

Review

Global Land Cover Mapping: A Review and Uncertainty Analysis

Russell G. Congalton ^{1,†,*}, Jianyu Gu ^{1,2,†}, Kamini Yadav ¹, Prasad Thenkabail ³ and Mutlu Ozdogan ⁴

¹ Department of Natural Resources and the Environment, University of New Hampshire, 56 College Road, Durham, NH 03824, USA; E-Mails: gujianyu23@gmail.com (J.G.); kaminiyadav.02@gmail.com (K.Y.)

² State Key Laboratory of Earth Surface Processes and Resource Ecology, College of Resources Science and Technology, Beijing Normal University, No. 19, XinJieKouWai St., HaiDian District, Beijing 100875, China

³ US Geological Survey, 2255 N. Gemini Drive, Flagstaff, AZ 86001, USA; E-Mail: thenkabail@gmail.com

⁴ Department of Forest and Wildlife Ecology, University of Wisconsin, 1710 University Ave., Room 285, Madison, WI 53726, USA; E-Mail: ozdogan@wisc.edu

† These authors contributed equally to this work.

* Author to whom correspondence should be addressed; E-Mail: russ.congalton@unh.edu; Tel.: +1-603-862-4644; Fax: +1-603-862-4976.

External Editors: Janet Nichol and Giles Food

Received: 10 September 2014; in revised form: 6 November 2014 / Accepted: 24 November 2014 / Published: 3 December 2014

Abstract: Given the advances in remotely sensed imagery and associated technologies, several global land cover maps have been produced in recent times including IGBP DISCover, UMD Land Cover, Global Land Cover 2000 and GlobCover 2009. However, the utility of these maps for specific applications has often been hampered due to considerable amounts of uncertainties and inconsistencies. A thorough review of these global land cover projects including evaluating the sources of error and uncertainty is prudent and enlightening. Therefore, this paper describes our work in which we compared, summarized and conducted an uncertainty analysis of the four global land cover mapping projects using an error budget approach. The results showed that the classification scheme and the validation methodology had the highest error contribution and implementation priority. A comparison of the

classification schemes showed that there are many inconsistencies between the definitions of the map classes. This is especially true for the mixed type classes for which thresholds vary for the attributes/discriminators used in the classification process. Examination of these four global mapping projects provided quite a few important lessons for the future global mapping projects including the need for clear and uniform definitions of the classification scheme and an efficient, practical, and valid design of the accuracy assessment.

Keywords: global land cover; uncertainty analysis; error budget; classification scheme; accuracy assessment

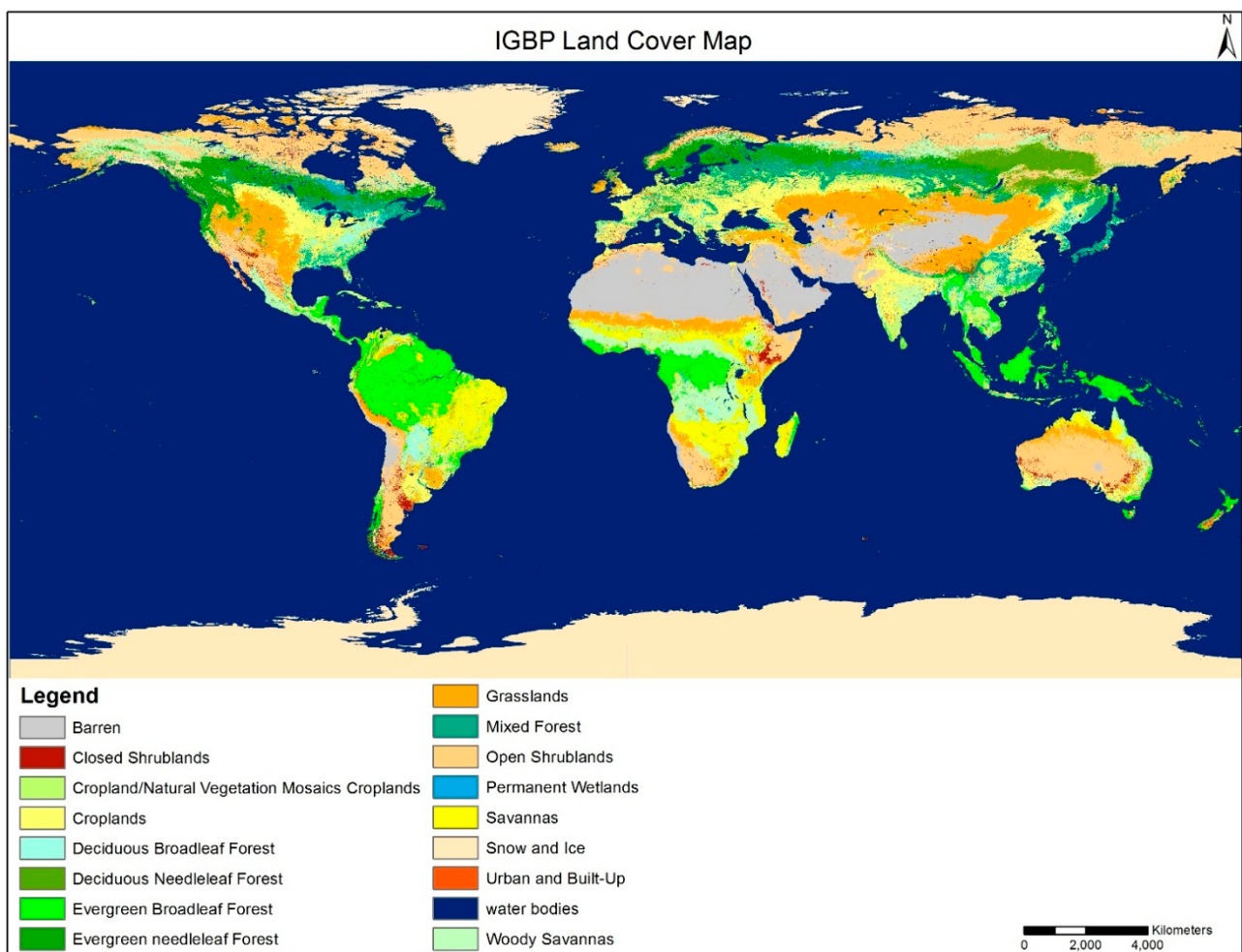
1. Introduction

Global land cover has been identified as one of the fundamental variables needed in order to study the morphological and functional changes occurring in the Earth's ecosystems and the environment including climate change and carbon circulation [1,2]. Compared with traditional methods (e.g., field surveys) to describe the earth's surface, remote sensing is more efficient and effective because of its ability to map and monitor the spatial distribution of land cover continuously and consistently at a variety of spatial and temporal scales. Remotely sensed imagery can also be used as an intermediate product to serve as a basis for spatial inferences (e.g., [3,4]). However, this paper concentrates on mapping spatial distribution of land cover. An upsurge in global land cover mapping began after the launch of the National Oceanographic and Atmospheric Administration (NOAA) satellite, equipped with Advanced Very High Resolution Radiometer (AVHRR) instrument, whose data are available at a global scale [5]. The subsequent rapid development of remotely sensed imagery and technologies offers more opportunities for national and international initiatives to implement global mapping projects using higher spatial, spectral, and temporal resolution images. Several global land cover maps have been produced in recent times including: IGBP DISCover (Figure 1, data obtained from [6]) [7–9], UMD Land Cover (Figure 2, data obtained from [10]) [11], Global Land Cover 2000 (Figure 3, data obtained from [12]) [13] and GlobCover 2009 (Figure 4, data obtained from [14]) [15].

One of the initial purposes of these global mapping projects was to serve the scientific and research communities by producing a variety of global land cover products. However, the users of these maps have often found it difficult to effectively apply these products to their specific applications due to compelling amounts of uncertainty and inconsistency that occurred in these maps [16]. The main reasons for these uncertainties and inconsistencies are: (1) these map products were based on different remote sensing collection devices (*i.e.*, sensors); (2) a variety of different methodologies were employed to create the map products; and (3) discrepant map class definitions (*i.e.*, using different classification schemes) were employed despite claims from some that they had followed the same classification scheme standard. Table 1 shows the map classes used for each of the different land cover classification schemes for each global land cover map investigated in this paper. Much effort has been made to compare these existing global land cover products, highlighting the strengths and weaknesses of each [16–19]. To facilitate these comparisons, a translation of classification schemes (or crosswalk) between the maps was necessary. The Land Cover Classification System (LCCS), developed by the Food and Agriculture Organization of

the United Nations (FAO) and the United Nations Environment Program (UNEP), has often been used as a general framework for translating between different map classification schemes [20]. These comparisons have resulted in some generalized and worthwhile conclusions such as: (1) demonstrating the spatial disagreement of existing maps [17,18,21]; (2) documenting the inability to discriminate the mixed classes [11,18]; and (3) demonstrating the strong relationship between the spatial heterogeneity and the resulting map accuracy [16,22]. These conclusions provide some beneficial suggestions for use in future global land cover mapping.

Figure 1. The IGBP Land Cover Map (figure generated from data obtained at [6]).

