Journal of Scientific Research & Reports
3(3): 457-478, 2014; Article No. JSRR.2014.007

SCIENCEDOMAIN ***international***
www.sciencedomain.org

Assessment of a Global Land Cover Classification Allocated Across the Landscape of Georgia (USA)

Pete Bettinger[1*], Krista Merry[1] and Zennure Ucar[1]

[1]School of Forestry and Natural Resources, University of Georgia, Athens, GA USA 30602, USA.

Authors' contributions

Author PB designed the study, performed the statistical analysis, assisted in writing the protocol, and wrote the first draft of the manuscript. Author KM assisted in the development of the protocol, the analysis of results, and the writing of the manuscript. Author ZU assisted in the analysis of results, and the writing of the manuscript. All authors read and approved the final manuscript

Original Research Article

Received 26th September 2013
Accepted 26th November 2013
Published 13th December 2013

ABSTRACT

Aims: To assess the agreement of a global land cover map to reference imagery when applied to a region (state) of the southern United States and to determine whether different sampling designs or the use of broader land class definitions can overcome problems associated with the inherent heterogeneity of land use in the region.
Study Design: We assessed the agreement of the Glob Cover 2009 global, medium resolution land cover assignments within the State of Georgia to USDA NAIP reference imagery. We performed the assessment using two statistically random sampling methods: pixel-based and block-based sampling. We then grouped some land classes according to *possibilities of agreement* relationships expressed by others, and assessed the agreement using these systems.
Place and Duration of Study: State of Georgia (USA). Imagery and reference data acquired in 2009.
Methodology: Sample: We examined 3,930 sample pixels or pixel blocks from 16 land cover classes. Each sample was allocated a land class in the GlobCover 2009 database. Each sample was interpreted as a land class through photo-interpretation of USDA NAIP

**Corresponding author: E-mail: pbettinger@warnell.uga.edu;*

imagery. An omission-commission matrix was developed from the relationship between land cover map and reference interpretation, as was an estimated population matrix. Statistics regarding agreement were developed using the latter matrix.

Results: Overall agreement for the state of Georgia was approximately 48% using both pixel- and block-based assessments. Agreement increased with the implementation of the *possibilities of agreement* relationships for both pixel- and block-based assessments. Three forested land cover types, representing about 78% of the Glob Cover land classes in Georgia, had agreement levels between 60 and 97% when *possibilities of agreement* were employed.

Conclusion: The use of the Glob Cover 2009 land cover classification may be well suited for broad, regional analysis and assessment of land cover trends. Moderate levels of classification agreement for important resources (forested areas) were estimated within the State of Georgia.

Keywords: Satellite imagery; agreement; stratified random sampling; producer's accuracy; user's accuracy.

1. INTRODUCTION

One of the most time-consuming aspects of broad-scale landscape analysis and modeling studies is the development of geographic data and associated information. Remotely sensed spectral properties of features on the surface of the Earth are widely used in land cover classification and categorization processes [1,2,3]. Land cover maps are one type of thematic map for representing land-based resources as geographic information, and for describing areas in an aggregated, socially constructed manner [4]. Global land cover databases are geographic information that attempt to characterize land features on a global scale and in a manner that could conceptually have a broad meaning [4]. For about two decades, advances have been made in the development of global land cover databases from remotely sensed imagery [5], with several global land cover databases available at no cost to the public [6].

One of the latest global databases is the GlobCover 2009 global land cover map developed in part by the European Space Agency [7].The GlobCover global land-cover map was developed from medium spatial resolution (roughly 300 m) remotely-sensed data collected by the Medium Resolution Imaging Spectrometer (MERIS) sensor that was installed within the ENVISAT satellite [8]. The GlobCover map was developed with an international audience in mind, therefore the land cover types are compatible with the Food and Agriculture Organization of the United Nations Land Cover Classification System [8]. While these types of products may be inadequate for detailed mapping of landscape features [9], they may be suitable for describing broad-scale landscape patterns and for conducting environmental analyses. For example, a portion of the GlobCover map covering China has been used to assess the potential scope of biomass production [10]. However, a map such as this is no more than an estimate of land use until it is compared to reference data [11].

Due to their broad coverage, global land cover maps can facilitate a number of types of landscape analysis, yet the applicability of the data to a particular problem depends on the quality and resolution of the data. It has been recognized that the methods employed in developing global land cover databases may not always be able to recognize and acknowledge special characteristics of different regions [12]. Because of this, Jung et al. [13]

suggested that some global land cover maps are not necessarily suited for certain uses, for example, as input for parameterizing carbon cycle models. Irrespective of the scope or extent of the proposed application, the agreement of global land cover maps to reference data should be evaluated prior to their use [6].

Assessments of agreement are often used to describe map quality pertaining to an entire set of land cover classes or specific sub-sets of land cover classes [1]. A common approach to the evaluation of the quality of thematic maps is to compare them to reference data within the confines of time, cost, and access restrictions on the analysis [14]. Rigorous assessments of the quality of global land cover maps have resulted in overall agreement levels ranging from about 66 to 80%, depending on the method employed and data assessed [15]. Arino et al. [16] assessed the GlobCover map using 2186 randomly located sample points and expert interpretation of the reference data. The overall agreement for what they considered the principal classes (cultivated and managed terrestrial land, natural and semi-natural terrestrial vegetation, natural and semi-natural aquatic vegetation, artificial surfaces, bare areas, and water, snow, and ice) was estimated to be 77.9%. Further, Yiming et al. [6] assessed the quality of the GlobCover map using 243 permanently-established research areas (e.g., FLUXNET meteorological tower sites), and estimated overall agreement to be around 65%. Portions of the GlobCover map have also been assessed for agreement at smaller scales: Song et al. [17] suggested that the GlobCover map overestimates forest cover in North America and Pérez-Hoyos et al. [18] suggested that the 2005 version of the GlobCover map underestimated agricultural areas in Europe. For three provinces in China, Li et al. [19] assessed the agreement of cropland represented in the GlobCover map and cropland identified in a Chinese national land cover database, and found regional differences due in part to heterogeneity of land uses across the landscape.

In this study, we performed an assessment of ability of the GlobCover map to describe land cover features within, and land uses applied to, the State of Georgia, which is located in the southern United States. The goal of this work was to assess the ability of the global land cover map to accurately portray individual land cover types, through a sampling strategy that provided an unbiased estimate of thematic agreement. The specific objectives were to (a) assess the overall agreement of the GlobCover map with other reference imagery for its ability to portray the resources and land use classes of the State; (b) determine whether pixel- or block-based sampling units provide better agreement; (c) assess both producer's and user's accuracies of the GlobCover map with respect to the various land use classes in the State; and (d) quantify the changes in agreement that may occur if suggested *possibilities of agreement* among classes [7] are employed.

2. METHODOLOGY

2.1 GlobCover 2009 Database

Surface (land) reflectance data in 15 different electromagnetic bands was collected by the MERIS sensor [20], and was orthorectified and atmospherically corrected prior to the development of the GlobCover map. The geolocation error of the publicly available map was reported to be acceptable (77 m RMS) for a database of this scope [21]. The spatial resolution of the raster data was reported to be 1/360º [7], or about 300 m, depending on latitude. In Georgia, the average size of a GlobalCover raster grid cell (pixel) in the 1984 World Geodetic System (WGS 84) reference coordinate system is approximately 266 *x* 307 m (width and length) in the southern part of the State, and 254 *x* 308 m in the northern part of the State.

