Collaborative validation of GlobeLand30: Methodology and practices

Jun Chen, Lijun Chen, Fei Chen, Yifang Ban, Songnian Li, Gang Han, Xiaohua Tong, Chuang Liu, Vanya Stamenova and Stefan Stamenov

ABSTRACT

30-m Global Land Cover (GLC) data products permit the detection of land cover changes at the scale of most human land activities, and are therefore used as fundamental information for sustainable development, environmental change studies, and many other societal benefit areas. In the past few years, increasing efforts have been devoted to the accuracy assessment of GlobeLand30 and other finer-resolution GLC data products. However, most of them were conducted either within a limited percentage of map sheets selected from a global scale or in some individual countries (areas), and there are still many areas where the uncertainty of 30-m resolution GLC data products remains to be validated and documented. In order to promote a comprehensive and collaborative validation of 30-m GLC data products, the GEO Global Land Cover Community Activity had organized a project from 2015 to 2017, to examine and explore its major problems, including the lack of international agreed validation guidelines and on-line tools for facilitating collaborative validation activities. With the joint effort of experts and users from 30 GEO member countries or participating organizations, a technical specification for 30-m GLC validation was developed based on the findings and experiences. An on-line validation tool, GLCVal, was developed by integrating land cover validation procedures with the service computing technologies. About 20 countries (regions) have completed the accuracy assessment of GlobeLand30 for their territories with the guidance of the technical specification and the support of GLCVal.

1. Introduction

Accurate characterization of Global Land Cover (GLC) is essential for sustainable development monitoring, environmental change studies, land resource management, and many other application areas (Foley et al. 2010; Roger and Pielke 2005; Running 2008; Grimm et al. 2008; Zell et al. 2012). In the past few years, a variety of operational or experimental approaches have been developed to produce global, regional and national land cover data products with different spatial-temporal resolutions and thematic accuracy (Loveland et al. 2000; Hansen et al. 2000; Friedl et al. 2010; Radoux et al. 2013; Latham et al. 2014; Ban and Jacob 2013; Gong et al. 2016). The thematic accuracy of these land cover data products can be documented through sample-based validation or comparison with existing reference datasets. The results will not only help the users to understand the uncertainty and the scope of data application, but also enable the producers to examine and analyze the types, sources and spatial distribution of errors (Strahler et al. 2008; Rwamba and Ndambuki 2017).

GlobeLand30 is a 30-m resolution global land cover (GLC) data product that was developed by the National Geomatics Center of China using a Pixel-Object-Knowledge (POK) approach (Chen et al. 2015), and its 2000 and 2010 versions were released in 2014, with 2020 version just released (Chen et al. 2015; Chen, Liao, and Chen et al. 2014a, 2014b, 2017a, 2017b). Such finer resolution GLC datasets provide more details of land cover patterns, and permit the detection of land cover changes at the scale of most human land activities. GlobeLand30-2000 and GlobeLand30-2010 achieved an overall classification accuracy of over 80% at a global scale, as a result of third-party expert’s assessment with a two-rank sampling strategy using 154,586 pixel samples (Chen et al. 2015). Moreover, many other experts or users validated GlobeLand30 for their own territories or interested areas through a sample-based validation or comparison with existing land cover or other datasets, resulting in an average accuracy of 80% for all classes or one single class (Brovelli et al. 2015; Manakos et al. 2015; Yang et al. 2017). However, these accuracy assessments were conducted either within 10% map sheets selected from a global scale or in some
individual countries (regions). There are still areas where the uncertainty of GlobeLand30 and other finer-resolution resolution GLC data products remain to be validated and documented.

In order to promote a comprehensive validation of GlobeLand30 and other finer-resolution GLC datasets, an international cooperative project was launched by the Global Land Cover Community Activity of the Group of Earth Observation (GEO) in 2014. This three-year project had the following three main objectives. The first objective was to develop a technical specification with appropriate methods for the validation of 30-m GLC datasets at a global scale. The second objective was to develop advanced tools for facilitating the collaborative validation processing. The third objective was to mobilize GEO members and participating organizations to join the cooperative validation by integrating their validation expertise and resources. This initiative was supported by the International Cooperation Program of the Ministry of Science and Technology of China, and about 30 GEO member countries and participating organizations joined this validation project.