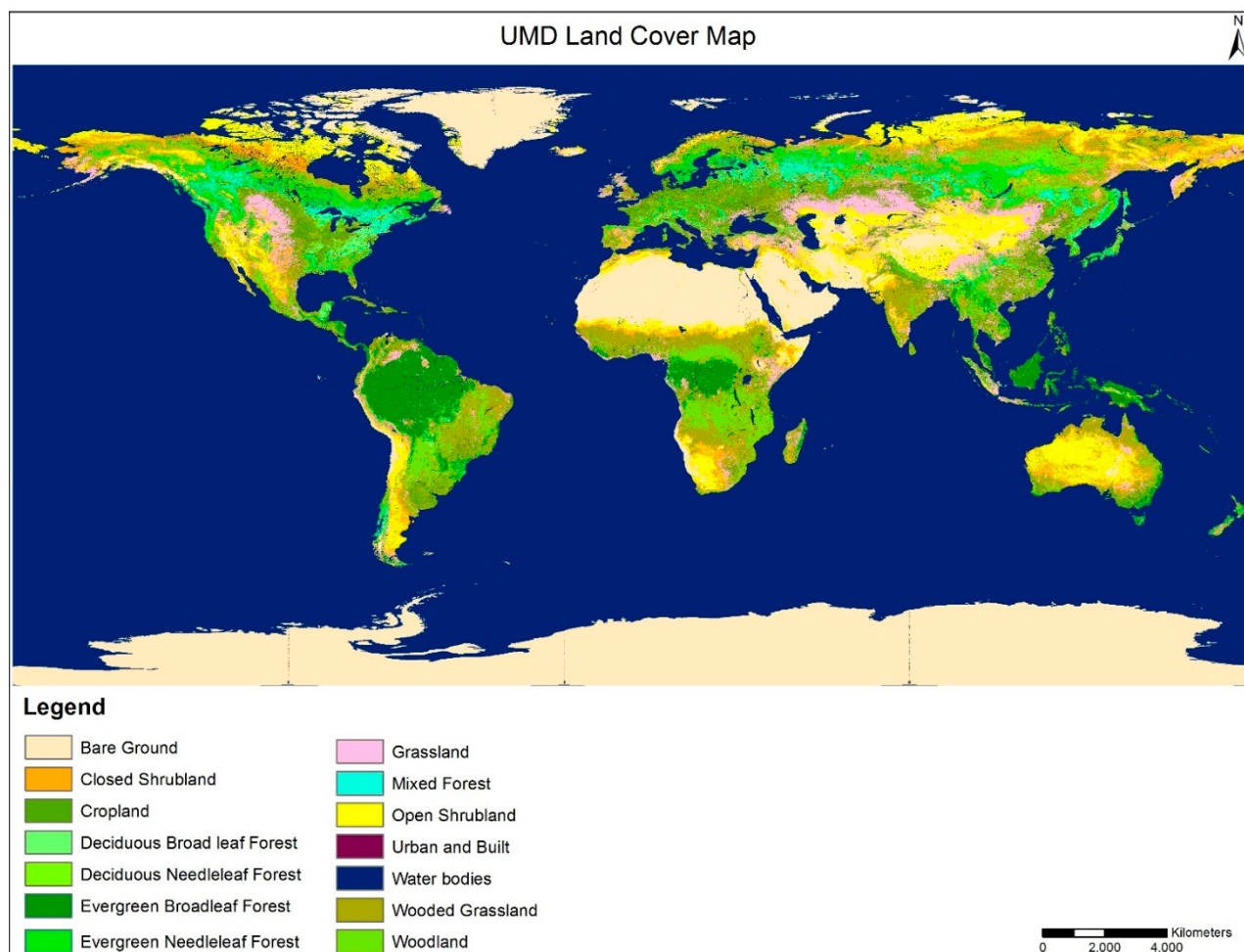


However, a great deal more can and should be gleaned by carefully reviewing these global land cover projects and evaluating the sources of error and uncertainty. A quote commonly attributed to George Santayana [23] states: *“Those who don’t know history are doomed to repeat it”*. In any future global land cover mapping effort it is absolutely critical to learn from the past so as to not make the same mistakes over again. Therefore, the specific objectives of this paper are:

(1) To intensively review these previous global land cover mapping projects to determine what lessons can be learned for future mapping projects.

(2) To perform an uncertainty analysis using an error budgeting approach of the mapping methods used to produce the spatial distribution of land cover types for these previous global mapping projects to better prepare for future projects.

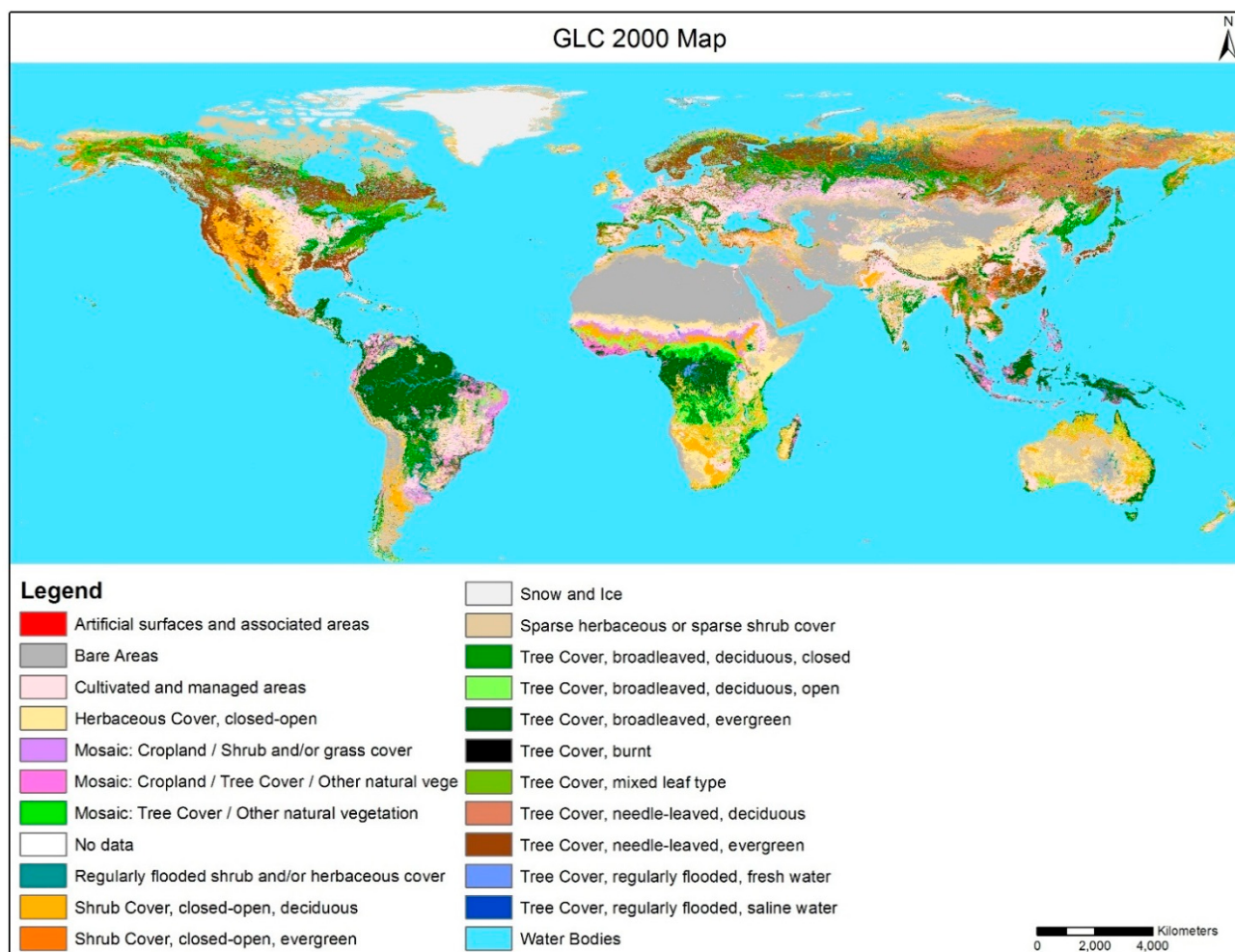
Figure 2. The UMD Land Cover Map (figure generated from data obtained at [10]).



2. Methods

The objectives of this paper can be satisfied using two main techniques. The first is a simple review and in-depth study of the selected global mapping projects to develop lessons learned for the future. The second is an uncertainty analysis conducted using an error budgeting approach [24,25]. We focused on the following four well known global land use land cover (LULC) mapping projects: IGBP DISCover [26–29], UMD Land Cover [11,18], Global Land Cover 2000 [13], and GlobCover 2009 [30–32].

