ISPRS Journal of Photogrammetry and Remote Sensing 125 (2017) 156–173

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Accuracy assessment of seven global land cover datasets over China

Yongke Yang, Pengfeng Xiao *, Xuezhi Feng, Haixing Li

Department of Geographic Information Science, Nanjing University, Nanjing, Jiangsu 210023, China
Collaborative Innovation Center of South China Sea Studies, Nanjing, Jiangsu 210023, China
Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing, Jiangsu 210023, China

A R T I C L E   I N F O

Article history:
Received 28 June 2016
Received in revised form 22 December 2016
Accepted 24 January 2017
Available online 3 February 2017

Keywords:
Global land cover dataset
Spatial agreement
Accuracy validation
Spatial variation in accuracy

A B S T R A C T

Land cover (LC) is the vital foundation to Earth science. Up to now, several global LC datasets have arisen with efforts of many scientific communities. To provide guidelines for data usage over China, nine LC maps from seven global LC datasets (IGBP DISCover, UMD, GLC, MCD12Q1, GLCNMO, CCI-LC, and GlobeLand30) were evaluated in this study. First, we compared their similarities and discrepancies in both area and spatial patterns, and analysed their inherent relations to data sources and classification schemes and methods. Next, five sets of validation sample units (VSUs) were collected to calculate their accuracy quantitatively. Further, we built a spatial analysis model and depicted their spatial variation in accuracy based on the five sets of VSUs. The results show that, there are evident discrepancies among these LC maps in both area and spatial patterns. For LC maps produced by different institutes, GLC 2000 and CCI-LC 2000 have the highest overall spatial agreement (53.8%). For LC maps produced by same institutes, overall spatial agreement of CCI-LC 2000 and 2010, and MCD12Q1 2001 and 2010 reach up to 99.8% and 73.2%, respectively; while more efforts are still needed if we hope to use these LC maps as time series data for model inputting, since both CCI-LC and MCD12Q1 fail to represent the rapid changing trend of several key LC classes in the early 21st century, in particular urban and built-up, snow and ice, water bodies, and permanent wetlands. With the highest spatial resolution, the overall accuracy of GlobeLand30 2010 is 82.39%. For the other six LC datasets with coarse resolution, CCI-LC 2010/2000 has the highest overall accuracy, and following are MCD12Q1 2010/2001, GLC 2000, GLCNMO 2008, IGBP DISCover, and UMD in turn. Beside that all maps exhibit high accuracy in homogeneous regions; local accuracies in other regions are quite different, particularly in Farming-Pastoral Zone of North China, mountains in Northeast China, and Southeast Hills. Special attention should be paid for data users who are interested in these regions.

1. Introduction

Land cover (LC) is one of the indispensable variables for the study of Earth system process. It provides thematic information of the Earth's surface that captures biotic and abiotic properties closely tied to the ecological condition of land areas (Friedl et al., 2010). LC mapping has been considered as the most important agent research on environmental change and sustainability (Foley et al., 2005; Running, 2008; Gong et al., 2013; Zhang et al., 2014).

To date, nine global LC datasets have arisen from different initiatives using satellite data, i.e. (1) International Geosphere-Biosphere Program Data and Information System's LC dataset (IGBP DISCover), 1992–1993, 1 km resolution (Loveland and Belward, 1997; Loveland et al., 2000), (2) University of Maryland LC dataset (UMD), 1992–1993, 1 km resolution (Hansen et al., 2000), (3) Global LC 2000 dataset (GLC) from the European Commission's Joint Research Center (JRC), 1 km resolution (Bartholomé and Belward 2005), (4) Moderate Resolution Imaging Spectroradiometer (MODIS) LC dataset (MOD12Q1 and MCD12Q1) available at annual scale from 2001, 500/1000 m resolution (Friedl et al., 2002; Friedl et al., 2010), (5) Global LC Map (GlobCover) from European Space Agency (ESA), 2005–2006/2009, 300 m resolution (Arino et al., 2008; Bontemps et al., 2010), (6) Climate Change Initiative LC dataset (CCI-LC) from ESA, 2000/2005/2010, 300 m resolution (Defourny et al., 2016), (7) Global Map–Global LC (GLCNMO) dataset from the International Steering Committee for Global Mapping, 2003/2008, 500 m resolution (Tateishi et al., 2014), (8) Finer Resolution Observation and Monitoring Global LC dataset (FROM-GLC) from China based on Landsat images from 1984 to 2011, 30 m