The GlobCover land classification system is hierarchical, follows a standardized classification approach, and is compatible with the GLC2000 global land cover classification system [22] which utilizes the Food and Agriculture Organization of the United Nations land cover classification system [8]. There are 21 land cover types represented in the GlobCover map; forested and agricultural classes dominant 16 of these classes found in the State of Georgia. The land cover map was produced by applying to cloud-free MERIS full resolution composites a regionally-tuned classification process [23]. Due to the medium spatial resolution (about 300 m), pixels can often represent land containing a mixture of cover types, which may contribute to land classification error [23]. Although the mission of the ENVISAT satellite ended in 2012, the exploitation of its archived data continues [24], and an assessment of this classification process can be of value to future endeavors.

2.2 Reference Data and Interpretation Protocol

Agreement between the global land cover map and the reference database is based on our assumption that the units are homogeneous within the reference data. In order to apply this assumption, our assignment rule indicated the dominant (primary) reference land cover class within each sample unit, either the category represented by a single GlobCover pixel or the category represented by a 3 x 3 pixel block. In this process, we determined a single, dominant land class for each sample unit area using the GlobCover criteria (described in [7]). While others (e.g., [20]) collected primary and secondary reference land cover classes for litigious samples, our interpretation of the GlobCover classes in relation to actual land uses in Georgia led to very few of these cases.

A pixel is a common assessment unit choice, and a block is another option; neither need to be equivalent in size to a minimum mapping unit of a land cover map [25]. The reference information was determined primarily from United States National Aerial Imagery Program (NAIP) imagery, which has a 1 m spatial resolution, secondarily from United States Geological Survey 1999 1-m color infrared orthophotograph quarter quadrangles, and finally from older NAIP imagery. The NAIP imagery were temporally consistent (to one year of data collection) with the date of collection of the MERIS data used to develop the GlobCover map, and with the latest imagery generally available through Google Earth. However, if recent land use activity were evident in the reference area within a sample unit, older NAIP imagery were used to estimate conditions prior to 2009. The few cases where this was necessary involved the final harvest (clear cutting) of some forested areas. However, other aerial imagery was also employed to assist in differentiating forest types where the reference interpretation was unclear (Fig. 1). These issues tended to occur in the northern portion of the state, and mostly on national forest land. Additionally, these areas also were ones that generally had continuous canopy cover and contained mostly natural forests (rather than plantations) where it was difficult to separate coniferous and deciduous forests in the true color NAIP imagery. In most of these cases, 1999 color infrared imagery captured during the winter season was used to help differentiate coniferous and deciduous trees. In nearly all of these cases, major changes in forest conditions (density, species composition) around the sample areas had not occurred even a decade after the images were captured.

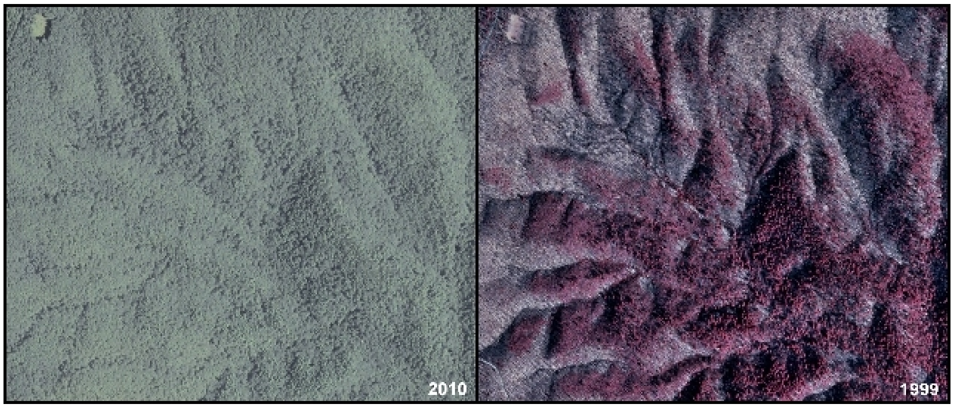

Fig. 1. Color aerial imagery (2010) and color infrared imagery (1999) for a forested area of the exact same location and extent in northern Georgia, USA.

Photo interpretation errors were minimized through the use of auxiliary (historical) imagery, yet they were also minimized by using a single photo interpreter having nearly 25 years experience using aerial images and teaching courses on aerial photogrammetry techniques. The use of a single interpreter is debatable, though texts (e.g., [26]) on the subject often allude to the use of teams, the training and auditing of teams, and the need for consistency among team members. Variation in the interpretation of land classes will certainly exist among teams members, and can exist within a single team member. The level of interpretation error among teams members may be uncertain, and when detected, corrective actions may need to be applied [26]. In Carrão et al. [20], four interpreters were used to independently confirm the interpretation of land cover types in the reference data, although this work does not conclude that four interpreters are more appropriate than a single, trained observer. Wickham et al. [27] used teams (prior experience unclear) who underwent training (to reduce individual subjectivity) and were guided by an experienced interpreter who overrode assignments when randomly checked samples were audited. While this division of labor may have been facilitated by regional knowledge of land by the teams, and may have provided work process efficiencies, no test was performed to declare the team approach to be more preferable than a single experienced individual. Zhu et al. [28] used two analysts to assess all sample areas and a third to mediate differences between if they occurred between the two assignments. This protocol was designed to minimize errors in photo interpretation. However, in all of these cases information was lacking concerning the experience and training of the interpreters, therefore drawing the conclusion that these designs are superior to the use of a single experienced professional is premature. Neither these previous works, nor the current work were designed to test this hypothesis.

Through careful attention to the distribution of land classes and forest types within the reference areas, a very high level of confidence concerning the interpretation of land classes from the reference data was realized. In most cases it was clear which land cover type in the reference data was the dominate (> 50% by area) type in each sample area. In cases where two land cover classes occupied large amounts of the sample area, careful measurement of area of each type, and careful interpretation of the GlobCover assignment rules determined which land cover type to assign. There were no doubtful reference points, and therefore none were removed from the analysis. However, we understood that locational error should be a concern around the boundaries of land cover classes [2,25]. Given the history of land use in Georgia, there were very few large areas of homogenous land cover, and therefore

GlobCover pixels generally contained areas with a great deal of edge (distinct differences between land cover classes). Through frequent verification of geospatial location of the validation data, misregistration was not evident. Other potential inconsistencies in the interpretation process were minimized though the use of a standard interpretation scale (1:4000). As suggested by Olofsson et al. [11], our guiding principle for the response design was to develop protocols that would be considered operationally practical and consistently implemented by the interpreter of the validation database.