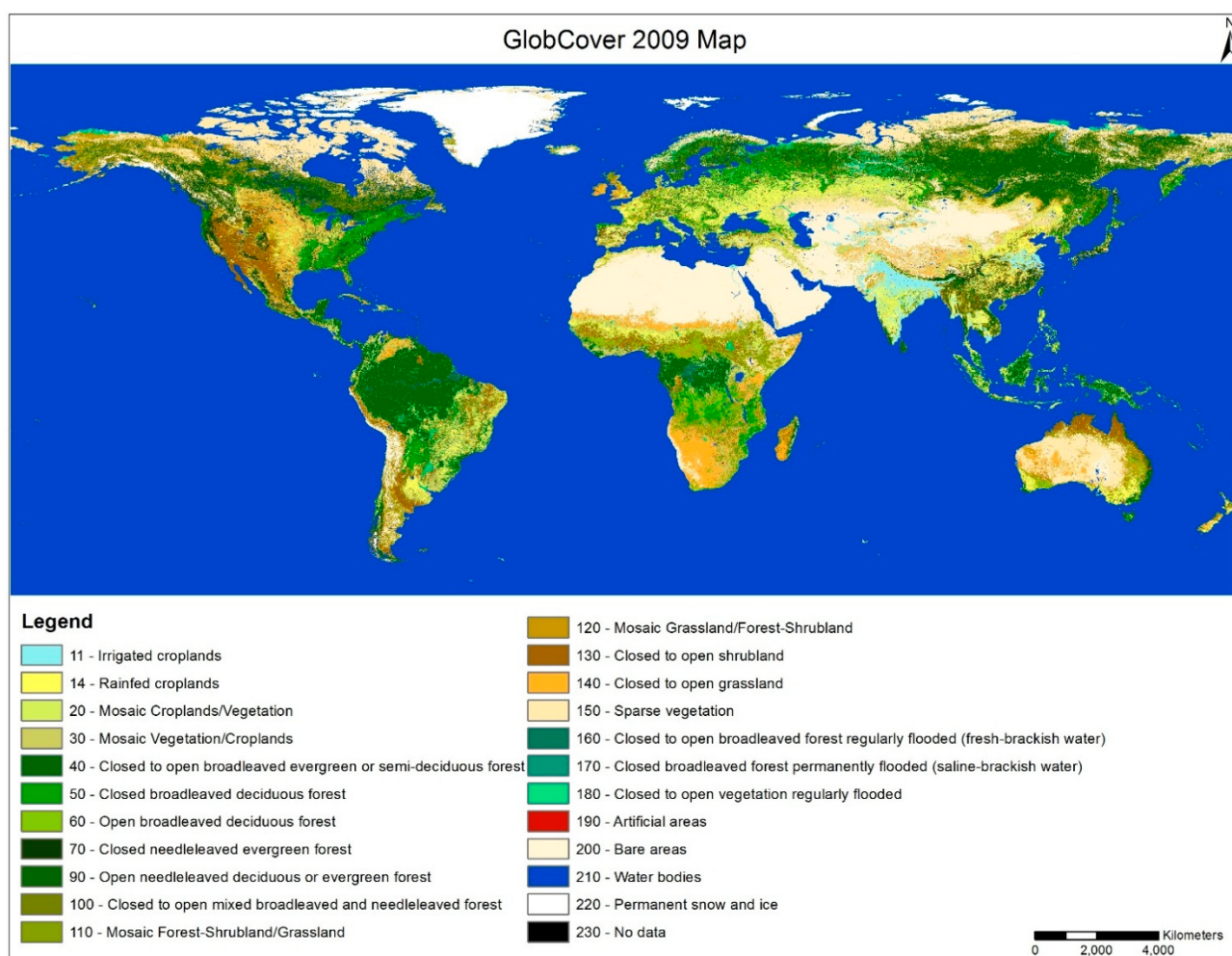
The review process summarized these four global mapping projects from the following aspects: producer, sensor, input data, preprocessing, classification, accuracy assessment, and the associated website. All information used in the evaluation and review was collected from the extensive project reports, published papers and websites developed for each project.

Figure 3. The GLC 2000 Land Cover Map (figure generated from data obtained at [12]).

The error budget procedure developed by [25] was adopted to analyze the potential uncertainty in each of the global mapping projects. The entire mapping project was divided into several component parts and the uncertainties that exist in each part of the global mapping process was evaluated and analyzed from the following three perspectives: (1) error contribution potential; (2) implementation difficulty; and (3) implementation priority. A relative rank was given for each of the components. Error contribution potential is defined as the degree of uncertainty that impacts the product. It helps us to understand which components contribute the most or least to the overall error. Implementation difficulty is characterized by the degree of difficulty to control or correct the uncertainty given existing technology. It provides understanding into which errors are easy to correct and which are difficult. Implementation priority is the combination of the potential errors and implementation difficulty. It is a useful indicator because some types of errors have greater potential to cause serious issues and are more difficult to fix while other may have the same error potential, but are much easier to correct. These errors that cause the most problems, but are easy to fix should be considered first in the future mapping projects. In this paper, we divided each global land cover mapping project into the following major components: systematic, natural, input data, ancillary data, preprocessing, classification method, processing sequence and accuracy assessment. The relative rank in error contribution potential, implementation difficulty and implementation priority was determined by review and evaluation of the existing issues extracted from published papers and by comparison with the global land covers mapping projects. It is recognized that the error budget

analysis is a subjective process and only provides relative answers. However, it remains a powerful tool as it encourages the mapping scientist to spend considerable time thinking about all the various aspects of the mapping project and therefore, results in selection of the best methods possible. Given limited time and resources, this approach reveals what aspects of the project will yield the most benefit for the least effort. Therefore, the error budgeting analysis helps the mapping scientist to be as efficient and as effective as possible.

Figure 4. The GLOB Cover 2009 Land Cover Map (figure generated from data obtained at [14]).



A classification scheme is a hierarchically structured group of classes with descriptive information based on their characteristics in common. In other words, it is a way of labeling information into groups so that the information can be effectively described, managed, or processed. For example, land cover can be divided into forests, grass, brush, water, *etc.* There are many different classification schemes that have been developed for a variety of purposes. The Land Cover Classification System (LCCS) [20] is one such system that has standardized classes with well-established definitions and thresholds. It has been widely used as a basic scheme for use in global land cover mapping and also acts as a general framework to translate (crosswalk) the land cover classes from different land cover datasets into a common set of labels.

Table 1. Different Global Land Cover legends, their class names and numbers.

	GLC 2000	Glob Cover 2009	IGBP	UMD
1	Tree Cover, broadleaved, evergreen	Closed to open (>15%) broadleaved evergreen or semi-deciduous forest (>5 m)	Evergreen Broadleaf Forests	Evergreen Broadleaf Forests
2	Tree Cover, broadleaved, deciduous, closed	Closed (>40%) broadleaved deciduous forest (>5 m)	Deciduous Broadleaf Forests	Deciduous Broadleaf Forests
3	Tree Cover, broadleaved, deciduous, open	Open (15%–40%) broadleaved deciduous forest/woodland (>5 m)		
4	Tree Cover, needle-leaved, evergreen	Closed (>40%) needle leaved evergreen forest (>5 m)	Evergreen Needle leaf Forests	Evergreen Needle leaf Forests
5	Tree Cover, needle-leaved, deciduous	Open (15%–40%) needle leaved deciduous or evergreen forest (>5 m)	Deciduous Needle leaf Forests	Deciduous Needle leaf Forests
6	Tree Cover, mixed leaf type	Closed to open (>15%) mixed broadleaved and needle leaved forest (>5 m)	Mixed Forests	Mixed Forests
7	Tree Cover, regularly flooded, fresh water			
8	Tree Cover, regularly flooded, saline water			
9	Mosaic: Tree Cover/other natural vegetation	Mosaic grassland (50%–70%)/forest or shrub land (20%–50%) Mosaic Forest/Shrubland (50%–70%)/Grassland (20%–50%)	Grasslands	Grasslands
10	Tree Cover, burnt			
11	Shrub Cover, closed–open, evergreen	Closed to open (>15%) (broadleaved or needle leaved, evergreen or deciduous) shrub land (<5 m)	Closed shrub lands	Closed Bush lands or Shrub lands
12	Shrub Cover, closed–open, deciduous		Open shrub lands	Open shrub lands Wooded Grasslands/shrub lands
13	Herbaceous Cover, closed–open	Closed to open (>15%) herbaceous vegetation (grassland, savannas or lichens/mosses)	Woody savannas	Woodlands
14	Sparse Herbaceous or sparse Shrub Cover	Sparse (<15%) vegetation (woody vegetation, shrubs, grassland)	Savannas	
15	Regularly flooded Shrub and/or Herbaceous Cover	Closed to open (>15%) grassland or woody vegetation on regularly flooded or waterlogged soil—Fresh, brackish or saline water Closed (>40%) broadleaved forest or shrub land permanently flooded—Saline or brackish Closed to open (>15%) broadleaved forest regularly flooded (semi-permanently or temporarily)—Fresh or brackish water	Permanent Wetlands	
16	Cultivated and managed areas	Post-flooding or irrigated croplands (or aquatic)	Croplands	

Table 1. Cont.