This paper presents the major considerations and preliminary results of this GEO-led validation project. Section 2 analyzes the major problems facing, including the lack of internationally agreed technical specification for 30-m GLC validation and an online tool for supporting collaborative validation activities. Section 3 presents the major technical considerations and development for formulating such a technical specification. The design and development of an online validation system, GLCVal, are then briefly described in Section 4. The collaborative validation practices are discussed in Section 5 with two case studies, followed by some concluding remarks.

2. Challenges facing global-scale validation

Land cover validation is a technical process where people select appropriate samples, at which reference data are collected and used to estimate the overall and class-specific accuracies of the given data product (Olofsson et al. 2014; Chen et al. 2016). As far as the sampling design is concerned, a number of approaches have been developed to determine the sample size and spatial distribution of samples, according to several fundamental criteria, such as probability, cost effectiveness and spatial balance. The stratified random sampling and two-stage cluster sampling are two commonly used approaches in GLC validation practices (Steelman and Czaplewski 1998; Steelman et al. 2012; Tong et al. 2011). Moreover, a few desktop and internet-based software tools have been developed to support the collection of reference data (Fritz et al. 2009; Clark et al. 2011). However, these sampling approaches and supporting tools could not satisfy the collaborative validation of 30-m resolution GLC products yet.

First, the 30-m resolution GLC data products present a strong spatial heterogeneity of land cover over the globe. The traditional sampling approaches cannot efficiently handle this spatial heterogeneity and may not produce credible spatial samplings. Typical problems include inappropriate sample sizes for target regions, under-represented sample numbers for rare classes, and irrational sample distribution in the geographical space (Chen et al. 2016). It is therefore necessary to take the spatial heterogeneity of land cover into consideration and to allow higher sample densities or larger sample sizes for more heterogeneous regions (Chen et al. 2016; Defourny et al. 2009; Congalton 1991). A Landscape Shape Index (LSI)-based approach was proposed to improve the quality of spatial sampling design for 30-m GLC validation, and its basic idea is to characterize the spatial heterogeneity with three-level LSIs for determining the subsequent sample sizes and their spatial distributions (Chen et al. 2016). Experimental tests showed that this LSI approach obtained more appropriate sample size for each region, sufficient sample numbers for rare classes, and optimal sample distributions in geographical space. The LSI-based sampling approach was therefore recommended to the GEO-led 30-m GLC validation project to solve the three problems associated with traditional sampling approaches.

Second, the collection of sample data over a large area is labor-intensive, costly and time-consuming. It is becoming even more difficult for GLC validation over the whole globe. With the rapid development of the internet technology, some web-based systems have been developed to enable people from anywhere in the world online access of reference data and other sample information (Stehman et al. 2012; Fritz et al. 2012; Han et al. 2015). The information includes very high-resolution satellite imagery, geo-tagged photos, and thematic maps. However, most of them are dedicated to some particular steps of validation, not supporting the entire procedure of validation. For instance, Geo-Wiki does help the image interpretation of reference data, but cannot support sample size calculation, automatic sample layout, precision calculations and other functions. It is therefore necessary to develop an online validation system (or tool) (See al. 2015). The major functions will include the integration and access of reference materials, estimation of total sample sizes and inter-class distribution, geospatial layout of samples, expert interaction checking, generating accuracy reports, online marking of errors, uploading and downloading of samples, and so on. It will help the experts and volunteers from different parts of the world to find and use reference information, upload sample information, make sample judgment and labeling, and annotate error messages.
Third, conducting a comprehensive accuracy assessment of finer-resolution GLC data products has now become a necessity for a number of reasons. One is the increasing demands from sustainable development monitoring, environmental change studies and many other application areas, where the accuracy assessment reports of large area or even global scale are requested. The other is the number of people working in the area of land cover validation (such as users and researchers) has increased rapidly and is forming a large voluntary validation community. They share common interests and have their own validation resources. Moving from individual validation practices to a GEO-led collaborative validation task force is now possible.

In order to promote and organize a GEO-led 30-m GLC validation, a technical specification needs to be developed to describe the appropriate approaches and procedures for sampling design, response, and analysis protocols. A web-based validation tool needs to be developed accordingly to enable the online and collaborative validation processes by experts from different parts of the world. The GEO members and participating organizations should be mobilized and invited to join the collaborative validation project with their expertise and available resources. Figure 1 illustrates the overarching conceptual framework of this GEO-led validation project.

3. Key methodological issues in developing the technical specification

In order to develop an internationally agreed technical specification on 30-m GLC validation, a number of technical issues were examined and explored, including LSI-based sampling design, reference data collection, sample judgment and labeling, and accuracy assessment.