2.3 Sampling Strategy

We developed a design-based assessment approach, using stratified random sampling, to arrive at statements of agreement for the land classes contained within the GlobCover map. The strata are the land cover classes of the GlobCover database that are found in the State of Georgia. Chen and Wei [14] suggest that there is no current consensus in the literature on the minimum sample size needed to adequately assess thematic map accuracy (agreement with reference imagery). The selection of sample method and sample size seems to depend on the differences in class proportion, spatial autocorrelation, and the type of agreement indices to be developed. Practical constraints often guide the choice of sampling unit [29] since agreement assessments can be rather expensive [11]. Further, opinions vary on the choice of spatial unit for the assessment of agreement [30,31]. As we alluded to earlier, our response design was developed as both (a) a pixel-based assessment of agreement and (b) a block-based assessment of agreement. The pixel-based approach is rather straightforward: a GlobCover pixel is selected at random within each stratum and compared to the reference land class covering the same area. Considering potential problems with thematic and positional agreement of the GlobCover map, we also developed a 3 x 3 block of pixels situated around a selected pixel, and compared the land cover type suggested by the center pixel to the dominant reference land class represented by the 9-cell area. However, when exploring the GlobCover map, the ability to place a 3 x 3 block of pixels totally (or mostly) within a single land cover type was difficult (in contrast to [20]) due to the "speckling" pattern of the land classes in the GlobCover map. Forcing this requirement would also have resulted in a biased sample that avoided areas of heterogeneity [26]. Part of the problem can be attributed to the inherent diversity of land cover classes spread across the landscape. Part of the problem can also be attributed to the fact that individual pixels were assumed independently assigned to a land cover class during the classification process, and often those that should have been considered similar to others nearby were classified differently, resulting in a speckling pattern [32]. We recognize that a cluster of mixed land classes can affect the outcome of the assessment of a single class, however the trade-off is to bias the assessment by only considering larger, homogeneously classed areas. Therefore, one sampling unit (spatial support unit) was defined as a single randomly selected GlobCover pixel, and the other a 3 x 3 block of pixels. The pixel-based sampling units and the center of each block were initially located as the pixels of the GlobCover map that intersected a random point, and as in Wickham et al. [33], and were then co-located on high-resolution imagery representing the reference condition on the ground.

A stratified random sampling method was used to select the sample units and to prepare the ground reference data. In general, this method allocates sample units to land cover classes based on the extent of the land cover classes found in the State of Georgia. Further, Lo and Watson [34] have suggested that this method works well in swampy coastal areas with both simple and complex spatial patterns of vegetation. However, if a sample is allocated in proportion to the area of each strata, some strata would have very few samples, and others would contain unnecessarily large samples [11]. With few samples located in areas that are small in proportion, the probability of obtaining an error-free sample or an disproportionately error-prone sample is high [35]. Chen and Wei [14] illustrated cases where minimum sample sizes were required for land cover classes, and Scepan [36] assumed a small minimum size (25 sample units) for an assessment of a global database. We decided that conservatively, 100 sample units was necessary to adequately sample the smaller land classes, therefore a minimum sample size for each land cover class was assumed (as in [37]). Congalton and Green [26] also suggest that a minimum of 100 samples are necessary when a large area is being assessed, or when there are a relatively large number of classes to assess. While we assumed a minimum number of sample points for the smaller land classes represented, we also assumed a variable number (1 point per 5,000 ha) for the larger land classes represented in Georgia (Table 1). We made this assumption in light of what others suggest - that additional samples beyond those reasonably needed to address agreement can add little to the analysis [38]. We viewed this assumption as one that provides a balance between the statistical rigor necessary and the time and cost of the assessment. Sample units were selected from the set of non-overlapping and spatially exhaustive units contained within each land class [31]. The total number of sample units selected with this process (3,930) is greater than the estimated number needed (873) for a worst-case scenario (as described in [26]) involving 16 classes, class proportions of 50%, and a desired precision of 5%. However, we felt that our chosen sample size was a practical compromise between time available and data necessary [39] to adequately represent each of the land classes across the expanse of the State.

While there is no general consensus on standard approaches to assessments of agreement [30,40], we provide the estimated population matrix and three measures of agreement as indicators of the classification agreement of the GlobCover 2009 map to USDA NAIP imagery. Agreement between the global land cover map and the reference data will be described by overall, producer's, and user's accuracies for both types of sampling units (a single randomly selected pixel, or a 3 *x* 3 block of pixels) in conjunction with an estimated population matrix. The estimated population matrix is derived from the omission-commission matrix, and illustrates the proportion of agreement for each land cover class and the proportion of area misclassified [31]. Overall agreement (percent of pixels correctly

classified) provides an indication of the probability that a randomly selected location within the map has been classified correctly [41]. Overall agreement [33] is computed as

$$\hat{O} = \left(\frac{1}{N}\right)\sum_{h=1}^{H} N_h \hat{p}_h \tag{1}$$

Here, N represents the total number of pixels in the State, h represents a land cover stratum, $\hat{p}_h$ represents the sample proportion of pixels in stratum h correctly classified, and N_h represents the number of pixels in the stratum. The result effectively represents an area-weighted overall agreement for the entire collection of strata. In order to develop user's accuracy measures for each land class, the following are defined using the omission-commission matrix values (from [26]):

$$n_{i+} = \sum_{j=1}^{k} n_{ij} \tag{2}$$

Table 1. Major GlobCover land classes in the State of Georgia, their total area, and the sample size used in the assessment

GlobCover class legend value	Class value	Area (1000 ha)	Sample size
Post-flooding or irrigated croplands	11	---	---
Rainfed cropland	14	51.6	100
Mosaic cropland (50-70%) / vegetation (20-50%)	20	245.0	100
Mosaic vegetation (50-70%) / cropland (20-50%)	30	744.0	149
Closed to open (>15%) broadleaved evergreen and/or Semi-deciduous forest	40	0.1	---
Closed (>40%) broadleaved deciduous forest	50	5,819.1	1,164
Open (15-40%) broadleaved deciduous forest	60	75.7	100
Closed (>40%) needleleaved coniferous forest	70	3,783.6	756
Open (15-40%) needleleaved coniferous forest	90	---	---
Closed to open (>15%) mixed forest	100	2,134.1	468
Mosaic forest/shrubland (50-70%) / grassland (20-50%)	110	157.6	100
Mosaic grassland (50-70%) / forest/shrubland (20-50%)	120	149.1	100
Closed to open (>15%) shrubland	130	15.4	100
Closed to open (>15%) grassland	140	1,464.9	293
Sparse (>15%) vegetation	150	0.1	---
Closed (>40%) broadleaved, regularly flooded, fresh water	160	12.0	100
Closed (>40%) broadleaved, semi-deciduous or evergreen, regularly flooded, saline water	140	161.5	100
Closed to open (>15%) vegetation on regularly flooded or waterlogged soil	180	3.1	100
Artificial surfaces and associated areas	190	55.0	100
Bare land	200	0.2	---
Water	210	156.1	100
Total		15,235.0	3,930

$$n_{+j} = \sum_{i=1}^{k} n_{ij} \tag{3}$$

The number of samples for each land cover class (Table 1) are the row totals represented by n_{i+}. The values n_{+j} represent column totals in the error matrix. The user's accuracy for land cover class *i* is

$$UA_i = \left(\frac{n_{ii}}{n_{i+}} \right) \tag{4}$$

In this case, the number of correctly classified pixels (n_{ii}) is divided by the row total (n_{i+}). The user's accuracy (an expression of the error of commission) provides an indication of the likelihood that a pixel classified as a certain land cover class actually is that land cover class in the validation (reference) database. In other words, it is an estimate of the probability that a given pixel will appear on the ground as the land class that it was assigned.