	GLC 2000	Glob Cover 2009	IGBP	UMD
17	Mosaic: Cropland/Tree Cover/other natural vegetation	Rain fed croplands	Cropland/Natural Vegetation Mosaics croplands	Croplands
18	Mosaic: Cropland/Shrub or Grass Cover	Mosaic cropland (50%–70%)/vegetation (20%–50%) Mosaic vegetation (50%–70%)/cropland (20%–50%)		
19	Bare areas	Bare areas	Barren	Barren
20	Water bodies	Water bodies	Water bodies	Water bodies
21	Snow and Ice	Permanent snow and ice	Snow and Ice	
22	Artificial surfaces and associated areas	Artificial surfaces and associated areas (Urban areas > 50%)	Urban and Built-up	Urban and Built-up

Previous research has shown that the classification scheme is one of the major sources of differences in global land cover mapping [16–18]. Inconsistencies in the class definitions among maps were widespread despite some of them using the class labels from a standardized classification scheme (e.g., LCCS). Many researchers have taken advantage of LCCS to study how to convert class labels from different land cover maps to improve the interoperability and compatibility. In this paper, we focused not on this translation but rather on the analysis of the uncertainty in the class definitions. We have broken each class definition into a set of independent attributes or discriminators (e.g., tree height and canopy cover percent) associated with the thresholds given by LCCS. We used this method to compare and analyze the differences in the classification schemes among the global land cover mapping projects. We also analyzed the “mappability” (ability to actually discern these attributes) of these independent diagnostic criteria. Because most of these criteria are defined from an ecological or environmental perspective, not all of them can be identified from optical remotely sensed imagery, especially at rather coarse spatial resolution. What further information is needed to enhance the effectiveness of each criterion was also considered as improved imagery and other geospatial information are rapidly becoming more available. We limited our analysis of the classification scheme to map classes within only the forest/trees and cropland categories because previous research has suggested that relatively more inconsistencies existed in these categories [18,33].

3. Results

3.1. Summary of Global Land Cover Mapping Databases

We reviewed and summarized IGBP DISCover, UMD Land Cover, Globe Land Cover 2000 and GlobCover 2009, as shown in Table 2 [10,12,14,34]. Generally, the institutions responsible for creating these four land cover databases employed different remotely sensed imagery (sensors), methodologies, and validation techniques to produce their global land cover maps. Table 2 provides a clear overview of these differences by category.

Table 2. Summary of the characteristics of global land cover mapping databases.

	IGBP	UMD	GLC 2000	GlobCover 2009
Producer	USGS, UNL, JRC	UMD	JRC	ESA
Sensor	AVHRR	AVHRR	SPOT VEGETATION-1	ENVISAT MERIS
Input Data	<i>Primary</i>	Monthly global NDVI	41 temporal metrics from	13 Spectral Bands
	<i>Input Data</i>	composites	spectral bands and NDVI	and NDVI composites
	<i>Collection</i>	April 1992–March 1993	November 1999—	January 2009—
	<i>Date</i>	April 1992–March 1993	December 2000	December 2009
	<i>Ancillary</i>	DEM Atlases of ecoregion, soils, vegetation Land cover maps	Radar DMSP Elevation Data (ETOPO5)	Altimeter Corrected Elevations (Getasse 30)
	<i>Spatial</i>	1 km	1 km	300 m
	<i>Resolution</i>			
Preprocessing	<i>Projection</i>	Goode Homolosine Equal Area projection	Goode Homolosine Equal Area projection	Lat-Lon
	<i>Geometric</i>	Geo-registered to Goode Homolosine equal area projection	Geo-registered to Goode Homolosine equal area projection	Ortho-rectification with ETOPO5, resampled by bi-cubic convolution
	<i>Correction</i>			Level 1B data corrected into Level 3 Mosaics using AMORGOS tool
	<i>Atmospheric</i>	Reduce atmospheric contamination and decrease off-nadir viewing effects by NDVI composition	Atmospherically corrected for ozone and Rayleigh scattering and solar zenith angle	Cloud screening Reduce Abrupt signal drops
Classification	<i>Correction</i>		Cloud screening Reduce Abrupt signal drops	Cloud screening Rayleigh scattering & Aerosol correction
	<i>Number of</i>	17	14	22
	<i>Classes</i>			
	<i>Training Sites</i>	N/A	37,249 training sites	N/A
	<i>Classification</i>	IGBP scheme	IGBP scheme	LCCS
	<i>Scheme</i>			
Accuracy Assessment	<i>Classification</i>	Unsupervised clustering with post-classification refinement	Supervised Decision tree	Unsupervised classification with ISODATA algorithm
	<i>Method</i>			Per-pixel supervised (urban and wetland) and unsupervised
	<i>Processing</i>	Continent-by-continent	Global	Region by region
	<i>Sequence</i>			Global
Accuracy Assessment	<i>Validation</i>	Statistical validation	None	Statistical Validation
	<i>Method</i>			
	<i>Sampling</i>	Stratified random sampling by classes		Two-stage stratified clustered sampling
	<i>Method</i>			Stratified random sampling by classes
	<i>Reference</i>	Landsat TM and Spot images	Landsat TM and Spot images	Reference dataset
Accuracy Assessment	<i>Data</i>			
	<i>Accuracy of</i>	Cropland: 85.7% Cropland/natural vegetation mosaic: 56.5%	76%	
	<i>Cropland</i>			
	<i>Total Accuracy</i>	66.9%	68.6%	67.5%
	<i>Web</i>	[34]	[10]	[12]
				[14]

3.1.1. IGBP-DISCover

IGBP-DISCover global land cover was created using 1 km Advanced Very High Resolution Radiometer (AVHRR) data spanning from April 1992 to March 1993 and was produced by the U.S. Geological Survey (USGS), the University of Nebraska-Lincoln (UNL), and the European Commission's Joint Research Center (JRC) [7–9]. A new land cover classification scheme was developed for this project and contained 17 land cover classes. The classification methodology employed was an unsupervised clustering followed by a post-classification refinement [26], which was applied continent-by-continent. Not all classes were labeled using this classification strategy. Instead, water bodies were masked by the hydrography layer of the Digital Chart of the World (DCW) [8,35] while the barren and snow and ice classes were generated by the threshold of the 12-month maximum NDVI composite [26,36].

Except for these three classes (water bodies, barren, snow and ice), each class was assessed using 25 random samples stratified by land cover type [29]. The reference cover types were interpreted from either Landsat TM or SPOT images by three interpreters. The IGBP Land Cover Working Group (LCWG) reported two versions of accuracy of this global land cover map according to degree of consensus on the reference land cover types [37]. In the first version, all three interpreters must agree on the reference label call. The sample point overall accuracy was 59.4 percent and area-weighted overall accuracy was 66.9 percent using this method. The second assessment version used a “Majority Rule” of the three interpreters and the resulting accuracy was 73.5 percent with an area weighted value of 78.7 percent [34].

3.1.2. UMD Land Cover

The University of Maryland created a global land cover map from 1 kilometer Advanced Very High Resolution Radiometer (AVHRR) spanning the April 1992 through March 1993 [11]. The input variables consisted of 41 data layers, all of which were transformed combinations of the individual spectral bands and NDVI values [11]. The data were radiometrically calibrated and geo-registered to the 1 km Goode's Interrupted Homolosine equal area projection [11,38,39]. The UMD classification scheme had 14 classes which were simplified from the original 17 IGBP land cover classes. In the UMD scheme, the crop-natural vegetation mosaic and wetlands were not included and the class of snow and ice was collapsed into the barren class. Minor differences in the definitions of some classes (e.g., tree height) exist between IGBP and UMD. Classification was performed used a decision tree that was pruned by visual interpretation of the preliminary results and then applied to determine the class membership. Not all classes were classified from the decision tree. The urban and built-up class was obtained from the existing 1km IGBP classification and the water class was labeled using the preliminary water mask from the MODIS sensor [8,11]. The training data for classification were originally produced for the 1984 8-km global land cover product by the National Aeronautics and Space Administration/National Oceanic and Atmospheric Administration Pathfinder Land (PAL) project and were generated from 156 Landsat MSS images [11,40].

No independent validation has been reported for this mapping; however some research by [11] has conducted an assessment using the existing training data and reported the accuracy was 69 percent.

3.1.3. Global Land Cover 2000

The Global Land Cover 2000 product was created using the 1 km VEGETATION sensor on-board SPOT 4 collected from November 1999 through December 2000 [13,41,42]. This project was an international partnership of 30 research groups coordinated by the European Commission's Joint Research Center (JRC) [13,43,44]. The project used a bottom up approach to global mapping of the world by dividing it into 19 regions and using local experts to complete the mapping [16,45,46]. The United Nation Land Cover Classification System (LCCS) was used. The maps contained two levels of land cover information—detailed, regionally optimized land cover classes (up to 44 classes) [43] for each continent and a less thematically detailed global classification scheme that harmonizes regional classes into one consistent product (22 classes) [41,46]. Top-of-canopy reflectance values were calculated using the water vapor, ozone climatology and aerosols generated by simple statistical modeling. Implementation of classification and image post processing methods were performed independently by local lead scientists. Regional monthly and seasonal mosaics were produced after implementing a wide range of statistical averaging techniques. An unsupervised classification (ISODATA) was applied and where necessary, ancillary data such as Defense Meteorological Satellite Program (DMSP) and Radar imagery were added to improve certain classes (e.g., urban and swamp forest) [13,42,44,47,48].