3.1. LSI-based sampling design

A LSI-based sampling design was adopted for 30-m GLC validation to guarantee samples with good spatial representation and even spatial distribution. The design satisfies with the following criteria: relatively high sample size or density for heterogeneous landscapes, spatially distributed well (spread throughout the region without large gaps), representative sample numbers for all classes, especially rare classes and fragmented classes and self-adaptive ability to different regions (Chen et al. 2016). As shown in Figure 2, three LSIs are calculated for three different geographical levels,
and used to estimate the sample size for regions, allocate sample to each class, and distribute sample sites.

(1) Determining sample size: A regional LSI (called as rLSI) is used for the determination of the sample size of a given target region and the formula is as the following (Chen et al. 2016):

\[
N_i = \frac{r_{LSI_i} \times A_i}{\sum_{j=1}^{n} r_{LSI_j} \times A_j} \times N, (i = 1, 2, \ldots, n) \quad (1)
\]

For any given region \( i \), its sample size \( N_i \) can be derived using its regional landscape shape index \( (r_{LSI_i}) \) and its area \( (A_i) \), as well as those of the other regions, and the total size \( N \). It ensures that the regions that are more heterogeneous would have larger sample sizes or higher sample densities.

(2) Allocating sample size to classes: A class-level LSI (called as cLSI) is used for allocating the derived sample size of a specific target region to its individual land cover classes. In Neyman optimal allocation formula, \( S \) as the standard error of class \( k \), is the key element as the measure of dispersion, spatial variability and inhomogeneity of the spatial distribution. The standard deviation is greater for the class with stronger heterogeneity. Thus, more samples should be allocated. Here, the cLSIs are used as a proxy for \( S \):

\[
cN_{i,k} = N_i \times \frac{c_{LSI_{i,k}} \times W_{i,k}^a}{\sum_{i=1}^{r=m} c_{LSI_{i,j}} \times W_{i,j}^a}, (k = 1, 2, \ldots, m) \quad (2)
\]

where \( c_{LSI_{i,k}} \) is the cLSI of class \( k \) in region \( i \). The sample number of class \( k \) \( (cN_{i,k}) \) can be derived according to regional sample size \( N_i \), the areal proportion \( (W_{i,k}) \) and \( c_{LSI_{i,k}} \). Experiments showed that the sample size of rare land-cover class have been increased using cLSI-based class sample size allocation formula (Chen et al. 2016).

(3) Selecting sample sites: The per-class LSI (named uLSI) is calculated for each sample unit and a specific uLSI-curve can be created according to the level of spatial heterogeneity by sorting all units by uLSI in descending order. The x-axis pertains to the serial number in the geographical units and the y-axis pertains to uLSI pertains to the corresponding uLSI. In this way, the uLSI-curve can serve as a spatial heterogeneous space, and a two-stage sampling using uLSI-curve was designed for selecting sample units (Brovelli et al. 2015). In the first stage, a primary sampling unit (PSU) is selected via equal step sampling protocol along the x-axis of uLSI-curve. In the second-stage, one pixel is selected as sample sites using random sampling protocol within each PSU. More sample sites are located in these small patches. It means that the sample location considers both heterogeneous and homogeneous landscapes.

3.2. Reference data collection

The collection and integration of various reference data and resources are essential for collaborative GLC validation. The reference resources include Very High Resolution (VHR) images, thematic maps, in-situ measurements, crowdsourcing and VGI data, and so on. Before using these reference data, proper evaluations need to be done to ensure the data quality, such as the completeness, logical consistency, positional accuracy, thematic accuracy, temporal quality and usability.

Geo-Wiki, OpenStreetMap, in-situ measurements and other public geographic data are among the valuable reference sources for GLC validation. Geo-Wiki attempts to integrate open access to high-resolution satellite imagery from Google Earth with crowdsourcing into a single Web 2.0 application (Fritz et al. 2009, 2012). The user can plot GLC cover maps on the top of Google Earth, and display the disagreement maps between any pair of land cover products. In-situ measurement data resulted from fieldwork is another valuable reference source for GLC validation. Since field data collection is time-consuming and labor-intensive, their use at the global scale is limited due to high costs. As a result, field data collection can act as supplementary means in some places where qualified images are missing or the detailed class information is needed.