The producer's accuracy (an expression of the error of omission) provides an indication of the percentage of the area of a certain land class in the validation (reference) database that was actually mapped as that land class. It is suggestive of the proportion of a given land cover class that was correctly classified. In order to determine the producer's accuracy for each land class, we developed an estimated population matrix by first applying the following adjustment to each entry in the omission-commission matrix in order to express each value as a proportion of the study area.

$$p_{ij} = \left(\frac{n_{ij}}{\sum_{j=1}^{k} n_{ij}}\right)\left(\frac{N_i}{\sum_{i=1}^{k} N_i}\right) \tag{5}$$

The producer's accuracy for land cover class *j* is then

$$PA_j = \left(\frac{p_{jj}}{p_{+j}}\right) \tag{6}$$

Confidence intervals at the 95% confidence level were computed for the overall, user's and producer's accuracies using the methodology described in [20]. Whether using the omission-commission (confusion) matrix or the estimated population matrix, the overall and user's accuracies are the same.

The methods employed in developing global land cover databases may not always be able to recognize and acknowledge special characteristics of different regions [12]. For land cover maps, the various statements of agreement can be sensitive to the level of aggregation of land classes. For example, maps employing a great number of classes are likely to be less accurate according to the *overall accuracy* metric, than similar maps employing a fewer number of classes [13]. Arino et al. [16] and Bontemps et al. [7] have suggested that *possibilities of agreement* measures would better represent the utility of the map to the user. Due to various sources of confusion, there may be cases where the validation database and the GlobCover database are in agreement when a class falls within a range of categories that include a mosaic class and complies with dominance criteria in the category definition. These *possibilities of agreement* therefore involve similarities between the GlobCover land cover map and the reference data that are different from the diagonal cells in the omission-commission matrix. We did not engage in land cover class aggregation *per se*, or in a simple expansion of the major diagonal [26], because the classification scheme is comprised of an ordered list of discrete classes that are not necessarily similar in theory. Therefore we followed the suggestion of Bontemps et al. [7] and recognized a number of *possibilities of agreement* relationships (Table 2) in a supplemental analysis of land cover classification agreement.

When the adjusted classification approach is employed,

$$\hat{p}_h = \sum_{b=1}^{Bh} f_{b+} \tag{7}$$

$$PA_j = \left(\frac{\sum_{c=1}^{Cj} g_{+c}}{p_{+j}} \right) \tag{8}$$

$$UA_i = \left(\frac{\sum_{b=1}^{Bi} f_{b+}}{p_{i+}} \right) \tag{9}$$

with each of these using the relationships

$$g_{+c} \subseteq p_{+j} \tag{10}$$

$$f_{b+} \subseteq p_{i+} \tag{11}$$

Here, *b* represents a row in the estimated population matrix, while *c* represents a column. The set g_{+c} (*Cj*) represents the reference classes that are recommended for an assessment approach using *possibilities of agreement* relationships (Table 2) to the determination of the producer's accuracy of land class *j*, while the set f_{b+} (*Bh=Bi*) represents the land cover classes that are recommended for a similar approach to the determination user's accuracy of land class *i*; f_{b+} and g_{+c}do not necessarily include all of the classes within a row (*b*) or column (*c*), but reflect the relationships described in Table 2 as well.

Finally, when it is presented, the *kappa* statistic is always smaller than the measure of overall agreement, yet it has been suggested to under-estimate the probability of a correct classification [30]. We acknowledge that the kappa statistic is not recommended by some [43], and have chosen not to present it here.

Table 2. Interactions between the GlobCover and the reference data that suggest a *possibilities of agreement* approach in the land cover analysis (from [7]).

GlobCover land class	Reference data
11 - Post-flooding or irrigated croplands	14 - Rainfed croplands
14 - Rainfed croplands	11 - Post-flooding or irrigated croplands
20 - Mosaic cropland / vegetation	11 - Post-flooding or irrigated croplands 14 - Rainfed croplands
30 - Mosaic vegetation / cropland	40 - Closed to open (>15%) broadleaved evergreen and/or semi-deciduous forest 50 - Closed (>40%) broadleaved deciduous forest 60 - Open (15-40%) broadleaved deciduous forest 70 - Closed (>40%) needle-leaved coniferous forest 80 - Open (15-40%) needle-leaved coniferous forest 100 - Closed to open (>15%) mixed forest 130 - Closed to open (>15%) shrubland 140 - Closed to open (>15%) broadleaved evergreen and/or semi-deciduous forest
110 - Mosaic forest/shrubland/grassland	40 - Closed to open (>15%) broadleaved evergreen and/or semi-deciduous forest 50 - Closed (>40%) broadleaved deciduous forest 60 - Open (15-40%) broadleaved deciduous forest 70 - Closed (>40%) needle-leaved coniferous forest 80 - Open (15-40%) needle-leaved coniferous forest 100 - Closed to open (>15%) mixed forest 130 - Closed to open (>15%) shrubland
40 - Closed to open (>15%) broadleaved evergreen and/or semi-deciduous forest 50 - Closed (>40%) broadleaved deciduous forest 60 - Open (15-40%) broadleaved deciduous forest 70 - Closed (>40%) needle-leaved coniferous forest 80 - Open (15-40%) needle-leaved coniferous forest	100 - Closed to open (>15%) mixed forest

3. RESULTS AND DISCUSSION

In general, the Glob Cover land cover map seems to represent certain broad sets of land classes of the State of Georgia in a logical manner (Figure 2). This non-quantitative perspective is informative for landscape-level descriptions of resources. In the following sections, we describe pixel and pixel-based and block-based assessments of land cover map and reference data agreement along with the measures of agreement described above.

3.1 Using Pixels as the Sampling Unit

The overall agreement of the GlobCover 2009 map for the State of Georgia was estimated to be about 47.8% when agreement was defined as a match in land class between the reference data and the GlobCover land class assigned to the sampled pixels (Table 3). An absolute precision of 1.65% at the 95% confidence level was obtained, thus the overall agreement was estimated to be within the range [46.2,49.5]. In examining the estimated population matrix (Table 4), one can reasonably arrive at a conclusion that there is a considerable amount of confusion among land cover classes. For example, the largest single-class sample of GlobCover pixels (GlobCover class 50 - closed, deciduous forests) was assigned to 15 different reference land classes, according to the reference data. As we noted, the producer's accuracy provides an indication of the percentage of the area of each land class in the reference data that was actually mapped as that land class. When using the pixel as a sample unit, estimated producer's accuracies (Table 5) were highest for closed, deciduous forests (86.0%), water (68.9%), closed broadleaved, flooded, saline forests (65.7%), and closed coniferous forests (62.6%). However, estimated producer's accuracies were also very low for the closed to open shrubland (0.7%), the mosaic grassland / forest-shrubland (8.0%), and the mosaic vegetation / cropland (7.0%). Each of these were often classed as closed coniferous or close deciduous forests in the GlobCover map. With the exception of closed to open shrubland, each were also often classed as closed to open grassland. Each of these were meant to represent heterogeneous land classes, and perhaps were more difficult to determine in the southern United States than in other areas of the world, even though an equal-reasoning regional stratification process was used in the development of the land cover map [7].

