovelli, A.M.; Molinari, E.M.; Hussein, E.; Chen, J.; Li, R. The First Comprehensive Accuracy Assessment of GlobeLand30 at a National Level: Methodology and Results. *Remote Sens.* **2015**, *7*. [[CrossRef](#)]
27. Mozak, S. Comparing Global Land Cover Datasets through the Eagle Matrix Land Cover Components for Continental Portugal. Master's Thesis, Nova Information Management School, Universitat Jaume, Lisbon, Portugal, 2016.
28. Jacobson, A.; Dhanota, J.; Godfrey, J.; Jacobson, H.; Rossman, Z.; Stanish, A.; Walker, H.; Riggio, J. A novel approach to mapping land conversion using Google Earth with an application to East Africa. *Environ. Model. Soft.* **2015**, *72*, 1–9. [[CrossRef](#)]
29. Yang, Y.; Xiao, P.; Feng, X.; Li, H. Accuracy assessment of seven global land cover datasets over China. *ISPRS J. Photogramm. Remote Sens.* **2017**, *125*, 156–173. [[CrossRef](#)]
30. Lu, M.; Wu, W.; Zhang, L.; Liao, A.; Peng, S.; Tang, H. A comparative analysis of five global cropland datasets in China. *Sci. China Earth Sci.* **2016**, *59*, 2307–2317. [[CrossRef](#)]
31. 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](#)]
32. Meng, W.; Tong, X.; Xie, H.; Wang, Z. Accuracy Assessment for Regional Land Cover Remote Sensing Mapping Product Based on Spatial Sampling: A Case Study of Shaanxi Province, China. *J. Geo-Inf. Sci.* **2015**, *17*, 742–749. [[CrossRef](#)]
33. Ma, J.; Sun, Q.; Xiao, Q.; Wen, B. Accuracy assessment and comparative analysis of GlobeLand30 dataset in Henan province. *J. Geo-Inf. Sci.* **2016**, *18*, 1563–1572. [[CrossRef](#)]
34. Wickham, J.; Homer, C.; Vogelmann, J.; McKerrow, A.; Mueller, R.; Herold, N.; Coulston, J. The multi-resolution land characteristics (MRLC) consortium—20 years of development and integration of USA national land cover data. *Remote Sens.* **2014**, *6*, 7424–7441. [[CrossRef](#)]
35. Wickham, J.D.; Stehman, S.V.; Smith, J.H.; Yang, L. Thematic accuracy of the 1992 National Land-Cover Data for the western United States. *Remote Sens. Environ.* **2004**, *91*, 452–468. [[CrossRef](#)]
36. Wickham, J.; Stehman, S.; Fry, J.A.; Smith, J.; Homer, C.G. Thematic accuracy of the NLCD 2001 land cover for the conterminous United States. *Remote Sens. Environ.* **2010**, *114*, 1286–1296. [[CrossRef](#)]
37. Wickham, J.D.; Stehman, S.V.; Gass, L.; Dewitz, J.; Fry, J.A.; Wade, T.G. Accuracy assessment of NLCD 2006 land cover and impervious surface. *Remote Sens. Environ.* **2013**, *130*, 294–304. [[CrossRef](#)]
38. Wickham, J.; Stehman, S.V.; Gass, L.; Dewitz, J.A.; Sorenson, D.G.; Granneman, B.J.; Poss, R.V.; Baer, L.A. Thematic accuracy assessment of the 2011 National Land Cover Database (NLCD). *Remote Sens. Environ.* **2017**, *191*, 328–341. [[CrossRef](#)]
39. Stehman, S.V.; Wickham, J.D.; Smith, J.H.; Yang, L. Thematic accuracy of the 1992 National Land-Cover Data for the eastern United States: Statistical methodology and regional results. *Remote Sens. Environ.* **2003**, *86*, 500–516. [[CrossRef](#)]
40. Stehman, S.V.; Wickham, J.D.; Wade, T.G.; Smith, J.H. Designing a Multi-Objective, Multi-Support Accuracy Assessment of the 2001 National Land Cover Data (NLCD 2001) of the Conterminous United States. *Photogramm. Eng. Remote Sens.* **2008**, *74*, 1561–1571. [[CrossRef](#)]
41. Olofsson, P.; Foody, G.M.; Herold, M.; Stehman, S.V.; Woodcock, C.E.; Wulder, M.A. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* **2014**, *148*, 42–57. [[CrossRef](#)]
42. Stehman, S.V.; Wickham, J.D. Pixels, blocks of pixels, and polygons: Choosing a spatial unit for thematic accuracy assessment. *Remote Sens. Environ.* **2011**, *115*, 3044–3055. [[CrossRef](#)]