3.3. Sample judgment and labeling

With the help of an online tool, the land cover types of the sample points can be judged by validation experts or volunteers through the interpretation of high-resolution images, or comparison with large-scale land cover maps, as well as in-situ measurement data. Some additional supporting information can be provided or uploaded, such as text commentaries and photos of the samples.

A two-scale window strategy was proposed for interpretation of land cover types from imagery. As shown in the upper-right part of Figure 3, each sample point recorded as SID (Samples’ ID) is enclosed by two square areas, that is a 30 m * 30 m window, and a 300 m * 300 m window. The labels of two land cover types are then interpreted and assigned to each sample point. The first is interpreted according to the land cover covering the biggest proportion of the 30 by 30 meters window, and the second is based on the majority land cover in the 300 by 300 meters window.
These two labels are saved as LCval30 and LCval300 in the attribute table of the sample judgment, as shown in Figure 4.

In order to examine the credibility of the above sample judgment and labeling, a four-level confidence measure was further proposed. Level A means the sample judgment is absolutely correct. Level B implies that the judgment is absolutely incomplete, while Level C means the sample judgment is uncertain, which is due to reference data quality. The sample judgment of Level D is uncertain, which may be due to personal knowledge and interpretation ability. Each sample may have three possible results by the use of judgment methods: 100% correct judgment, 100% wrong judgment and uncertainty judgment. With the four levels of the degree of trust, the results of sample judgment can be evaluated and demonstrate the accuracy assessment in different trust. This can enable the elimination of potential false judgment samples that will influence the final accuracy evaluation, and guarantee the credibility of the GLC validation.

### 3.4. Accuracy assessment

Accuracy assessment aims to provide an index of how closely the derived class allocations depicted in the thematic land cover map represent reality. The commonly adopted method is confusion matrix and its derived basic descriptive measures, including overall accuracy, user and producer accuracy, Kappa and so on. Overall accuracy is a general measure for the entire map. It is simply the sum of major diagonal (i.e. the correctly classified sample units) divided by the number of sample units in the error matrix. User’s accuracy means the probability that a pixel classified on the map represents the class on the ground whereas producer’s accuracy indicates the probability of a reference pixel being correctly classified (Congalton 1991). By examining relationships between the two measures, the map user gains insight about the varied reliabilities of classes on the map, and the analyst learns about the performance of the process that generated the maps.

The Kappa analysis is a discrete multivariate technique used in accuracy assessment to statistically determine if an error matrix is significantly different from another. Not all agreements can be attributed to the success of the classification. Kappa attempts to provide a measure of agreement that is adjusted for chance agreement.

### 4. Online validation tool – GLCVal

A web-based system (GLCVal) was designed and developed to support the international collaborative validation of GlobeLand30. The basic idea is to integrate the concept of "Internet plus" and service computing technologies with land cover validation procedures (Chen et al. 2018). Figure 5 illustrates its major functionalities, supporting data services and algorithms. This online system (http://glcval.geo_compass.com) serves as a platform to assemble land cover validation methods and workflows into a single portal, including the storage and online access of all the available reference data, “wizards” guided validation workflow, self-adaptive sampling designs, user-friendly interpretation of samples from imagery, and generation of accuracy assessment report.

The GLCVal allows users to carry out their validation tasks through a step-by-step guidance. After logging into GLCVal, a new validation task can be created with a supporting functional sequence. A land cover product needs to be selected or uploaded and displayed. Users can choose any administrative or geographical region of interest or define the area of interest by uploading an administrative boundary data for validation. Then, they may choose LSI-based
| SID          | LCType | LCval30 | LCval300 | Thumbnail          |
|--------------|--------|---------|----------|-------------------|
| N34352010536993 | 90     | 20      | 90       | ![thumbnail](image) |
| N34352010541433 | 20     | 20      | 20       | ![thumbnail](image) |

**Figure 4.** Two-scale windows for sample judgment.

**Figure 5.** GLCVal- Function design.
sampling or other sampling approach to generate the sample dataset, which will be displayed on the screen. Users can interpret the land cover type of each sample point by comparing higher resolution image of the same place, as shown in Figure 6. When the judgment and labeling of all the samples are completed, an accuracy assessment report will be generated automatically.

The labeled samples are one of the important outputs of GLC validation. A sample metadata was developed and all the valid sample points were integrated, stored for further utilization and sharing.

5. Collaborative validation practices: case studies

About 30 GEO member countries and participating organizations joined this GEO-led international collaborative validation. The technical specification was used for guiding the collaborative validation of GlobeLand30 with the support of online tool GLCVal. Here, we present two case studies conducted in Bulgaria and Sweden.