A two-stage stratified clustered sampling approach was used based on class priority and complexity of the landscape [48]. Quality control procedures and a quantitative accuracy assessment were implemented by the selection of random sample units interpreted from predominantly Landsat imagery or SPOT HRV images, where necessary [21]. A total of 1265 sample sites were interpreted for accuracy assessment. The results indicate an overall accuracy of 68.6% [48].

3.1.4. GlobCover 2009

The GlobCover 2009 land cover was created from 300 m MEdium-spectral Resolution Imaging Spectrometer (MERIS) imagery onboard the Environmental Satellite (ENVISAT) collected from January 2009 to December 2009 [32,49]. GlobCover 2009 was produced by the European Space Agency (ESA) and Université Catholique de Louvain (UCL) [30]. The project divided the world into 22 equal areas for analysis [14] and included two modules: pre-processing and classification [32]. The pre-processing module included a set of corrections such as cloud detection, atmospheric correction, geometric corrections and reclassification of the land and water classes using the MERIS level 1B Land/Ocean mask [32]. The classification module consisted of per-pixel supervised classification for the urban and wetland classes plus an unsupervised classification for remaining classes to create similar spectral and temporal clusters. The labelling of the clusters was performed based on the correspondence between spectral-temporal class and the reference land cover class. The global reference land cover classification scheme was based on LCCS with 22 global classes and up to 51 classes regionally [49,50]. Some post classification editing such as gap filling was performed to update the GlobCover 2009 using a reference land cover database [31].

The validation process included reference data collection, sampling strategy and accuracy assessment [50]. The GlobCover validation data set contained 2190 samples collected and labeled by

16 international experts based on LCCS Classification scheme. The samples were selected using a stratified random sampling. The area weighted overall accuracy was 67.5% [50].

3.2. Error Budget for Global Land Cover Datasets

Table 3 summarizes the error budget for the four global land cover datasets reviewed in this paper. Instead of creating four separate tables that would produce rather similar results, we chose to create on analysis that is the combination of the four mapping projects. If the producers of each of these maps conducted this error budgeting analysis, their results (tables) might differ more because of their additional insights into their procedures. However, even the combined analysis provides some interesting insights. The analysis consists of 8 major components and uses three evaluation criteria: error contribution, implementation difficulty and implementation priority. The results of the comparison of these four land cover datasets showed a higher error contribution potential due to the classification system and the accuracy assessment procedures; medium error contribution potential in the natural and input data procedures; and lower error contribution potential in the systematic, ancillary data, and preprocessing procedures. Implementation difficulty of each procedure/component was ranked from 1 (easy) to 5 (difficult). The accuracy assessment and classification system procedures have lower implementation difficulties while the errors in systematic and natural components are more difficult to correct. The components with lower error potential have higher implementation difficulty (e.g., systematic and natural) while those with higher error potential have lower implementation difficulty (e.g., accuracy assessment and classification system). Implementation priority was determined by the combination of the error potential contribution and implementation difficulty where the rank of 1 represents highest priority to be fixed in the future while 21 represents the lowest priority.

Table 3. Error Budget for the global land cover mapping databases.

N.		Error Contribution Potential	Implementation Difficulty	Implementation Priority
1	Systematic	Low	5	
1.1	<i>Spatial resolution</i>	<i>Low</i>	5	21
1.2	<i>Spectral resolution</i>	<i>Low</i>	5	20
2	Natural	Medium	4	
2.1	<i>Atmosphere</i>	<i>Medium</i>	4	19
3	Input data	Medium	2	
3.1	<i>Temporal NDVI</i>	<i>Medium</i>	3	18
3.2	<i>Spectral bands</i>	<i>Medium</i>	3	17
4	Ancillary data	Low	2	
4.1	<i>SAR</i>	<i>Low</i>	3	15
4.2	<i>Regional land cover maps</i>	<i>Medium</i>	2	16
4.3	<i>High resolution images</i>	<i>Low</i>	2	14

Table 3. Cont.

N.		Error Contribution Potential	Implementation Difficulty	Implementation Priority
5	Preprocessing	Low	2	
5.1	<i>Geometric correction</i>	<i>Low</i>	<i>2</i>	<i>11</i>
5.2	<i>Atmospheric correction</i>	<i>Low</i>	<i>2</i>	<i>10</i>
5.3	<i>Cloud mask</i>	<i>Low</i>	<i>1</i>	<i>9</i>
5.4	<i>water mask</i>	<i>Medium</i>	<i>1</i>	<i>13</i>
5.5	<i>Snow mask</i>	<i>Medium</i>	<i>1</i>	<i>12</i>
6	Classification system	High	3	
6.1	<i>Classification scheme</i>	<i>High</i>	<i>3</i>	<i>5</i>
6.2	<i>Training sites</i>	<i>High</i>	<i>2</i>	<i>3</i>
6.3	<i>Number of classes</i>	<i>Medium</i>	<i>3</i>	<i>7</i>
6.4	<i>Classification method</i>	<i>Medium</i>	<i>2</i>	<i>6</i>
7	Processing sequence	Medium	1	8
8	Accuracy assessment	High	1	
8.1	<i>Sampling scheme</i>	<i>High</i>	<i>1</i>	<i>1</i>
8.2	<i>Reference data</i>	<i>High</i>	<i>2</i>	<i>4</i>
8.3	<i>Interpreters' skill</i>	<i>High</i>	<i>1</i>	<i>2</i>

Note: Implementation difficulty: 1—easy; 5—difficult Implementation priority; 1—higher priority; 21—lower priority.

3.3. Summary of Classification Scheme for Global Land Cover Mapping Databases

Table 4 presents the various classification scheme definitions for the just the forest/trees and croplands categories from the four global land cover projects. Only these land cover classes were analyzed here since previous research showed the most inconsistencies occur in these classes. GLC 2000 and GlobCover 2009 used more classes in these categories than did IGBP and UMD. GLC 2000 and GlobCover 2009 followed the LCCS and the definitions between these projects are similar. IGBP and UMD followed the IGBP scheme and the class definitions between these projects are similar. However, the definitions between the two basic schemes (LCCS vs. IGBP) do not match well. Most of the inconsistencies are evident in the mixed class types. For example, each database has defined a mixed/mosaic class having a combination of other classes with no clear majority. However, the definitions do not clearly define the percentages of each class which leads to confusion about the spatial distribution and therefore, to inaccurate classification. This problem exists in both the forest/tree and cropland categories.

Table 4. Four global land cover legends with definitions for just the forest/trees and cropland categories.

IGBP	UMD	GLC 2000	GlobCover 2009
Name & Description	Name & Description	Name & Description	Name & Description
Trees			
Evergreen Broadleaf Forests (Height > 2 m, Canopy > 60%)	Evergreen Broadleaf Forests (Height > 5 m, Canopy > 60%)	Tree Cover, broadleaved, evergreen, closed to open (Height > 3–30 m, Canopy > 15%)	Broadleaved evergreen or semi-deciduous forest (>5 m) closed to open (Canopy > 15%)
Deciduous Broadleaf Forests (Height > 2m, Canopy > 60%)	Deciduous Broadleaf Forests (Height > 5 m, Canopy > 60%)	Tree Cover, broadleaved, deciduous, Closed (Height > 3–30 m, Canopy > 40%)	Broadleaved deciduous (>5 m) Closed (Canopy > 40%)
		Tree Cover, broadleaved, deciduous, open (Height > 3–30 m, Canopy 15%–40%)	Broadleaved deciduous forest/woodland (>5 m), open (Canopy 15%–40%)
Evergreen needle leaf Forests (Height > 2 m, Canopy > 60%)	Evergreen needle leaf Forests (Height > 5m, Canopy > 60%)	Tree Cover, needle leaved, evergreen, closed to open (Height > 3–30 m, Canopy > 15%)	Needle leaved evergreen forest (> 5 m) Closed (Canopy > 40%)
Deciduous needle leaf Forests (Height > 2 m, Canopy > 60%)	Deciduous Needle leaf Forests (Height > 5 m, Canopy > 60%)	Tree Cover, needle leaved, deciduous, closed to open (Height > 3–30 m, Canopy > 15%)	Needle leaved deciduous or evergreen forest (>5 m) Open (Canopy 15%–40%)
Mixtures or mosaics of the other four forest cover types with none of the forest types > 60% (Height > 2 m, Canopy > 60%)	Mixtures or mosaics of needle leaf and broadleaf with neither type has <25% or >75% trees (Height > 5 m, Canopy > 60%)	Tree Cover, mixed leaf type, closed to open (Height > 3–30 m, Canopy > 15%)	Closed to open (Canopy > 15%) mixed broadleaved and needle leaved forest (>5 m)
		Tree Cover, closed to open (Height > 3–30 m, Canopy > 15%) regularly flooded, fresh or brackish water: Swamp Forests	Closed to open (Canopy > 15%) broadleaved forest regularly flooded (semi-permanently or temporarily)—Fresh or brackish water
		Tree Cover, closed to open (Height > 3–30 m, Canopy > 15%), regularly flooded, saline water: Mangrove forests	Closed (Canopy > 40%) broadleaved forest or shrub land permanently flooded—Saline or brackish water
		Tree Cover, burnt (mainly boreal forests)	

Table 4. Cont.

