

The spatial local accuracy of land cover datasets over the Qiangtang Plateau, High Asia

LIU Qionghuan^{1,2}, *ZHANG Yili^{1,2,3}, LIU Linshan¹, LI Lanhui^{1,2}, QI Wei¹

1. Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China;
2. University of Chinese Academy of Sciences, Beijing 100049, China;
3. CAS Center for Excellence in Tibetan Plateau Earth Sciences, Beijing 100101, China

Abstract: We analyzed the spatial local accuracy of land cover (LC) datasets for the Qiangtang Plateau, High Asia, incorporating 923 field sampling points and seven LC compilations including the International Geosphere Biosphere Programme Data and Information System (IGBPDIS), Global Land cover mapping at 30 m resolution (GlobeLand30), MODIS Land Cover Type product (MCD12Q1), Climate Change Initiative Land Cover (CCI-LC), Global Land Cover 2000 (GLC2000), University of Maryland (UMD), and GlobCover 2009 (GlobCover). We initially compared resultant similarities and differences in both area and spatial patterns and analyzed inherent relationships with data sources. We then applied a geographically weighted regression (GWR) approach to predict local accuracy variation. The results of this study reveal that distinct differences, even inverse time series trends, in LC data between CCI-LC and MCD12Q1 were present between 2001 and 2015, with the exception of category areal discordance between the seven datasets. We also show a series of evident discrepancies amongst the LC datasets sampled here in terms of spatial patterns, that is, high spatial congruence is mainly seen in the homogeneous southeastern region of the study area while a low degree of spatial congruence is widely distributed across heterogeneous northwestern and northeastern regions. The overall combined spatial accuracy of the seven LC datasets considered here is less than 70%, and the GlobeLand30 and CCI-LC datasets exhibit higher local accuracy than their counterparts, yielding maximum overall accuracy (OA) values of 77.39% and 61.43%, respectively. Finally, 5.63% of this area is characterized by both high assessment and accuracy (HH) values, mainly located in central and eastern regions of the Qiangtang Plateau, while most low accuracy regions are found in northern, northeastern, and western regions.

Keywords: land cover datasets; spatial accuracy assessment; remote sensing; Qiangtang Plateau; High Asia

Received: 2018-05-27 **Accepted:** 2019-02-12

Foundation: The Strategic Priority Research Program of the Chinese Academy of Sciences, Nos. XDA20040200, XDB03030500; Key Foundation Project of Basic Work of the Ministry of Science and Technology of China, No. 2012FY111400; National Key Technologies R&D Program, No. 2012BC06B00

Author: Liu Qionghuan, PhD Candidate, E-mail: liuqionghuan@yeah.net

***Corresponding author:** Zhang Yili, Professor, specialized in physical geography and biogeography.

E-mail: zhangyl@igsnr.ac.cn

Partial results from this study have already appeared in *Geographical Research* (in Chinese); these are partially revised in this study.

1 Introduction

Land cover (LC) information is basis data for various scientific research endeavors, including environmental science and ecology. These data are also key for understanding the sustainable use of land resource as well as for government decision-making and also exert important impacts on downstream analyses. Classified LC maps have become one of the most important products of remote sensing (RS) surveys, enabling environmental and natural resources monitoring, modeling and management from local-to-global scales (Sutherland *et al.*, 2009; Estes *et al.*, 2018). In concert with the development of RS and GIS technologies, numerous global and regional LC products have emerged that utilize RS images as a data source. Indeed, more than 20 global and regional scale LC products were available and freely accessible in 2016 (Grekousis *et al.*, 2015). However, due to inconsistencies in classification systems and technologies as well as image acquisition time and spatial resolution, clear differences in accuracy are seen when different data are applied at regional or global scales (Herold *et al.*, 2008; Gong *et al.*, 2013). These discrepancies have seriously undermined the effective use of LC datasets for research in many fields (Congalton *et al.*, 2014; Tsendbazar *et al.*, 2015) and means that it is very important to analyze the accuracy of existing information in this area (Foody, 2002; Moristette *et al.*, 2002; Wulder *et al.*, 2014).

Previous researchers have chosen to utilize relative assessment methods to compare overall area and spatial consistencies as well as differences between multiple LC datasets (McCallum *et al.*, 2006; Herold *et al.*, 2008). Studies conducted to date have also compared the spatial consistency of global or regional LC datasets based on different products and with various resolutions (Fritz *et al.*, 2011; Kaptué Tchuenté *et al.*, 2011) while others have selected a particular product for reference, Global Land Cover 2000 (GLC2000), for example, to compare and assess the overall accuracy of other global datasets (Schultz *et al.*, 2015). In other research, scholars have utilized sample points for reference in order to assess and validate the accuracy of global LC datasets (Stehman, 2009, 2014; Gong *et al.*, 2013; Bai *et al.*, 2015; Lei *et al.*, 2016; Wickham *et al.*, 2017).