The Republic of Bulgaria is located in the Southeastern Europe, the middle of the Balkan Peninsula, and covers an area of 110,994 km². The territory of Bulgaria is characterized by significant geodiversity and heterogeneous landscape determined by the variety of landforms, the high-level segmentation of the relief and the specifics of the climatic and soil types. The location of Bulgaria in the southern parts of the temperate climatic zone, the Black sea and the proximity to the Mediterranean Sea all influence the climatic elements and determine the transition from temperate continental climate in north, through transition climate zone and continental Mediterranean climate in the southern part. The transitivity of natural components and the diversity of the relief influence on the landscape structure and heterogeneity.

Sweden is located in northern Europe, and is one of the world’s northernmost countries. Stretching 1,574 km from north to south and 499 km from east to west, Sweden has a total area of 450,295 km² and is the third largest country in the European Union. Sweden is characterized by its long coastline, extensive forests and numerous lakes. 70% of Sweden is forest, while only 8% of land cover is farmland. Sweden is divided into three major regions: to the north is the vast mountain and forest region; central Sweden consists of lowland in the east and highland in the west; and southern Sweden includes highlands in Småland and agricultural plains in Skåne. The climate varies across Sweden, but it is mainly temperate in the south and subarctic in the north.

Using the GLCVal system, 382 and 394 sample points were generated for Bulgaria and Sweden, respectively, with LSI-based sampling approach at 95% confidence level, as shown in Figure 7. These sample points were judged and labeled by the experts from the Academy of Bulgaria and by research assistants at KTH Royal Institute of Technology following the guidelines of the technical specification for 30-m resolution GLC validation. Additional reference data were used, including local orthophotos, high-resolution satellite images, and national land cover data. The color orthophoto maps of Bulgaria have a spatial resolution between 0.4 and 0.5 m, acquired in 2006 and 2010–2011. For Sweden, high-resolution satellite images for the period 2009–2011 from Google Earth at a spatial resolution between 0.6 and 2.5 m were utilized for validation.

The validation result of GlobeLand30 dataset is presented as an error matrix in Tables 1 and 2. The Overall Accuracy (OA) of the GlobeLand30 in Bulgaria and Sweden are 79.84% and 88.58% (kappa: 0.87), respectively. The Producer Accuracy (PA) and the User Accuracy (UA) are presented for each of the land cover classes. For Bulgaria, the land cover classes...
with highest producer accuracy (over 80%) are the cultivated land, forest and waterbodies and artificial surfaces, while the lowest results are observed in shrubland and wetland classes. The user accuracy shows that the class with highest values is forest (92.2%), followed by the artificial areas. The cultivated land, shrubland and waterbodies are assessed with user accuracy over 70%, while lowest values are again observed for the wetland land cover class. For Sweden, almost all land cover classes achieved over 80% producer accuracy except bareland. The 100% accuracy for the water class is outstanding for Sweden as it has around 100,000 lakes with different conditions. Bareland has the lowest classification accuracy and was confused with permanent snow and ice, tundra, and artificial surfaces. This is likely due to the lack of images acquired at the suitable season to map permanent snow and ice in addition to the spectral similarities of bareland, tundra and artificial surfaces. In terms of user accuracy, all land cover classes yielded over 80% accuracy except permanent snow and ice. The low user accuracy for permanent snow and ice was caused by commission errors from bareland and tundra.

Table 1. Error matrix for GlobeLand30 dataset in Bulgaria.

| Land cover       | 10  | 20  | 30  | 40  | 50  | 60  | 80  | 90  | All | UA (%) |
|------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|--------|
| 10–cultivated land | 75  | 10  | 7   | 10  | 4   | 106 | 167 | 56.5 |
| 20–forest        | 5   | 154 | 4   | 4   | 5   | 167 | 92.2|
| 30–grassland     | 2   | 9   | 26  | 7   | 1   | 1   | 46  |
| 40–shrubland     | 1   | 3   | 2   | 23  | 1   | 1   | 29  |
| 50–wetland       | 1   | 1   | 4   | 21  | 1   | 1   | 24  |
| 60–water bodies  | 1   | 1   | 5   | 7   | 2   | 50.0|
| 80–artificial surfaces | 1  | 1   | 1   | 21  | 0   | 1   | 87.5|
| 90–bareland      | 1   | 1   | 1   | 46  | 0   | 1   | 91.7|
| All              | 84  | 177 | 46  | 4   | 2   | 6   | 26  | 382 | 0.0   |
| PA (%)           | 89.3| 87.0| 63.4| 50.0| 50.0| 83.3| 80.8| 0.0  | 0.0   |