The user's accuracy provides an indication of the likelihood that a pixel classified as a certain land cover class actually is that land cover class on the ground. The estimated user's accuracy when individual pixels were used as the sampling unit was highest for closed to open vegetation on regularly flooded soil (93%), artificial surfaces (89%), water (89%), and rainfed cropland (87%). The estimated user's accuracy for the large classes was moderate: closed deciduous forests (47.7%), closed coniferous forests (57.7%), closed to open mixed forests (67.5%). The estimated user's accuracy for several classes on the landscape was low: closed deciduous forests in regularly flooded, saline water areas (6.0%), closed to open grassland (14.3%), closed to open shrubland (18.0%). In the case of closed deciduous forests in regularly flooded, saline water areas, these were mostly considered vegetation on flooded land (not necessarily saline water), even though the data were derived through direct association with reference data in the land classification process [7]. The GlobCover *closed to open shrubland* class was often misapplied to artificial surfaces or bare land. The GlobCover *closed to open grassland* class was often misapplied to rainfed cropland, mosaic cropland / vegetation (or vice versa). Confidence intervals ranged from 3 to 10% around the estimate, with tighter intervals suggested for those classes that had larger sample sizes.

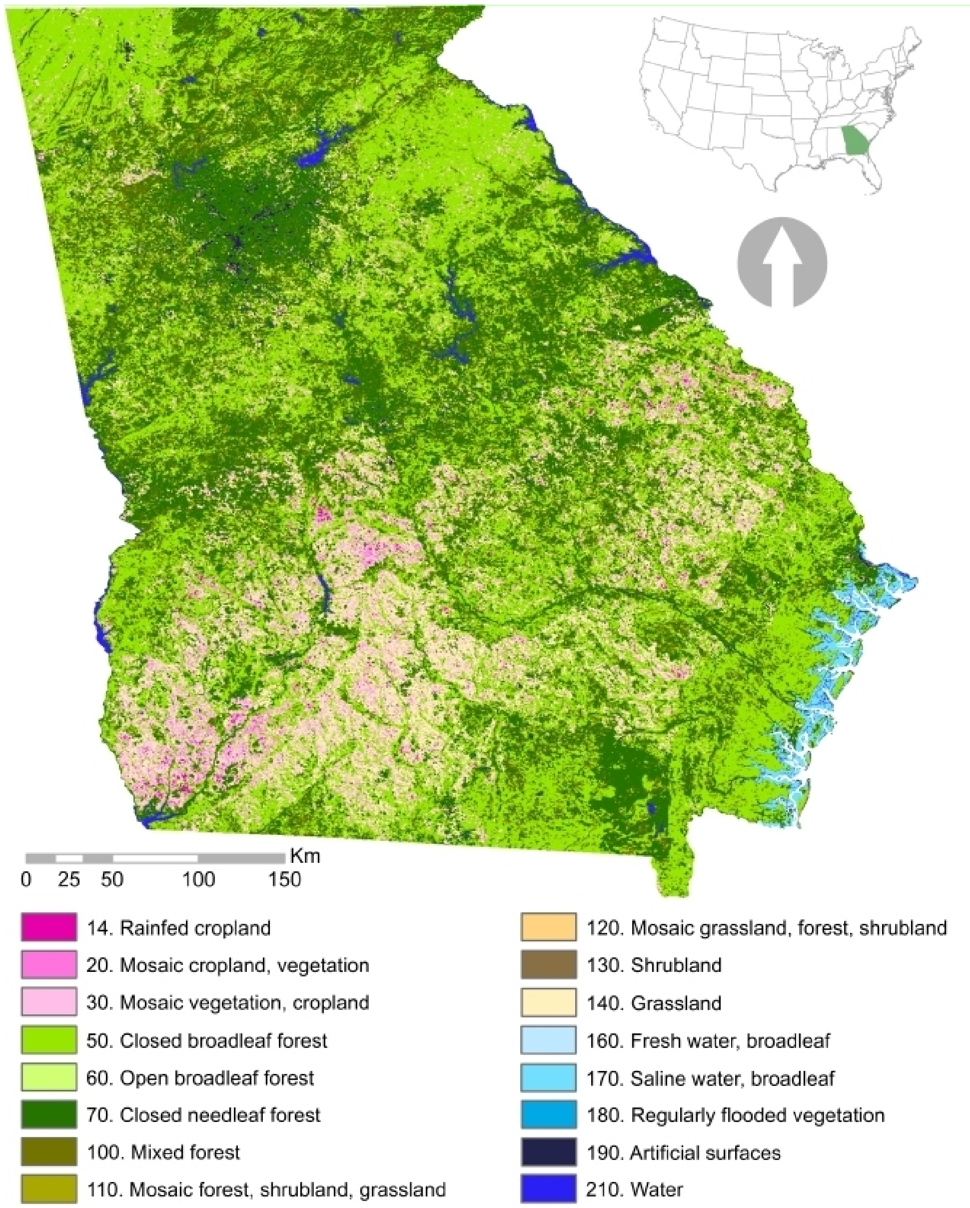

Fig. 2. Allocation of broad land cover classes from the GlobCover 2009 database for the State of Georgia

Table 3. Overall agreement and precision for the analyses conducted

Sample Unit	Class aggregation	Overall agreement (%)	Overall agreement absolute precision (%)
Pixel	No	47.8	1.65
Pixel	Yes	57.1	1.59
Block	No	47.9	1.62
Block	Yes	59.1	1.54

[a]*Utilizing the possibilities of agreement noted in Table 2.*

3.2 Using Pixel Blocks as the Sampling Unit

When pixel blocks were used as the sampling unit, the overall agreement of the GlobCover 2009 map for the State of Georgia increased slightly, to about 47.9% (Table 3). The absolute precision of the overall agreement was estimated to be 1.59% at the 95% confidence level, thus the overall agreement was estimated to be within the range [46.4,49.5]. This information suggests that the GlobCover land classification moderately agrees with the reference data. As when using individual pixels as the sample units, when pixel blocks are used as the sampling units, a considerable amount of confusion among land cover classes occurs. The estimated producer's accuracies (Table 5) were highest in closed broadleaved, regularly flooded, fresh water forests (97.9%), grasslands (92.9%), and closed deciduous forests (92.1%). However, as before, the estimated producer's accuracies were also very low for the closed to open shrubland (0.7%), the mosaic grassland / forest-shrubland (4.9%), and the mosaic vegetation / cropland (5.2%) class. Each of these were often classed as closed coniferous or closed deciduous forests in the GlobCover map. And as before, with the exception of closed to open shrubland, each were also often classed as closed to open grassland.

The estimated user's accuracy when pixels blocks were used as the sampling unit was highest for closed to open vegetation on regularly flooded soil (90%), and artificial surfaces (98%). The estimated user's accuracy for the large classes was moderate: closed deciduous forests (40.8%), closed coniferous forests (58.5%), closed to open mixed forests (78.2%). As in the previous analysis, the estimated user's accuracy for several classes on the landscape was low: closed deciduous forests in regularly flooded, saline water areas (5.0%), closed to open grassland (4.8%), closed to open shrubland (8.0%). Again, similar to the previous analysis, (a) in the case of closed deciduous forests in regularly flooded, saline water areas, these were mostly considered vegetation on flooded land (not necessarily saline water); (b) the *closedto open shrubland* class was often misapplied to artificial surfaces or bare land, and (c) the *closed to open grassland* class was often misapplied to rainfed cropland, mosaic cropland / vegetation (or vice versa). As with previous cases, the confidence intervals for these classes ranged from 3 to 10% around the estimate, with tighter intervals suggested for those classes that had larger sample sizes.