Error matrix values as well as those for overall accuracy (OA), producer's accuracy (PA), and user's accuracy (UA), have proven to be the most precise and therefore commonly applied statistical approaches for evaluating LC accuracy. These measures and the way in which they are calculated, have, however, been criticized in the past for not providing any indication of error spatial distribution, because non-stationary error distributions occur, for example, in the presence of heteroscedastic residual distributions (Comber *et al.*, 2012). The appropriateness of information conveyed by an error matrix may therefore be limited when specific local conditions vary. In this paper, therefore, we focused on spatial variations in classification accuracy at the level of individual pixels (Comber *et al.*, 2012; Khatami *et al.*, 2017a; Khatami *et al.*, 2017b), and the analyses presented here are also based on field points and sample blocks.

Grassland degradation, glacier retreat, lake expansions, and other eco-environmental changes on the Qinghai-Tibet Plateau have received increasing levels of scientific attention over the last 30 years (Liu *et al.*, 2008; Yao *et al.*, 2012; Nie *et al.*, 2017). Indeed, land cover change (LCC) on the Qiangtang Plateau, the hinterland of the Qinghai-Tibet Plateau, plays a key role in the regional eco-environment and is thus an important component of the overall

plateau ecosystem. It is therefore necessary for us to take full advantage of existing data as well as the accuracy of currently available numerical simulations in tandem with RS inversion data in order to develop comprehensive assessments (Zhang *et al.*, 2013). A limited number of scientific investigations have, however, been conducted in this region because of the extremely harsh natural environment of the Qiangtang Plateau, and basic research data is therefore lacking for the development of accurate spatial predictions.

The analyses presented in this study are based on 923 sampling points from across the Qiangtang Plateau as well as seven large-scale LC datasets. These datasets comprise the International Geosphere Biosphere Programme Data and Information System (IGBPDIS) compilation, Global Land cover mapping at 30 m resolution (GlobeLand30), the MODIS Land Cover Type product (MCD12Q1), Climate Change Initiative Land Cover (CCI-LC), Global Land Cover 2000 (GLC2000), data collected by the University of Maryland (UMD), and GlobCover 2009 (GlobCover). We utilized GWR to assess local spatial accuracy and sought to address a number of research questions including what is the overall spatial accuracy of LC data available for the Qiangtang Plateau? Which areas are characterized by high levels of local accuracy? And which LC data compilations are more accurate for particular local areas on the plateau?

2 Study area

The Qiangtang Plateau is located between 32°10'–36°32'N and 79°42'–92°05'E and encompasses the northern flank of the Gangdise-Nyainqentanglha mountains, the southern Kunlun Mountains, the eastern Karakoram Mountains, and the western Tanggula Mountains (Figure 1). A series of complex landforms and a cold-dry climatic regime characterizes this region. The Qiangtang Plateau has an average altitude of 5025 m, ranging between 4153 m and 6854 m, an annual rainfall of 154.9 mm (Li *et al.*, 2017), and an average temperature below zero degree centigrade (Yao *et al.*, 2015). This region encompasses a vast area of alpine desert grassland, being characteristic as simple vegetation type and low fractional cover of vegetation. In addition, the vegetation of the Qiangtang Plateau is extremely fragile because of the harsh weather conditions that characterize this region, and once threatened, this biotope cannot be restored.

3 Data sources and methods

3.1 Sampling points

The sampling points used for this analysis are uniformly distributed across the entire region of interest. However, because of the uneven distribution of each class and field condition limitations, we are unable to guarantee that the same numbers of sampling points occur within each category. We utilized a total of 923 sampling points, of which 163 are derived from previous research (Gong *et al.*, 2013; Yang *et al.*, 2017), alongside 152 water bodies data points based on Google Earth images and Panoramio photographs that are each 2 km by 2 km in size. We also included 110 glacier samples extracted from the Second Glacier Inventory Dataset of China published by the Cold and Arid Regions Scientific Data Center (Guo *et al.*, 2014) that are also 2 km by 2 km in size. The remaining 498 sample points used

here are the result of field investigations carried out by our research team between 2013 and 2016. Our sample spatial distribution is summarized in Figure 1 while detailed descriptions of these data are presented in Table 1.