* Corresponding author at: Department of Geographic Information Science, Nanjing University, Nanjing, Jiangsu 210023, China.
E-mail address: xiaopf@nju.edu.cn (P. Xiao).

http://dx.doi.org/10.1016/j.isprsjprs.2017.01.016
0924-2716/

resolution (Gong et al., 2013; Yu et al., 2013), and (9) 30 m resolution Global LC dataset (Globeland30) from National Geomatics Center of China, 2000/2010 (Chen et al., 2015).

Although these LC datasets provide a typically reflection of the Earth's surface, there must be some differences among these LC datasets derived from different satellite data, classification scheme and approach. Thus, accuracy assessment is needed for better using of these LC datasets. The literature of existing accuracy assessment works for these LC datasets are listed in Table 1. Two commonly used methods for accuracy assessment of LC datasets are (1) crossing with validations sample pixels within a confusion matrix, and (2) measuring agreement and disagreement with existing LC maps or statistical LC information. E.g., the overall accuracies were derived from confusion matrices for the following maps: 66.9% for IGBP DISCover (Scepan et al., 1999), 68.6% for GLC 2000 (Mayaux et al., 2006), 78.3% for MCD12Q1 2010 (Friedl et al., 2002), 77.9% for GLCNMO 2008 (Tateishi et al., 2014), 73.2% for CCI-LC (Defourny et al., 2016), 64.9% for FROM-GLC (Gong et al., 2013), and 80.3% for GlobeLand30 (Chen et al., 2015). Besides, some studies show that there are large areas of disagreement among these datasets, e.g. in the west of China (Giri et al., 2005; Herold et al., 2008; McCallum et al., 2006).

Many efforts have been paid by Chinese scientists to evaluate the accuracies of the global LC datasets over China. E.g., Wu et al. (2008) and Ran et al. (2010) validated IGBP DISCover, UMD, MOD12Q1 2001, and GLC 2000 using the same validation database, i.e. China's LC map of 2000 produced by Liu et al. (2003), respectively. The results of the former indicate that MOD12Q1 2001 has the best fit in depicting China's croplands. The results of the later indicate that GLC 2000 has the highest overall accuracy (59.3%). Niu et al. (2012) assessed the accuracy of wetlands in GlobCover 2009 using the wetland map of China in 2008, and found that the overall accuracy of wetlands in GlobCover 2009 is only 32.0%, and the wetland area of GlobCover 2009 is far smaller than that in reference data. Note that some global LC datasets use a regional stratification classification approach (e.g. IGBP DISCover, GLC, GLCNMO, and CCI-LC), so that their accuracy over China may not be identical elsewhere. Moreover, in the case that FROM-GLC and GlobeLand30 were produced by Chinese institutes, their accuracy over China can be expected to be higher.

Table 1
Literature of existing accuracy assessment works for global LC datasets.