43. Congalton, R.G.; Green, K. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 2008; pp. 63–83.

44. Congalton, R.G. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens. Environ.* **1991**, *37*, 35–46. [[CrossRef](#)]
45. Olofsson, P.; Foody, G.M.; Stehman, S.V.; Woodcock, C.E. Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sens. Environ.* **2013**, *129*, 122–131. [[CrossRef](#)]
46. Stehman, S.V. Estimating area and map accuracy for stratified random sampling when the strata are different from the map classes. *Int. J. Remote Sens.* **2014**, *35*, 4923–4939. [[CrossRef](#)]
47. Smith, J.H.; Wickham, J.D.; Stehman, S.V.; Yang, L. Impacts of patch size and land-cover heterogeneity on thematic image classification accuracy. *Photogramm. Eng. Remote Sens.* **2002**, *68*, 65–70.
48. Fonte, C.C.; Minghini, M.; Patriarca, J.; Antoniou, V.; See, L.; Skopeliti, A. Generating Up-to-Date and Detailed Land Use and Land Cover Maps Using OpenStreetMap and GlobeLand30. *ISPRS Int. J. Geo-Inf.* **2017**, *6*. [[CrossRef](#)]
49. Cao, X.; Chen, X.; Zhang, W.; Liao, A.; Chen, L.; Chen, Z.; Chen, J. Global cultivated land mapping at 30 m spatial resolution. *Sci. China Earth Sci.* **2016**, *46*, 1426–1435. [[CrossRef](#)]
50. Vogelmann, J.E.; Howard, S.M.; Yang, L.; Larson, C.R.; Wylie, B.K.; Van Driel, J.N. Completion of the 1990s National Land Cover Data set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. *Photogramm. Eng. Remote Sens.* **2001**, *67*, 650–662.
51. Yu, L.; Wang, J.; Gong, P. Improving 30 m global land-cover map FROM-GLC with time series MODIS and auxiliary data sets: A segmentation-based approach. *Int. J. Remote Sens.* **2013**, *34*, 5851–5867. [[CrossRef](#)]
52. Homer, C.G.; Dewitz, J.; Fry, J.; Coan, M.; Hossain, N.; Larson, C.; Herold, N.; McKerrow, A.; VanDriel, J.N.; Wickham, J. Completion of the 2001 National Land Cover Database for the conterminous United States. *Photogramm. Eng. Remote Sens.* **2007**, *73*, 337–341.
53. Fry, J.; Xian, G.Z.; Jin, S.; Dewitz, J.; Homer, C.G.; Yang, L.; Barnes, C.A.; Herold, N.D.; Wickham, J.D. Completion of the 2006 national land cover database for the conterminous united states. *Photogramm. Eng. Remote Sens.* **2011**, *77*, 858–864.
54. Homer, C.; Dewitz, J.; Yang, L.; Jin, S.; Danielson, P.; Xian, G.; Coulston, J.; Herold, N.; Wickham, J.; Megown, K. Completion of the 2011 National Land Cover Database for the Conterminous United States—Representing a Decade of Land Cover Change Information. *Photogramm. Eng. Remote Sens.* **2015**, *81*, 345–354.
55. Chen, J.; Cao, X.; Peng, S.; Ren, H.; Chen, J.; Cao, X.; Peng, S.; Ren, H. Analysis and Applications of GlobeLand30: A Review. *Int. J. Geo-Inf.* **2017**, *6*, 230. [[CrossRef](#)]



© 2018 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 (<http://creativecommons.org/licenses/by/4.0/>).