Table 2. Error matrix for GlobeLand30 dataset in Sweden.

| Land cover       | 10  | 20  | 30  | 40  | 50  | 60  | 80  | 90  | 100 | All | UA (%) |
|------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--------|
| 10–cultivated land | 40  | 6   | 2   | 48  | 83.3|
| 20–forest        | 51  | 7   | 58  | 87.9|
| 30–grassland     | 1   | 10  | 1   | 12  | 83.3|
| 40–shrubland     | 3   | 49  | 1   | 52  | 94.2|
| 50–wetland       | 1   | 2   | 44  | 1   | 48  | 91.7|
| 60–water bodies  | 1   | 1   | 45  | 48  | 93.8|
| 70–tundra        | 1   | 37  | 2   | 40  | 92.5|
| 80–artificial surfaces | 1  | 1   | 2   | 31  | 1   | 36  | 86.1|
| 90–bareland      | 1   | 22  | 1   | 23  | 95.7|
| 100–permanent snow and ice | 5 | 4 | 20 | 29 | 69.0 |
| All              | 42  | 65  | 10  | 61  | 47  | 45  | 42  | 33  | 29  | 394 | 0.0   |
| PA (%)           | 95.2| 78.5| 100 | 80.3| 93.6| 93.6| 90.0| 75.9| 100 | 0.0 | 0.0   |

Figure 7. Distribution of the LSI-based sample points in Bulgaria (left) and Sweden (right).
6. Conclusions

A technical specification and an online tool are two essential ingredients for a successful collaborative 30-m GLC validation. From 2015 to 2017, the GEO Global Land Cover Community Activity organized a project to examine the major technical issues, develop methods and tool(s), and carry out case studies. About 30 countries and organizations joined this GEO-led international cooperative. A technical specification for 30-m GLC validation was developed by considering the findings and experiences of international experts and users. An online validation tool, GLCVal, was developed by integrating land cover validation procedures with the service computing technologies. About ten workshops and seminars were organized to discuss the methodology adopted and disseminate the results. About 20 countries have completed the accuracy assessment of GlobeLand30 for their territories with the guidance of the technical specification and the support of GLCVal.

As the world’s first 30-m GLC data product, GlobeLand30 has been used by scientists and users from more than 130 countries for environmental change analysis, geographical condition monitoring, urban and rural management, earth surface process modeling, and sustainable development. The 2020 version of GlobeLand30 was released recently for open access and was shared on the occasion of the 75th anniversary of the United Nations. It can be expected that in the future more and more GLC data products at higher spatial, temporal, and thematic resolutions will be released as a support to the United Nations 2030 Sustainable Development Goals (SDGs). The results of this GEO-led project will provide a solid basis for promoting a comprehensive and collaborative validation of these new GLC products. It will also be a good opportunity to further improve the efficacy and efficiency of global-scale land cover validation.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Data Availability Statement

The data that support the findings of this study are openly available in “Global Change Data Repository” at http://dx.doi.org/10.3974/geodb.2014.02.01.V1, http://dx.doi.org/10.3974/geodb.2014.02.02.V1, http://dx.doi.org/10.3974/geodb.2018.02.05.V1.

Funding

This work is funded by the National Natural Science Foundation of China [Grant Nos. 41930650, 41631178] and the Program of International S&T Cooperation, the Ministry of Science and Technology of China [Grant No. 2015DFA11360]. The authors would like to thank all the participants who contributed to the GEO-led 30-m resolution validation tasks.

Notes on contributors

Jun Chen is an academician of Chinese Academy of Engineering, and a professor of National Geomatics Centre of China. His research interests are land cover mapping, validation, service and application.

Fei Chen is an associate professor of East China University of Technology, Nanchang Province, China. She received the PhD degree from Central South University, China. Her research interests are land cover validation methods.

Lijun Chen is a professor in Department of Global Geo-information Resources, National Geomatics Centre of China. His research interests are land cover mapping and remote sensing image processing methods.

Yifang Ban is a professor in Division of Geoinformatics, KTH Royal Institute of Technology, Sweden. Her research interests are land cover mapping and validation and remote sensing image processing methods.