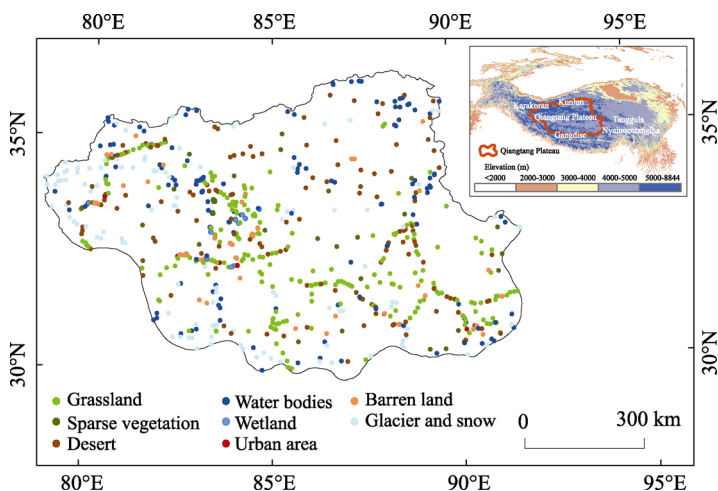


Figure 1 Map showing the study area discussed in this analysis as well as the distribution of sampling points on the Qiangtang Plateau

3.2 Land cover datasets

Although more than 20 global LC products are currently available within 10 data series (Grekousis *et al.*, 2015), seven were selected for this analysis (Table 2). Five of these data series, GlobeLand30, CCI-LC, GlobCover, MCD12Q1, and Global Land Cover 250 m China (GLC250 m_CN), comprise 22 datasets for different years, and one (GLC250 m_CN) is not currently shared publicly. We therefore selected four datasets that capture the most recent years from the remaining four compilations above (i.e., those released publicly prior to December 2016) alongside three others that comprise different series (i.e., UMD, IGBPDIS, and GLC2000) that have been widely applied for long-term LCC research. However, the International Satellite Land Surface Climatology Project Initiative II (ISLSCP II) dataset was not utilized as it actually comprises 12 distinct datasets. A further available data series (Geo-Wiki) was also not selected for analysis because of inherent temporal uncertainties. The LCCS (Land Cover Classification System) includes FAO (Food and Agriculture Organization, FAO), USGS (United States Geological Survey, USGS), IGBP, Simplified IGBP and UN (United Nations, UN).

3.3 Land cover data processing

Projections based on the seven LC datasets utilized here were initially converted to Albers projection (Albers_Conic_Equal_Area). Next, because the classification criteria used for LC types were different in each case and therefore restrict direct comparisons, correspondence was assessed using the approach developed by Giri (2005), Herold (2008), and Ran (2010) to establish linkage between different classification systems. We referred to existing linkages between classification systems (Giri *et al.*, 2005; Ran *et al.*, 2010; Kaptué Tchuenté *et al.*,

Table 1 Class description of field sample points over the Qiangtang Plateau*

LC type	Number of sample points	Definition	LC type	Number of sample points	Definition
Grassland	219	Region mainly covered with a community of cold-tolerant perennial herbaceous plants. In this region, this LC type mainly comprises meadows covered with <i>Kobresia littledalei</i> and <i>K. pygmaea</i> , a plant coverage area dominated by taxa with some cold tolerance, especially xerophytic perennial herbaceous species.	Wetland	80	Broad areas covered with herbs or woody plants, usually transitional zones between land and water.
Sparse vegetation	53	Plants comprise continuous vegetation that extends up to the permanent snow line, as well as a zone encompassing coverage between 5% and 40% of total surface that consists of cold-adapted plants such as cold habitat perennial axis-shaft root grasses, cushion plants, and lichens. One area, for example, contains cushion plants such as <i>Arenaria serpyllifolia</i> and <i>Androsace tapete</i> .	Urban area	26	Land covered with buildings.
Desert	45	Desert areas are widely distributed in this region and are characterized by the presence of semi-shrubs and their dwarf counterparts (e.g., <i>Ceratocarpus latens</i> , <i>Ajania pallasiana</i>) as well as <i>C. compacta</i> .	Barren land	142	Barren land, sand, rock, and saline areas with vegetation coverage less than, or equal to, 10%.
Water bodies	152	Long-strip depressions that naturally form along the ground surface as well as land below the perennial water level developed under natural conditions.	Glacier and snow	110	Land that is perennially covered with snow or ice.
Total	923				

*We selected locations for sampling based on vegetation type investigations in order to encompass uniform land cover, neighborhood consistency, and good representation. Representative spatial range varies according to the distribution of LC types, however; in the case of large areas of grassland or desert, a circular area with a radius between 500 m and 1,000 m was selected and sampling was carried out at a central point. Sampling points in each case can therefore be represented by a circular area between 30 m and 50 m in diameter.

2011) in combination with the specific definitions utilized by various data products and field investigations to develop a LC type classification system of the Qiangtang Plateau that is based on a sample point survey encompassing eight primary types. The corresponding relationships between our classification system and the four utilized in the seven datasets we evaluated are discussed elsewhere (Liu *et al.*, 2017). Finally, we resampled and matched cell locations between CCI-LC and MCD12Q1, CCI-LC/GlobCover and MCD12Q1, and CCI-LC/MCD12Q1 and GlobeLand30, in order to compare similarities and discrepancies in LC dataset spatial patterns.