Table 4. Estimated population matrix for the case when a single pixel was employed as the sampling unit, indicating sample proportional agreement between the reference data and the GlobCover 2009 map.

GlobCover class	Reference data																			
	14	20	30	50	60	70	90	100	110	120	130	140	150	160	170	180	190	200	210	Total
14	0.29	0.01	0.02	--	--	--	--	--	--	--	--	0.01	--	--	--	--	0.01	--	--	0.34
20	0.43	0.58	0.47	--	--	--	--	--	0.03	0.03	0.02	0.02	--	--	--	--	0.03	--	--	1.61
30	1.05	1.21	1.31	0.03	--	0.13	--	0.20	0.29	0.26	--	0.20	--	--	--	--	0.16	--	0.03	4.88
50	0.30	0.49	1.35	18.21	0.03	5.25	0.20	4.89	3.08	2.26	0.82	0.95	--	--	--	0.16	0.16	0.03	--	38.20
60	0.00	--	0.02	0.02	0.15	0.08	0.01	0.12	0.02	0.03	0.02	--	--	--	--	--	0.00	--	--	0.50
70	0.26	0.43	0.62	1.61	0.16	13.34	0.33	3.06	0.56	0.82	1.25	0.23	--	--	0.03	0.13	1.61	0.23	0.16	24.84
90	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
100	0.03	0.13	0.16	0.79	--	1.87	0.13	10.38	0.43	0.39	0.49	0.03	--	0.03	--	0.16	0.26	0.07	--	15.37
110	0.08	0.08	0.03	--	0.01	0.08	0.01	0.08	0.48	0.03	0.04	0.03	--	--	--	--	0.05	0.01	0.01	1.03
120	0.08	0.04	0.08	0.01	--	0.04	0.01	0.12	0.16	0.34	0.03	0.02	--	--	--	--	0.06	--	--	0.98
130	0.01	--	0.00	0.00	--	0.00	--	0.00	0.00	--	0.02	--	0.00	--	--	0.01	0.04	0.02	0.01	0.10
140	1.97	0.85	1.58	0.46	--	0.46	--	0.43	0.89	0.72	--	1.38	--	0.10	--	--	0.62	0.07	0.10	9.62
150	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
160	0.00	0.00	0.00	--	--	0.01	--	0.00	--	0.00	--	0.00	--	0.02	--	--	0.03	--	0.00	0.08
170	--	--	--	0.01	--	--	--	0.04	--	--	--	--	--	--	0.06	0.85	--	--	0.10	1.06
180	--	--	--	--	--	--	--	--	--	--	--	--	--	0.00	0.00	0.02	--	--	0.00	0.02
190	--	--	--	--	--	0.02	--	0.01	--	0.00	--	--	--	--	--	--	0.32	--	--	0.36
200	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
210	--	--	0.01	0.02	--	0.02	--	--	--	--	--	--	0.01	--	--	--	0.04	0.01	0.91	1.02
Total	4.51	3.83	5.65	21.17	0.36	21.31	0.69	19.32	5.94	4.91	2.69	2.87	0.01	0.15	0.10	1.33	3.41	0.43	1.32	100.0

Table 5. Producers and users accuracy for GlobCover land classes in the State of Georgia, using a single pixel and a pixel block as the sample unit.

	Single pixel		Pixel block	
GlobCover class value	Producer's accuracy estimate (%)	User's accuracy estimate (%)	Producer's accuracy estimate (%)	User's accuracy estimate (%)
14	6.5	87.0	12.3	65.0
20	15.1	36.0	19.6	56.0
30	23.2	26.8	20.0	44.3
50	86.0	47.7	92.1	40.8
60	41.8	30.0	35.1	19.0
70	62.6	53.7	76.9	58.5
100	53.7	67.5	51.9	78.2
110	8.0	46.0	4.9	53.0
120	7.0	35.0	5.2	21.0
130	0.7	18.0	0.7	8.0
140	48.1	14.3	92.9	4.8
160	13.9	27.0	97.9	12.0
170	65.7	6.0	61.4	5.0
180	1.4	93.0	1.3	90.0
190	9.4	89.0	9.6	98.0
210	68.9	89.0	79.6	77.0

3.3 Possibilities of Agreement among Classified Pixels and Pixel Blocks

As we expected, when the *possibilities of agreement* relationships (Table 2) are employed, measures of estimated agreement increase. When individual pixels were assumed to be the sampling units, estimated overall agreement increased to 57.1%. The absolute precision of the overall agreement was estimated here to be 1.62% at the 95% confidence level, thus the overall agreement was estimated to be within the range [55.5,58.7]. These results represent nearly a 10% increase in estimated agreement. Given the relationships noted in Table 2, one notable change in estimated agreement arose: the producer's accuracy of the closed to open mixed forest increased from 53.7% to 96.9%. The user's accuracies of the closed deciduous forest rose from 47.7% to 60.5%, and the closed coniferous forest rose from 53.7% to 66.0%. Other minor increases in estimated user's and producer's accuracy were also noted.

When pixel blocks were assumed to be the sampling units, overall agreement increased to 59.1%. The absolute precision of the overall agreement was 1.54% at the 95% confidence level, thus the overall agreement was estimated to be within the range [57.5,60.6]. This represents nearly an 11-12% increase in estimated agreement. Notable changes in estimated producer's and user's accuracies include the following: (a) the estimated producer's accuracy of the closed to open mixed forest class increased from 51.9% to 98.7%; (b) the estimated user's accuracy of the closed deciduous forest increased from 40.8% to 59.8%; and (c) the estimated user's accuracy of the closed coniferous forest class increased from 58.5% to 72.0%. Confidence intervals for these estimates were very similar to those developed earlier for the single pixel and pixel block assessments of land cover classes.

3.4 Discussion of Findings

Our results are based on a sample size larger than other reported studies involving the GlobCover global landcover map [6,16,17,18,19]. While our study was situated in a region where land cover agreement had not previously been assessed, the overall agreement of our assessment is slightly lower than most of these previous assessments. We agree with [19] that some of the problems in land cover classification are due in part to the heterogeneity of land uses across the landscape. The State of Georgia contains a high level of heterogeneity of land cover types, due in part to the systems employed to distribute land to colonists in the 17th and 18th centuries. As Smith et al. [44] suggested, when the heterogeneity of land cover classes increases, the probability that a pixel in a database (particularly one larger than the average patch size) will be misclassified should increase. Pflugmacher et al.[5] similarly suggested that agreement tends to be lower in areas with complex, heterogeneous land uses and where there exist spectrally similar land classes. In areas where the patches are larger than average, the probability of misclassification of a pixel should decrease [44]. While there were some areas of contiguous large patches in the northern mountains and along the coasts, these were not in sufficient supply to influence the overall agreement of the database with the reference imagery. Areas of disagreement between the GlobCover global land cover map and the validation data reflect issues that perhaps need more attention on behalf of the producers of the map [29]. These issues include (a) shrubland and mosaic vegetation classes being misclassified as deciduous or coniferous forest classes; and (b) misapplying shrubland or grassland class labels to bare land, cropland, or artificial surfaces.