Songnian Li is a professor in the Department of Civil Engineering, Ryerson University, Toronto, Canada. His research interests are geographic information services and platforms for supporting land cover validation.

Gang Han is a senior engineer of National Geomatics Centre of China. His research interest is research and development of geographic information service website.

Xiaohua Tong is a professor of College of Surveying and Geo-Informatics, Tongji University, China. His research interests are validation methods of geographic information products.

Chuang Liu is a professor of Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, China. Her research interests are big data analysis and land cover validation.

Vanya Stamenova is a professor in Department of Remote Sensing and GIS, Space Research and Technology Institute, Bulgarian Academy of Sciences, Bulgaria. His research interests are land cover mapping and remote sensing image processing methods.

Stefan Stamenov is an associate professor in Space Research and Technology Institute, Bulgarian Academy of Sciences, Bulgaria. Her research interests are land cover mapping and remote sensing image processing methods.

ORCID

Jun Chen http://orcid.org/0000-0002-5933-0026
Lijun Chen http://orcid.org/0000-0001-7817-8595
Fei Chen http://orcid.org/0000-0002-1675-3555
Yifang Ban http://orcid.org/0000-0003-1369-3216
Songnian Li http://orcid.org/0000-0002-8244-5681
Gang Han http://orcid.org/0000-0001-5124-0635
Xiaohua Tong http://orcid.org/0000-0002-1045-3797
Chuang Liu http://orcid.org/0000-0002-4728-670X
Vanya Stamenova http://orcid.org/0000-0001-5421-8732
Stefan Stamenov http://orcid.org/0000-0001-6010-8142
References

Ban, Y. F., and A. Jacob. 2013. “Object-Based Fusion of Multitemporal Multispectral ENVISAT ASAR and HJ-1B Multispectral Data for Urban Land-Cover Mapping.” IEEE Trans. Geoscience Remote Sensing 51 (4): 1998–2006. doi:10.1109/TGRS.2012.2236560.

Brodelli, M. A., M. E. Molinari, E. Hussein, J. Chen, and J. Li. 2015. “The First Comprehensive Accuracy Assessment of GlobelAnd30 at a National Level: Methodology and Results.” Remote Sensing 7 (4): 4191–4212. doi:10.3390/rs70404191.

Chen, F., J. Chen, H. Wu, D. Y. Hou, W. W. Zhang, J. Zhang, X. G. Zhou, and L. J. Chen. 2016. “A Landscape Shape Index-based Sampling Approach for Land Cover Accuracy Assessment.” Science China Earth Sciences 59 (12): 2263–2274. DOI:10.1007/s11430-015-5280-5.

Chen, J., A. P. Liao, L. J. Chen, H. W. Zhang, C. Y. He, S. Peng, H. Wu, et al. 2014a. “Global Land Surface Water Dataset at 30m Resolution (2010) [DB/OL].” Global Change Data Repository. doi:10.3974/geodb.2014.02.V1.

Chen, J., A. P. Liao, L. J. Chen, H. W. Zhang, C. Y. He, S. Peng, H. Wu, et al. 2014b. “Global Artificial Land Surface Dataset at 30 M Resolution (2010)[db/ol], global.” Change Data Repository. doi:10.3974/geodb.2014.02.V1.

Chen, J., A. P. Liao, L. J. Chen, H. W. Zhang, C. Y. He, S. Peng, H. Wu, et al. 2017a. “Content and Composition of the Global Land Surface Water Dataset at 30 M Resolution (2010) [[J]]. Journal of Global Change Data & Discovery 1 (1): 27–39. doi:10.3974/geodb.2017.01.05.

Chen, J., A. P. Liao, L. J. Chen, H. W. Zhang, C. Y. He, S. Peng, H. Wu, et al. 2017b. “Content and Composition of Global Artificial Land Surface Dataset at 30 M Resolution (2010) [[J]].” Journal of Global Change Data & Discovery 1 (2): 136–148. doi:10.3974/geodb.2017.02.02.

Chen, J., F. Chen, H. Wu, L. J. Chen, and G. Han. 2018. “Internet+ Land Cover Validation: Methodology and Practice.” Geomatics and Information Science of Wuhan University 43 (12): 2225–2232. (in Chines).

Chen, J., J. Chen, A. Liao, X. Cao, L. Chen, X. Chen, C. He, et al. 2015. “Global Land Cover Mapping at 30m Resolution: A POK-based Operational Approach.” ISPRS Journal of Photogrammetry and Remote Sensing 103 (5): 7–27. DOI:10.1016/j.isprsjprs.2014.09.002.