IGBP	UMD	GLC 2000	GlobCover 2009
Name & Description	Name & Description	Name & Description	Name & Description
Woody savannas Herbaceous and other understory systems (Height > 2 m, Canopy 30%–60%)	Woodlands Herbaceous or woody understories and tree, evergreen or deciduous (Height > 5 m, canopy 40%–60%)	Mosaic: Tree Cover/other natural vegetation (crop component possible) (Height > 3–30 m, Canopy > 60%–70%)	Mosaic grassland (50%–70%)/forest or shrub land (20%–50%)
Savannas Herbaceous and other understory systems (Height > 2 m, Canopy 10%–30%)	Wooded Grasslands/Shrub land Herbaceous or woody understories, evergreen or deciduous (Height > 5 m, canopy 10%–40%)		
Cropland			
Croplands: temporary crops followed by harvest and a bare soil period.	Crop producing > 80%,	Cultivated and managed areas (upland crops or inundated/flooded crops as, e.g., rice)	Post-flooding or irrigated croplands (or aquatic)
		Mosaic: cropland/tree cover/other natural vegetation	Rainfed croplands
Mosaic of croplands, forest, shrub lands, and grasslands, no component > 60%		Mosaic: cropland/shrub or grass cover	Mosaic cropland (50%–70%)/vegetation (20%–50%)
			Mosaic vegetation (50%–70%)/cropland (20%–50%)

Table 5 presents the eight attributes or discriminators used in the classification process and a comparison of the associated thresholds used between the LCCS and the four global land cover datasets. This comparison provides a measure of “mappability” or ability of the imagery to discern these thresholds. The more recent global land cover datasets (GLC 2000 and GlobCover 2009) incorporated water seasonality, water quality and water supply in their classification scheme while IGBP and UMD did not. Many inconsistencies are evident among the thresholds for tree height, canopy cover and spatial distribution while the leaf type and leaf phenology appear to be more consistent. The definitions of mixed class types which are based on spatial distribution are not clear. Not all attributes have the same “mappability”. Tree height, spatial distribution and water quality were most difficult thresholds to be detected by the coarse resolution images used in these mapping projects.

Table 5. Thresholds for the eight attributes/discriminators used in the four global land cover databases and the LCCS definition with their resulting mappability for just the forest/trees and croplands categories.

No.	Attributes	LCCS Definition	GLC 2000	GlobCover 2009	IGBP	UMD	Mappability
Basic Scheme			LCCS	LCCS	IGBP	IGBP	
Tree							
1	Height	>3 m–30 m	>3 m–30 m	>5 m	>2 m	>5 m	Low
2	Canopy Cover	Open (60%–70% to 20%–10%)	15%–40%	15%–40%	>60%	>60%	High
		Closed (>60%–70%)	>40%	>40%			
		Sparse (20%–10% to 1%)	×	×			
		Closed to open (15% to 100%)	>15%	>15%			
3	Leaf type	Broadleaf	Broadleaf	Broadleaf	Broadleaf	Broadleaf	Medium
		Needleleaf	Needleleaf	Needleleaf	Needleleaf	Needleleaf	
		Mixed	Mixed	Mixed	Mixed	Mixed	
4	Leaf phonology	Evergreen	Evergreen	Evergreen	Evergreen	Evergreen	High
		Deciduous	Deciduous	Deciduous	Deciduous	Deciduous	
5	Spatial distribution (Macropattern)	Continuous (>80%)	unknown	unknown	unknown	unknown	Low
		Fragmented (20%–80%)		20%–50%			
		Parklike Patches		×			
6	Water seasonality	Permanent	×	×	×	×	High
		Temporary	Regular	Regular			
		Waterlogged	×	×			
7	Water quality	Fresh water (<1000 ppm)	Fresh or brackish	Fresh or brackish	×	×	Medium
		Brackish water (1000 ppm–10,000 ppm)					
		Saline water (>10,000 ppm)		Saline			
Crop							
8	Water supply	Rainfed cultivation	Rainfed	×	×	×	Medium
		Post-flooding Cultivation	Post-flooded or irrigated				
		Irrigated					

Unknown means the value is not clearly defined in the database while × represents that this information is not employed by the database.

4. Discussion and Recommendations

4.1. Analysis of the Characteristics of Global Land Cover Mapping

The in depth review and analysis of the four global land cover mapping projects was quite revealing. These projects occurred as a progression in time and it is clear that multiple interested parties participated in more than one of these. It is logical to conclude that previous projects were reviewed before a new project was begun to, at the very least, see what previous researchers had done and, optimistically, use a similar approach (especially the classification scheme) so that the projects could be directly comparable. However, it is strikingly obvious that less coordination between projects actually occurred. Even a quick look at Table 2 shows many more differences than similarities.

4.2. Analysis of the Error Budget

The error budget analysis of the global land cover mapping projects showed that the validation was a crucial issue with the highest implementation priority. The four global land cover products employed different validation methods which resulted in conflicting interpretations and conclusions [51]. UMD did not conduct a statistically rigorous accuracy assessment for their product although a few researchers have compared it with existing regional datasets created from high-resolution imagery [11,18]. The failure to achieve a high accuracy is an indication of usefulness of the land cover information [52]. Global Land Cover 2000, GlobCover 2009 and IGBP completed their accuracy assessments using an independent statistical validation [48], but the majority of the reference sites were collected from medium resolution (e.g., Landsat TM). The resulting assessment is then limited by the quality, availability, and interpreter's skills in labeling this reference data. The sampling design also varied for each project and the choice of sampling design influenced the reliability of the accuracy assessment [3,53,54]. GlobCover 2009 was validated with data from 2008 [49] and multiple areas of the world were either not sampled or were poorly sampled. All these issues point to a need for a requirement of a general framework of accuracy assessment for global land cover mapping in the future including adopting a common sampling method, effective reference data collection, and standardized reporting of accuracy measures.

In addition to the accuracy assessment methods, the classification system also has a high implementation priority. The lack of consistent training data, classification scheme, number of classes and classification algorithm give rise to spatial disagreement among the mapping products and difficulty comparing them with each other. While it is more understood how to select the size and quality of training data on a regional basis [55,56], it is certain that these issues are less understood globally and have contributed uncertainty in the final land cover maps. GlobCover 2009 and GLC 2000 adopted the LCCS classification scheme while UMD and IGBP were based on IGBP classification scheme. The lack of consistent classification labels and especially definitions resulted in difficulty with interoperability and compatibility. Although translation based on LCCS is possible, some classes such as mixed classes are hard to crosswalk between schemes. These mixed classes and classes that use a variety of discriminators to label are subject to increased error. IGBP and GLC 2000 also collapsed the classes generated at the regional level to produce the global map. Different regions adopted different numbers of classes. The class definitions were also different for the various local/regional areas which then contribute to certain amounts of

uncertainty when combining on a global scale. A major source of classification error is associated with the allocation of similar land cover types to different classes. There is a certain amount of uncertainty in translating the classification legends/schemes which should be resolved based on a standard class definition system. Future global land cover mapping projects should strictly follow the available standardized classification scheme (e.g., LCCS) and all classes of interest must be defined clearly [24].

The other components of global land cover mapping project including systematic, natural, input data, ancillary data and preprocessing have relatively lower implementation priority because the processing methods are standardized and most of the errors such as geometric errors are unavoidable and quite difficult to improve.

4.3. Analysis of Classification Scheme

The comparison of the classification scheme showed that there are many inconsistencies among the four datasets, especially between the datasets following different basic schemes. This implied that LCCS is very different from the IGBP scheme. Future global land cover mapping should consider which scheme to follow and not change the thresholds defined in the scheme because despite UMD and IGBP using the same basic scheme, minor differences (e.g., tree height) existed that resulted in major differences between GLC2000 and GlobCover 2009.

Major differences occurred between the thresholds for the attributes/discriminators used in the classification process for spatial distribution, water seasonality, water quality and water supply while minor differences existed in the tree height, canopy cover, leaf type and leaf phenology. Spatial distribution is the underlying attribute used to define the mixed classes (e.g., the mosaic class in GLC2000 is defined as tree cover mixed with other natural vegetation). However none of the classification schemes have clearly provided the percent information for the spatial distribution of these mixes. Even the GLC 2000 and GlobCover 2009, which have followed the LCCS standard, use different thresholds for each attribute because the LCCS specifies a range of thresholds instead of a single one. These inconsistencies foster the risk of misunderstanding by the map producers and raise the uncertainties of the map products, especially for those products (e.g., GLC2000) which are created regionally by local researchers and then joined together [16,45,46]. Every researcher can then potentially have their own definitions of mixed classes because a single definition is not clearly presented. The ambiguity of these definitions also reduces interoperability and compatibility of these products, because translation of the mixed classes is extremely difficult [16,17]. The definition of mixed type classes should be of great consideration because previous research has proven that mixed pixels greatly decrease the accuracy of the land cover products [22,57]. Compared to IGBP and UMD, GLC2000 and GlobCover 2009 employed water seasonality, water quality and water supply in the definition of trees and crops. This implies that as higher resolution images and more advanced remote sensing technology become available, recent global land cover projects are attempting to separate the general classes into more detail to meet increasing needs from the land cover information.