Other areas of confusion can arise between forest successional stages. Forest management is a cyclical, progressive process that occurs over a long period of time on a specific area of land, thus forests may be considered open and forested when the trees are very young (recently planted or regenerated) or very old [45], and closed during the intermediate stages. In our analysis, if at any time it were clear that trees were being grown on a piece of land, it was considered forested. However, the period immediately after harvest and during or after site preparation (if artificially regenerated) could have easily been confused with other land classes. Further, mixed forest types can become confused with other forest types, non-tree vegetation, and even cultivated land [15]. Most global land cover maps are unable to adequately discriminate among these classes due to the mixture of different vegetative life forms and their reflective qualities. Mapping a continuum of land cover features as a single discrete category can therefore be problematic [1], particularly as the heterogeneity of landscape features is greater than the spatial resolution of the remote sensor can accommodate [44]. The ambiguous signature described by a pixel can test the sensitivity of a classification process to different combinations of classes [13]. This separability of classes with similar spectral signatures seems to be a common problem for global land cover mapping projects [13].

In terms of sampling design, one could have arrived at measures reflecting greater agreement by restricting the placement of pixels and blocks in a non-probabilistic (and undesirable) manner to areas across the landscape that were very homogenous in nature [25], however, we chose to adhere to the random nature of sample point placement. The statistical basis for the inferences derived is ensured through the use of probability sampling [44]. In many cases, the characteristic resolution of features in urbanized areas is relatively small compared to the size of a GlobCover pixel, therefore a large number of mixed pixels can be located in urban areas, leading to a high degree of map variability [47]. Stehman and Wickham [31] also note that pixel-based assessments are generally more sensitive to location error (which is around 77 m in MERIS data [16]) than assessments using larger spatial units such as blocks. While our measures of agreement tended to increase when using a block of pixels, the increase was marginal. Therefore, in areas where there is considerable heterogeneity in land cover classes across the landscape, the use of a larger spatial sampling unit does not seem to sufficiently increase the efficiency or level of agreement from the assessment process.

We noted through our literature review that a global land cover map can be used to assess land-cover trends, to study managed and natural ecosystems, and to facilitate regional and global sustainability and climate change modeling [16]. However, Jung et al. [13] found that the cropland and natural vegetation mosaics were the least reliable land cover classes in an assessment of several global land cover maps. Further, others have indicated that the classification agreement of some categories (e.g., wetland and mixed forests) can be sensitive to the landscape patch size [48]. We have attempted to illustrate that while the GlobCover 2009 map may in fact be of value for facilitating an assessment of resources on a regional level, the amount of heterogeneity inherent across the landscape can cause the misclassification of land classes and the misapplication of class labels to certain land classes.

4. CONCLUSION

Although several validation exercises involving the GlobCover 2009 global land cover map have been previously performed, this is perhaps the most extensive examination in terms of validation points, and the first exclusively within the United States. This study focused on the ability of a land cover map, developed to represent global land resources, to represent well the resources of a specific yet important forest area in the southern United States. While a relatively robust statistical design was employed, the estimated classification agreement was moderate (yet below 50%) when either an individual pixel or a block of pixels were used as the sampling units. Closed canopy mixed forest areas, coniferous forest areas, and deciduous forest areas, three of the broadest land types in the State of Georgia, and representing about 78% of the GlobCover land classes in the State, had estimated user's and producer's accuracies ranging from 60 to 97% when a classification approach was employed using *possibilities of agreement* among land classes.

ACKNOWLEDGEMENTS

This work was supported by the Warnell School of Forestry and Natural Resources at the University of Georgia. The GlobCover 2009 global land cover map was obtained from European Space Agency and University Catholique de Louvain.

COMPETING INTERESTS

There are no financial and personal relationships with other people or organizations that could inappropriately influence (bias) this work. Therefore, no competing interests exist.

REFERENCES

1. Foody GM. Status of land cover classification accuracy assessment. Remote Sens Environ. 2002; 80:185-201. doi:10.1016/S0034-4257(01)00295-4.
2. Selkowitz DJ, Stehman SV. Thematic accuracy of the National Land Cover Database (NLCD) 2001 land cover for Alaska. Remote Sens Environ. 2011;115:1401-1407. doi:10.1016/j.rse.2011.01.020.
3. Vintrou E, Desbrosse A, Bégué A, Traoré S, Baron C, Seen DL. Crop area mapping in West Africa using landscape stratification of MODIS time series and comparison with existing global land products. Int J Appl Earth Obs Geoinf. 2012;14:83-93. doi:10.1016/j.jag.201.06.010.
4. Comber A, Fisher P, Wadsworth R. What is land cover? Environ Plann B Plann Des. 2005;32:199-209. doi:10.1068/b31135.
5. P flugmacher D, Krankina ON, Cohen WB, Friedl MA, Sulla-Menashe D, Kennedy RE, *et al*. Comparison and assessment of coarse resolution land cover maps for Northern Eurasia. Remote Sens Environ. 2011;115: 3539-3553. doi:10.1016/j.rse.2011.08.016.
6. Yiming A, Wenwu Z, Yinhui Z. Accuracy assessments of the GLOBCOVER dataset using global statistical inventories and FLUXNET site data. Acta Ecologica Sinica. 2012;32:314-320. doi:10.1016/j.chnaes.2012.09.001.
7. Bontemps S, Defourny P, van Bogaert E, Arino O, Kalogirou V, Perez JR. 2011. GlobCover 2009, Products description and validation report. Frascati, Italy: European Space Agency, and Louvain-la-Neuve, Belgium: Université Catholique de Louvain; 2011.
8. Arino O, Trebossen H, Achard F, Leroy M, Brockman C, Defourny P, *et al*. The GlobCover initiative. Paper presented at MERIS (A) ATSR Workshop 2005, Frascati, Italy, September 26-30. Publication ESA SP-597. Noordwijk, The Netherlands: European Space Agency; 2005.
9. Cleve C, Kelly M, Kearns FR, Moritz M. Classification of the wildland–urban interface: A comparison of pixel- and object-based classifications using high-resolution aerial photography. Comput Environ Urban Syst. 2008;32:317-326. doi:10.1016/j.compenvurbsys.2007.10.001.
10. Schweers W, Bai Z, Campbell E, Hennenberg K, Fritsche U, Mang H-P, *et al*. Identification of potential areas for biomass production in China: Discussion of a recent approach and future challenges. Biomass Bioenergy. 2011;35:2268-2279. doi:10.1016/j.biombioe.2011.02.034.
11. Olofsson P, Stehman SV, Woodcock CE, Sulla-Menashe D, Sibley AM, Newell JD, *et al*. A global land-cover validation data set, part I: Fundamental design principles. Int J Remote Sens. 2012; 33: 5768-5788. doi:10.1080/01431161.2012.674230.
12. Miettinen J, Shi C, Tan WJ, Liew SC. 2010 land cover map of insular Southeast Asia in 250-m spatial resolution. Remote Sens Lett. 2012;3:11-20. doi:10.1080/01431161.2010.526971.
13. 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. doi:10.1016/j.rse.2006.01.020.
14. Chen D, Wei H. The effect of spatial autocorrelation and class proportion on the accuracy measures from different sampling designs. ISPRS J Photogramm Remote Sens. 2009; 64:140-150.doi:10.1016/j.osprsjprs2008.07.004.