Validation method	LC dataset	Literature
Crossing with validation sample pixels (measuring overall accuracy, producer's and user's accuracy, and Kappa coefficient)	IGBP DISCover	Scepan et al. (1999)
	GLC	Mayaux et al. (2006)
	GlobCover	Bicheron et al. (2008)
	CCI-LC	Defourny et al. (2016)
	GLCNMO	Tateishi et al. (2014)
	IGBP DISCover, UMD, GLC, MODIS	Latifovic and Olthof (2004), Herold et al. (2008)
	IGBP DISCover, UMD, MODIS	Gong (2009)
	FROM-GLC	Gong et al. (2013)
	GlobeLand30	Chen et al. (2015)
	MODIS	Friedl et al. (2002)
Measuring agreement/ disagreement with existing LC maps or statistical data	IGBP DISCover and UMD	Hansen and Reed (2000)
	GLC and MODIS	Fritz and See (2008), Fritz et al. (2010), Gao and Jia (2012)
	IGBP DISCover, UMD, GLC, MODIS	McCallum et al. (2006), Jung et al. (2006), Wu et al. (2008), Ran et al.(2010)
	MODIS	Sedano et al. (2005)
	GLC, MODIS, GlobCover	Tchuenté et al. (2011)
	IGBP DISCover, UMD, GLC, MODIS, GlobCover	Pérez-Hoyos et al. (2012a); Kuenzer et al. (2014)
	GlobCover	Niu et al. (2012), An et al. (2012)

Unfortunately, these existing works did not refer to the accuracy of evergreen needleleaf forests, evergreen broadleaf forests, deciduous needleleaf forests, deciduous broadleaf forests, mixed forests, cropland/natural vegetation mosaic, and snow and ice over. Since there is no available LC map of China containing these LC classes and it is expensive and time-consuming to collect new validation database. However, it is meaningful to obtain a more detailed description about the accuracies of these global LC datasets.

While valuable for guiding data usage, existing accuracy information, e.g. the accuracy indexes derived from confusion matrix, provides only a global estimation for the entire map, and does not represent the spatial variation in classification errors (Foody, 2002; Foody, 2005; Strahler et al., 2006; Comber et al., 2012). Instead of being random or stationary in space, however, the classification errors may vary within map area (Campbell, 1981; Steele et al., 1998; Foody, 2005). Many users, especially those only interested in parts of the mapped area or those using maps within spatially distributed models, may benefit from a spatial analysis of the classification accuracy (Kyriakidis et al., 1999; Strahler et al., 2006). Thus, describing the spatial variation of classification accuracy is one of our motivations for this study.

The existing methods used to describe the spatial variation in classification accuracy can be divided into two types: (1) confidence-based method (generate a strength map of the class membership) (Maselli et al., 1994; Colditz et al., 2012; Pérez-Hoyos et al., 2012b), and (2) geo-statistically based method (Congalton, 1988; Meer, 1996; Atkinson and Lewis, 2000; Kyriakidis and Dungan, 2001; Foody, 2005; Comber et al., 2012). The former method is based on the development of a confidence map to describe the spatial variation in classification errors. However, as an ancillary dataset provided by data producers, confidence map can only be obtained along with the development of the LC map. The later has been proved to be an available and meaningful method to depict the spatial variation in classification errors. E.g., Foody (2005) attempted to compute local accuracy of a LC map by geographically constraining the validation samples used for each target location. Comber et al. (2012) used a local binary logistical regression method to reveal the local accuracy. For each target location, a kernel function was applied to define validation sample and weighting used for local accuracy assessment; thus the parameter is a dynamic variables related to the geographic space rather than constants.

The goal of this study is to evaluate the similarities and differences of these global LC datasets, and their classification accuracy over China, by conducting a comparative assessment through three aspects: (1) comparing the areal and spatial agreement among different LC maps, and analysing their inherent relation with data sources and classification schemes and methods, (2) calculating the overall accuracy, user's and producer's accuracy, and Kappa coefficient for each LC maps based on the validation samples, which are collected by visual interpretation of high resolution images, and (3) depicting the spatial variations in classification accuracy for each LC maps via a geographically weighted statistical model.

2. Data and methodology

2.1. Global land cover datasets

This study refers to nine LC maps (IGBP DISCover, UMD, GLC 2000, MCD12Q1 2001, MCD12Q1 2010, GLCNMO 2008, CCI-LC