We also found that the four datasets adopted some attributes/discriminators to define the class types, which are of greater difficulty to be identified by the spatially coarse resolution images including height, water quality and spatial distribution. Height and spatial distribution represent the physical characteristics of the land cover but are hard to be interpreted from the imagery itself. Height is used to discriminate

trees from the herbaceous vegetation. Some new technologies such as Lidar [58–60] could be used to determine height, but these data are very expensive to obtain globally at sufficient resolution to be useful. Spatial distribution is limited by the spatial resolution of the images and by the classification method. Most of the global land cover classification methods are pixel-based which eliminates any knowledge about the spatial distribution of information within the pixel. Our review clearly shows that future global land cover mapping project should consider the mappability of these attributes/discriminators in the classification process and that a compromise between the costs of new technologies and the goals of the mapping project is warranted.

4.4. Lessons Learned

There are quite a few important lessons that were learned from the examination of these four global mapping projects. They include:

(1) The classification scheme must be carefully chosen at the beginning of the project. If the maps are to be compared to previous projects, then the scheme must match exactly the previous map. Use of a crosswalk to reconcile differences between schemes is not usually effective. The scheme must have not only map labels, but also clear and concise definitions of each of the map classes.

The classification scheme should be appropriate for use with remotely sensed imagery if the project involves such imagery to create the land cover map. Using a scheme that relies on information that must be collected on the ground that is smaller than the spatial resolution of the imagery dooms the project to failure. Every effort must be taken to insure that the scheme is appropriate for the resolution of the imagery used in the project.

(2) The use of ancillary data can seriously improve the accuracy of the map. However, issues arise when the ancillary data (or quality of data) are not uniformly available globally.

(3) Many global mapping projects are done by region. Again, care must be taken to make sure that the quality of the final global map is consistent and uniform. Coordination between regions must be vigilant in order for this to occur. It may not be necessary to employ the same classification algorithm for each region, but it is critical that the same, uniform classification scheme be used.

(4) Accuracy assessment has become a widely accepted component of every mapping project. Global maps offer unique challenges, but an efficient, practical, and statistically valid assessment must be designed early in the project in cooperation with all the participants.

(5) The entire mapping process must be well documented and transparent. Details must be recorded and available. Transparency aids greatly in the comparison of the map with other maps.

5. Conclusions

This paper presents the results of an in-depth review and error budgeting analysis of four global land cover maps. The review was conducted in order to summarize lessons learned from past global mapping projects to potentially improve future mapping exercises. The primary lesson learned was the importance of a consistent and well-defined classification scheme. In addition, the need for an efficient, yet statistically valid plan, for assessing the map accuracy was also discovered to be very valuable. Some important work has begun in this area [61]. The Global Observation of Forest Cover and Land Dynamics

(GOFC-GOLD) project [61] and other work being performed by USGS, NASA, and other groups is a great start at building global reference data sets and procedures. The lessons learned from reviewing the global mapping projects described in this paper will aid these efforts to make sure they are effective and usable.

An error budget was performed using an uncertainty analysis that showed which components of a mapping project were most subject to error and which could be most easily improved. Finally, a detailed analysis of issues with the classification schemes between maps demonstrated the need for consistency and highlighted the impacts when varying the attributes/discriminators used in mapping the forest/trees and crop land cover types. Careful consideration of the issues and analysis described in this paper will result in improved global land cover mapping in the future.

Acknowledgments

The authors would like to thank NASA Making Earth Science Data Records for Use in Research Environments (MEaSUREs) for funding (Grant Number: NNH13AV82I) this research. U.S. Geological Survey (USGS), and in particular USGS Western Geographic Science Center (WGSC) provided support and facility for the work in numerous different ways (administrative, technical, management). We are grateful for this support. Finally, we would like to thank the four anonymous reviewers that helped improve this paper with their comments and suggestions.

Author Contributions

Russell G. Congalton, Prasad Thenkabail, and Mutlu Ozdogan conceived the idea for this paper. Preliminary work was done by these three authors. Russell G. Congalton then extended the work and involved Jianyu Gu and Kamini Yadav in the compilation of data and the data analysis. Tables and figures that resulted from the analysis were generated by Jianyu Gu and Kamini Yadav along with the first draft of the writing. The final paper was written by Russell G. Congalton and then sent to all authors for comments and edits. Russell G. Congalton, Jianyu Gu, and Kamini Yadav compiled all the edits and produced the final paper. Jianyu Gu converted the paper to the final format for this journal.

Conflicts of Interest

The authors declare no conflict of interest.

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.
2. Sellers, P.J.; Dickinson, R.E.; Randall, D.A.; Betts, A.K.; Hall, F.G.; Berry, J.A.; Collatz, G.J.; Denning, A.S.; Mooney, H.A.; Nobre, C.A.; *et al.* Modeling the exchanges of energy, water, and carbon between continents and the atmosphere. *Science* **1997**, *275*, 502–509.
3. Stehman, S.V. Basic probability sampling designs for thematic map accuracy assessment. *Int. J. Remote Sens.* **1999**, *20*, 2423–2441.

4. McRoberts, R.E. Probability- and model-based approaches to inference for proportion forest using satellite imagery as ancillary data. *Remote Sens. Environ.* **2010**, *114*, 1017–1025.
5. DeFries, R.; Hansen, M.; Townshend, J. Global discrimination of land cover types from metrics derived from AVHRR pathfinder data. *Remote Sens. Environ.* **1995**, *54*, 209–222.
6. ORNL DAAC Distributed Active Center for Biogeochemical Dynamics. Available online: http://webmap.ornl.gov/wcsdown/dataset.jsp?ds_id=10004 (accessed on 5 August 2014).
7. Stone, T.A.; Schlesinger, P.; Houghton, R.A.; Woodwell, G.M. A map of the vegetation of South-America based on satellite imagery. *Photogramm. Eng. Remote Sens.* **1994**, *60*, 541–551.
8. Loveland, T.; Reed, B.; Brown, J.; Ohlen, D.; Zhu, Z.; Yang, L.; Merchant, J. Development of a global land cover characteristics database and IGBP discover from 1 km AVHRR data. *Int. J. Remote Sens.* **2000**, *21*, 1303–1330.
9. Loveland, T.R.; Merchant, J.W.; Brown, J.F.; Ohlen, D.O.; Reed, B.C.; Olson, P.; Hutchinson, J. Map supplement: Seasonal land-cover regions of the United States. *Ann. Assoc. Am. Geogr.* **1995**, *85*, 339–355.
10. Global Land Cover Facility (UMD Land Cover Facility). Available online: <http://glcf.umd.edu/data/landcover/> (accessed on 6 August 2014).
11. 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.
12. Joint Research Center Land Resource Management Unit Global Land Cover 2000-Products. Available online: <http://bioval.jrc.ec.europa.eu/products/glc2000/products.php> (accessed on 6 August 2014).
13. 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.
14. Esa Data User Element (GlobCover). Available online: <http://due.esrin.esa.int/globcover/> (accessed on 6 August 2014).
15. Sophie, B.; Pierre, D. *GlobCover 2009 Products Description and Validation Report*; European Space Agency: Paris, France, 2010.
16. Herold, M.; Mayaux, P.; Woodcock, C.E.; Baccini, A.; Schmullius, C. Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets. *Remote Sens. Environ.* **2008**, *112*, 2538–2556.
17. Giri, C.; Zhu, Z.; Reed, B. A comparative analysis of the Global Land Cover 2000 and MODIS land cover data sets. *Remote Sens. Environ.* **2005**, *94*, 123–132.
18. Hansen, M.C.; Reed, B. A comparison of the IGBP discover and university of Maryland 1 km global land cover products. *Int. J. Remote Sens.* **2000**, *21*, 1365–1373.
19. Pérez-Hoyos, A.; García-Haro, F.J.; San-Miguel-Ayanz, J. Conventional and fuzzy comparisons of large scale land cover products: Application to CORINE, GLC2000, MODIS and GlobCover in Europe. *ISPRS J. Photogramm. Remote Sens.* **2012**, *74*, 185–201.
20. Di Gregorio, A.; Jansen, L.J. *Land Cover Classification System (LCCS): Classification Concepts and User Manual*; Food and Agriculture Organization of the United Nations: Rome, Italy, 2000.
21. McCallum, I.; Obersteiner, M.; Nilsson, S.; Shvidenko, A. A spatial comparison of four satellite derived 1 km global land cover datasets. *Int. J. Appl. Earth Obs. Geoinf.* **2006**, *8*, 246–255.