15. Herold M, Mayaux P, Woodcock CE, 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. doi:10.1016.j.rse.2007.11.013.
16. Arino O, Bicheron P, Achard F, Latham J, Witt R, Weber J-L. GlobCover, the most detailed portrait of Earth. Bulletin 136. Noordwijk, The Netherlands: European Space Agency; 2008.
17. Song X-P, Huang C, Sexton JO, Feng M, Narasimhan R, Channan S, Townshend JR. An assessment of global forest cover maps using regional higher-resolution reference data sets. In Proceedings of the 2011 IEEE International Geoscience and Remote Sensing Symposium, 752-755. New York: Institute of Electrical and Electronics Engineers; 2011.
18. Pérez-Hoyos A, García-Haro FJ, 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. doi:10.1016/j.jsprsjprs.2012.09.006.
19. Li Z, Wu W, Zhou Q, Chen Z, Li Z. Assess the accuracy of the Globcover cultivated land data in Northeast China. In Proceedings of the First International Conference on Agro-Geoinformatics, 1-5. New York: Institute of Electrical and Electronics Engineers; 2012.
20. Carrão H, Araújo A, Gonçalves P, Caetano M. Multitemporal MERIS images for land-cover mapping at a national scale: A case study of Portugal. Int J Remote Sens. 2010; 31: 2063-2082. doi:10.1080/01431160902942910.
21. Bicheron P, Amberg V, Bourg L, Petit D, Huc M, Miras B, Brockmann C, et al. Geolocation assessment of MERIS GlobCover orthorectified products. IEEE Trans Geosci Remote Sens. 2011; 49: 2972-2982. doi: 10.1109/TGRS.2011.2122337.
22. Latifovic R, Zhu Z, Chilar J, Giri C, Olthof I. Land cover mapping of North and Central America - Global Land Cover. Remote Sens Environ. 2004;89:116-127. doi: 10.1016/j.rse.2003.11.002.
23. Li M, Mao L, Zhou C, Vogelmann JE, Zhu Z. Comparing forest fragmentation and its drivers in China and the USA with Globcover v2.2. J Environ Manage. 2010; 91: 2572-2580. doi: 10.1016/j.jenvman.2010.07.010.
24. Pierdicca N, Bignami C, Roca M, Féménias P, Fascetti M, Mazzetta M, et al. Transponder calibration of the Envisat RA-2 altimeter Ku band sigma naught. Adv Space Res. 2013; 51: 1478-1491. doi:10.1016./j.asr.2012.12.014.
25. Stehman SV. Sampling designs for accuracy assessment of land cover. Int J Remote Sens. 2009; 30: 5243-5272. doi:10/1080/01431160903131000.
26. Congalton RG, Green K. 2009. Assessing the accuracy of remotely sensed data, principles and practices. Second edition. Boca Raton, Florida: CRC Press; 2009.
27. Wickham JD, Stehman SV, Smith JH, Yang L. Thematic accuracy of the 1992 National Land-Cover Data for the western United States. Remote Sens Environ. 2004;91:452-468. doi:10.1016/j.rse.2004.04.002.
28. Zhu Z, Yang L, Stehman SV, Czaplewski RL. Accuracy assessment for the U.S. Geological Survey regional land-cover mapping program: New York and New Jersey region. Photogramm Eng RemoteSensing. 2000;66:1425-1435. doi:0099-1112/00/6612-1425.

29. Tchuenté ATK, Roujean J-L, De Jong SM. Comparison and relative quality assessment of the GLC2000, GLOBCOVER, MODIS and ECOCLIMAP land cover data sets at the African continental scale. Int J Appl Earth Obs Geoinf. 2011;13:207-219. doi:10.1016/jag.2010.11.005.
30. Stehman SV. Selecting and interpreting measures of thematic classification accuracy. Remote Sens Environ. 1997;62:77-89. doi:10.1016/S0034-4257(97)00083-7.
31. Stehman SV, Wickham JD. Pixels, blocks of pixels, and polygons: Choosing a spatial unit for thematic accuracy assessment. Remote Sens Environ. 2011; 115: 3044-3055. doi:10.1016/j.rse.2011.06.007.
32. Kelly M, Shaari D, Guo Q, Liu D. A comparison of standard and hybrid classifier methods for mapping hardwood mortality in areas affected by “sudden oak death." Photogramm Eng RemoteSensing. 2004;70:1229-1239.
33. Wickham JD, Stehman SV, Gass L, Dewitz J, Fry JA, Wade TG. Accuracy assessment of NLCD 2006 land cover and impervious surface. Remote Sens Environ. 2013;130: 294-304. doi:10.1016/j.rse.2012.12.001.
34. Lo CP, Watson LJ. The influence of geographic sampling methods on vegetation map accuracy evaluation in a swampy environment. Photogramm Eng RemoteSensing. 1998;64:1189-1200.
35. van Genderen JL, Lock BF, Vass PA. Remote sensing: Statistical testing of thematic map accuracy. Remote Sens Environ. 1978;7:3-14. doi:10.1013/0034-4257(78)90003-2.
36. Scepan J. Thematic validation of high-resolution global land-cover data sets. Photogramm Eng RemoteSensing. 1999;65:1051-1060.
37. Thapa RB, Murayama Y. Urban mapping, accuracy, & image classification: A comparison of multiple approaches in Tsukuba City, Japan. Appl Geogr. 2009; 29: 135-144. doi:10.1016/j.apgeog.2008.08.001.
38. Wulder MA, Franklin SE, White JC, Linke J, Magnussen S. An accuracy assessment framework for large-area land cover classification products derived from medium-resolution satellite data. Int J Remote Sens. 2006;27:663-683. doi:10.1080/01431160500185284.
39. Stehman SV, Czaplewski RL. Design and analysis for thematic map accuracy assessment: Fundamental principles. Remote Sens Environ. 1998;64:331-344. doi:10.1016/S0034-4257(98)00010-8.
40. Liu C, Frazier P, Kumar L. Comparative assessment of the measures of thematic classification accuracy. Remote Sens Environ. 2007;107:606-616. doi:10.1016/j.rse.2006.10.010.
41. Olofsson P, Foody GM, Stehman SV, Woodcock CE. 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. doi:10.1016/j.rse.2012.10.031.
42. Foody GM. On the compensation for chance agreement in image classification accuracy assessment. Photogramm Eng Remote Sensing. 1992;58:1459-1460.
43. Pontius RG Jr, Millones M. Death to the Kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment. Int J Remote Sens. 2011;32:4407-4429. doi: 10.1080/01431161.2011.552923.
44. Smith JH, Wickham JD, Stehman SV, Yang L. Impacts of patch size and land-cover heterogeneity on thematic image classification accuracy. Photogramm Eng Remote Sensing. 2002;68:65-70.

45. Bennett B. What is a forest? On the vagueness of certain geographic concepts. Topoi. 2001;20:189-201. doi: 10.1023/A:1017965025666.
46. Stehman SV. Practical implications of design-based sampling inference for thematic map accuracy assessment. Remote Sens Environ. 2000;79:35-45. doi:10.1016/S0034-4257(99)00090-5.
47. Potere D, Schneider A, Angel S, Civco DL. Mapping urban areas on a global scale: which of the eight maps now available is more accurate? Int J Remote Sens. 2009;30: 6531-6558. doi:10.1080/01431160903121134.
48. Smith JH, Stehman SV, Wickham JD, Yang L. Effects of landscape characteristics on land-cover class accuracy. Remote Sens Environ. 2003;84:342-349. doi:10.1016/S0034-4257(02)00126-8.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sciencedomain.org/review-history.php?iid=359&id=22&aid=2736