Clark, M. L., and A. T. Mitchell. 2011. “Virtual Interpretation of Earth Web-Interface Tool (VIEW-IT) for Collecting Land-Use/Land-Cover Reference Data.” Remote Sensing 3 (3): 601–620. doi:10.3390/rs3030601.

Congalton, R. G. 1991. “A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data.” Remote Sensing of Environment 37 (1): 270–279. doi:10.1016/0034-4257(91)90048-B.

Defourny, P., L. Schouten, S. Bartalev, S. Bontemps, and O. Arino. 2009. “Accuracy Assessment of a 300 M Global Land Cover Map: The GlobCover Experience.” New Lib. World 112 (5–6): 236–247.

Foley, M. M., B. S. Halpern, F. Micheli, M. H. Armsby, M. R. Caldwell, C. M. Cain, E. Prahler, et al. 2010. “Guiding Ecological Principles for Marine Spatial Planning.” Marine Policy 34 (3): 955–966. doi:10.1016/j.marpol.2010.02.001.

Friedl, M. A., D. Sulla-Menashe, B. Tan, A. Schneider, N. Ramankutty, A. Sibley, X. Huang, et al. 2010. “MODIS Collection 5 Global Land Cover: Algorithm Refinements and Characterization of New Datasets.” Remote Sens. Environment 114 (1): 168–182. DOI:10.1016/j.rse.2009.08.016.

Fritz, S., I. Mccallum, C. Schill, C. Perger, L. See, D. Schepaschenko, M. Van Der Velde, et al. 2012. “Geo-Wiki: An Online Platform for Improving Global Land Cover.” Environmental Modelling & Software 31 (7): 110–123. DOI:10.1016/j.envsoft.2011.11.015.

Fritz, S., I. Mccallum, C. Schill, C. Perger, R. Grillmayer, F. Achard, F. Kraxner, and M. Obersteiner. 2009. “Geo-Wiki.Org: The Use of Crowdsourcing to Improve Global Land Cover.” Remote Sensing 1 (3): 345–354. doi:10.3390/rs1030345.
Using Remote Sensing and GIS.” International Journal of Geosciences 8 (4): 611–622. doi:10.4236/ijg.2017.84033.
See, L., C. Perger, M. Hofer, J. Weichselbaum, C. Dresel, and S. Fritz 2015. “Laco-Wiki an Open Access Online Portal for Land Cover Validation. II-3/W5.” p. 167–171.
Stehman, S. V., P. Olofsson, C. E. Woodcock, M. Herold, and M. A. Friedl. 2012. “A Global Land-cover Validation Data Set, II: Augmenting A Stratified Sampling Design to Estimate Accuracy by Region and Land-cover Class.” J Remote Sensing 33 (22): 6975–6993. doi:10.1080/01431161.2012.695092.
Stehman, S. V., and R. L. Czaplewski. 1998. “Design and Analysis for Thematic Map Accuracy Assessment: Fundamental Principles [J].” Remote Sensing of Environment 64 (3): 331–344. doi:10.1016/S0034-4257(98)00010-8.
Strahler, A., L. Boschetti, G. Foody, M. Friedl, M. Hansen, M. Herold, P. Mayaux, J. Morisette, S. Stehman, and C. Woodcock. 2008. Global Land Cover Validation: Recommendations for Evaluation and Accuracy Assessment of Global Land Cover Maps, 25. Luxembourg: European Communities.
Tong, X. H., Z. H. Wang, H. Xie, D. Liang, Z. Q. Jiang, J. Li, and J. Li. 2011. “Designing a Two-rank Acceptance Sampling Plan for Quality Inspection of Geospatial Data Products.” Computer Geoscience 37 (10): 1570–1583. doi:10.1016/j.cageo.2011.02.006.
Yang, Y., P. Xiao, X. Feng, and H. Li. 2017. “Accuracy Assessment of Seven Global Land Cover Datasets over China.” ISPRS J. Photogramm. Remote Sensing 125: 156–173. doi:10.1016/j.isprsjprs.2017.01.016.
Zell, E., A. K. Huff, A. T. Carpenter, and L. A. Friedl. 2012. “A User-Driven Approach to Determining Critical Earth Observation Priorities for Societal Benefit.” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sensing 5 (6): 1594–1602. doi:10.1109/JSTARS.2012.2199467.