22. Latifovic, R.; Olthof, I. Accuracy assessment using sub-pixel fractional error matrices of global land cover products derived from satellite data. *Remote Sens. Environ.* **2004**, *90*, 153–165.
23. Santayana, G. *The Life of Reason: Or the Phases of Human Progress, Volume 1, Reason in Common Sense*; Dover: England, UK, 1905.
24. Congalton, R.G.; Green, K. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*; CRC Press: Boca Raton, FL, USA, 2008.
25. Congalton, R.; Brennan, M. Error in remotely sensed data analysis: Evaluation and reduction. In Proceedings of the Sixty Fifth Annual Meeting of the American Society of Photogrammetry and Remote Sensing, Portland, OR, USA, 17–21 May 1999; Volume 729.
26. Loveland, T.R.; Zhu, Z.L.; Ohlen, D.O.; Brown, J.F.; Reed, B.C.; Yang, L.M. An analysis of the IGBP global land-cover characterization process. *Photogramm. Eng. Remote Sens.* **1999**, *65*, 1021–1032.
27. Brown, J.F.; Loveland, T.R.; Ohlen, D.O.; Zhu, Z.L. The global land-cover characteristics database: The users' perspective. *Photogramm. Eng. Remote Sens.* **1999**, *65*, 1069–1074.
28. Loveland, T.R.; Belward, A.S. The IGBP-DIS global 1 km land cover data set, discover: First results. *Int. J. Remote Sens.* **1997**, *18*, 3289–3295.
29. Belward, A.S.; Estes, J.E.; Kline, K.D. The IGBP-DIS global 1-km land-cover data set discover: A project overview. *Photogramm. Eng. Remote Sens.* **1999**, *65*, 1013–1020.
30. Arino, O.; Ramos Perez, J.; Kalogirou, V.; Bontemps, S.; Defourny, P.; van Bogaert, E. *Global Land Cover Map for 2009 (GlobCover 2009)*; European Space Agency (ESA) & Université Catholique de Louvain (UCL): Frascati, Italy, 2012.
31. Bontemps, S.; Defourny, P.; van Bogaert, E.; Olivier, A.; Vasileios, K. *GlobCover 2009—Product Description Manual*; ESA: Paris, France, 2009; p.17.
32. Bontemps, S.; Defourny, P.; van Bogaert, E.; Arino, O.; Kalogirou, V.; Ramos-Perez, J. *GlobCover 2009—Products Description and Validation Report*; ESA: Paris, France, 2011.
33. Fritz, S.; See, L. Identifying and quantifying uncertainty and spatial disagreement in the comparison of global land cover for different applications. *Glob. Chang. Biol.* **2008**, *14*, 1057–1075.
34. Global Land Cover Characterization (USGS). Available online: <http://edc2.usgs.gov/glcc/glcc.php> (accessed on 7 August 2014).
35. Danko, D.M. The Digital Chart of the World project. *Photogramm. Eng. Remote Sens.* **1992**, *58*, 1125–1128.
36. Tucker, C.J.; Justice, C.O.; Prince, S.D. Monitoring the grasslands of the Sahel 1984–1985. *Int. J. Remote Sens.* **1986**, *7*, 1571–1581.
37. Scepan, J.; Menz, G.; Hansen, M.C. The discover validation image interpretation process. *Photogramm. Eng. Remote Sens.* **1999**, *65*, 1075–1081.
38. Townshend, J.R.G.; Justice, C.O.; Skole, D.; Malingreau, J.P.; Cihlar, J.; Teillet, P.; Sadowski, F.; Ruttenberg, S. The 1 km resolution global data set: Needs of the international geosphere biosphere programme. *Int. J. Remote Sens.* **1994**, *15*, 3417–3441.
39. Eidenshink, J.C.; Faundeen, J.L. The 1km AVHRR global land data set—1st stages in implementation. *Int. J. Remote Sens.* **1994**, *15*, 3443–3462.

40. De Fries, R.S.; Hansen, M.; Townshend, J.R.G.; Sohlberg, R. Global land cover classifications at 8 km spatial resolution: The use of training data derived from Landsat imagery in decision tree classifiers. *Int. J. Remote Sens.* **1998**, *19*, 3141–3168.
41. Pflugmacher, D.; Krankina, O.N.; Cohen, W.B.; Friedl, M.A.; Sulla-Menashe, D.; Kennedy, R.E.; Nelson, P.; Loboda, T.V.; Kuemmerle, T.; Dyukarev, E.; *et al.* Comparison and assessment of coarse resolution land cover maps for northern Eurasia. *Remote Sens. Environ.* **2011**, *115*, 3539–3553.
42. Mayaux, P.; Bartholomé, E.; Fritz, S.; Belward, A. A new land-cover map of Africa for the year 2000. *J. Biogeog.* **2004**, *31*, 861–877.
43. Fritz, S.; See, L.; Rembold, F. Comparison of global and regional land cover maps with statistical information for the agricultural domain in Africa. *Int. J. Remote Sens.* **2010**, *31*, 2237–2256.
44. Bartalev, S.A.; Belward, A.S.; Erchov, D.V.; Isaev, A.S. A new Spot4-vegetation derived land cover map of northern Eurasia. *Int. J. Remote Sens.* **2003**, *24*, 1977–1982.
45. Jung, M.; Henkel, K.; Herold, M.; Churkina, G. Exploiting synergies of global land cover products for carbon cycle modeling. *Remote Sens. Environ.* **2006**, *101*, 534–553.
46. Fritz, S.; Bartholomé, E.; Belward, A.; Hartley, A.; Stibig, H.; Eva, H.; Mayaux, P.; Bartalev, S.; Latifovic, R.; Kolmert, S. *Harmonisation, Mosaicing and Production of the Global Land Cover 2000 Database (Beta Version)*; Office for Official Publications of the European Communities: Luxembourg, 2003.
47. Eva, H.D.; Belward, A.S.; de Miranda, E.E.; di Bella, C.M.; Gond, V.; Huber, O.; Jones, S.; Sgrenzaroli, M.; Fritz, S. A land cover map of South America. *Glob. Chang. Biol.* **2004**, *10*, 731–744.
48. Mayaux, P.; Eva, H.; Gallego, J.; Strahler, A.H.; Herold, M.; Agrawal, S.; Naumov, S.; de Miranda, E.E.; di Bella, C.M.; Ordoyne, C.; *et al.* Validation of the global land cover 2000 map. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 1728–1739.
49. Bicheron, P.; Defourny, P.; Brockmann, C.; Schouten, L.; Vancutsem, C.; Huc, M.; Bontemps, S.; Leroy, M.; Archard, F.; Herold, M. *GlobCover: Products Description and Validation Report* [Online]; Medias France: Toulouse, France, 2008. Available online: <http://publications.jrc.ec.europa.eu/repository/handle/JRC49240> (accessed on 8 August 2014).
50. Defourny, P.; Bontemps, S.; Obsomer, V.; Schouten, L.; Bartalev, S.; Herold, M.; Bicheron, P.; Bogaert, E.V.; Leroy, M.; Arino, O. Accuracy assessment of global land cover maps: Lessons learnt from the GlobCover and GlobCorine Experiences. In Proceedings of 2010 European Space Agency Living Planet Symposium, Bergen, Norway, 28 June–2 July 2010.
51. Stehman, S.V. Selecting and interpreting measures of thematic classification accuracy. *Remote Sens. Environ.* **1997**, *62*, 77–89.
52. McRoberts, R.E. Satellite image-based maps: Scientific inference or pretty pictures? *Remote Sens. Environ.* **2011**, *115*, 715–724.
53. Muller, S.; Walker, D.; Nelson, F.; Auerback, N.; Bockheim, J.; Guyer, S.; Sherba, D. Accuracy assessment of a land-cover map of the Kuparuk River basin, Alaska: Considerations for remote regions. *Photogramm. Eng. Remote Sens.* **1998**, *64*, 619–628.
54. Stehman, S.V. Statistical rigor and practical utility in thematic map accuracy assessment. *Photogramm. Eng. Remote Sens.* **2001**, *67*, 727–734.

55. Kavzoglu, T. Increasing the accuracy of neural network classification using refined training data. *Environ. Model. Softw.* **2009**, *24*, 850–858.
56. Shao, Y.; Lunetta, R.S. Sub-pixel mapping of tree canopy, impervious surfaces, and cropland in the Laurentian Great Lakes basin using MODIS time-series data. *IEEE J-Stars* **2011**, *4*, 336–347.
57. Smith, J.H.; Stehman, S.V.; Wickham, J.D.; Yang, L.M. Effects of landscape characteristics on land-cover class accuracy. *Remote Sens. Environ.* **2003**, *84*, 342–349.
58. Dubayah, R.; Knox, R.; Hofton, M.; Blair, J.B.; Drake, J. Land surface characterization using lidar remote sensing. In *Spatial Information for Land Use Management*; Hill, M.J., Aspinall, R.J., Eds.; CRC Press: New York, NY, USA, 2000; pp. 25–38.
59. Yamamoto, K.; Takahashi, T.; Miyachi, Y.; Kondo, N.; Morita, S.; Nakao, M.; Shibayama, T.; Takaichi, Y.; Tsuzuku, M.; Murate, N. Estimation of mean tree height using small-footprint airborne Lidar without a digital terrain model. *J. For. Res.* **2011**, *16*, 425–431.
60. Andersen, H.-E.; McGaughey, R.J.; Reutebuch, S.E. Estimating forest canopy fuel parameters using lidar data. *Remote Sens. Environ.* **2005**, *94*, 441–449.
61. Olofsson, P.; Foody, G.; Herold, M.; Stehman, S.V.; Woodcock, C.E.; Wulder, M. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* **2014**, *148*, 42–57.

© 2014 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 license (<http://creativecommons.org/licenses/by/4.0/>).