3.4 Spatial variation analysis of boolean classification data

The number of samples (or patches) is always far smaller in actuality than the number of pixels within an entire study area. At the same time, however, random single point accuracy

Table 2 LC dataset characteristics

Dataset	OA (%)	Verification method	Sensor	Classification method	Resolution	Time	Classification system (Number of types)	URL for download	Reference
GLC2000	68.6	Confidence value statistical sampling	SPOT4 VEGETATION	Unsupervised classification	1 km	1999-2000	FAO LCCS (23 classes)	http://bioval.jrc.ec.europa.eu/products/glc2000/products.php	Bartholom <i>et al.</i> , 2005
IGBPDIS	66.9	Statistical sampling by validation working group	AVHRR	Unsupervised classification	1 km	1992-1993	USGS IGBP (17 classes)	http://edc2.usgs.gov/glcc/tabguide_globe.php	Loveland <i>et al.</i> , 2000
UMD	65.0	Evaluated using other digital datasets	AVHRR	Unsupervised classification, decision tree classification	1 km	1992-1993	Simplified IGBP (14 classes)	http://www.landcover.com/data/landcover/index.shtml	Hansen <i>et al.</i> , 2000
MCD12Q1	74.8	Cross-validation	MODIS	Supervised classification, decision tree classification, neural network	500 m	2001-2016	IGBP (17 classes)	http://e4ftl01.cr.usgs.gov/MOTA/MCD12Q1.006/	Friedl <i>et al.</i> , 2010, 2011
GlobCover	67.5	Statistical sampling expert's judgement	MERIS FR	Supervised classification, unsupervised classification	300 m	2009	UN LCCS (22 classes)	http://due.esrin.esa.int/globcover/	Bon-temps <i>et al.</i> , 2011
CCI-LC	74.1	Sampling-based labeling approach	MERIS Full and Reduced Resolution/ SPOT	Unsupervised classification	300 m	1992-2015	UN LCCS (22 classes)	http://maps.elie.ucl.ac.be/CCI/viewer/index.php	Belgium <i>et al.</i> , 2016
GlobeLand30	80.0	Knowledge-based interactive verification	Landsat TM, ETM7, HJ-1A/b/	Integration of pixel- and object-based methods with knowledge (pok-based)	30 m	2000, 2010	11 classes	http://www.globallandcover.com	Chen <i>et al.</i> , 2015

is both accidental and often wrong. For these reasons, spatial analysis can be used to determine how single, whole-map global measures of accuracy vary. We utilized GWR to estimate OA at specific locations (i, j) with no existing samples calculated. A GWR approach is similar to an ordinary logistic regression model but coefficient estimates are allowed to vary geographically using a kernel function (Comber *et al.*, 2012). A moving window then allows a local regression analysis to be computed in each case with points further away from the specific location under consideration contributing less to the solution.

Local spatial accuracy based on Equation (1) is weighted OA based on geographically constrained samples. This means that Equation (2) enables samples that are further away from a given location (i, j) to less significantly influence accuracy, rather than making an equal contribution. In other words, a weighted regression is performed out in cases where the weight $(W_{k(i,j)})$ associated with each location (i, j) is a decreasing function of d_i , the

distance from the window center to (i, j) (Comber *et al.*, 2012; Yang *et al.*, 2017). These relationships are expressed as follows:

$$OA_{(i,j)} = \frac{\sum_1^N bW_{k(i,j)}}{\sum_1^N W_{k(i,j)}} * 100\% \quad (1)$$

$$W_{k(i,j)} = \begin{cases} \left(1 - \frac{d_i^2}{h^2}\right)^2, & \text{if } d_i < h \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

where $OA_{(i,j)}$ denotes local spatial accuracy, while $W_{k(i,j)}$ is the contribution made by the reference point to location (i, j) , d_i refers to the distance between the reference point and location (i, j) , b is a binary variable, and 1 indicates that the sample type is the same as that approximated by LC data at the reference point (otherwise b will be denoted 0). In addition, N denotes the number of samples used for local accuracy assessment at location (i, j) where h refers to the bandwidth of GWR calibration. This latter measure can either be fixed or adaptive to a subset of the nearest n sample points and different kernel functions (shapes) can be used for distance weighting. As larger bandwidths generally result in a greater degree of spatial smoothing (Gollini *et al.*, 2015), an adaptive value of 15% was applied in Bi-squared kernel function.

We also applied an estimated kappa coefficient to further explore accuracy in this study, with $Kappa_{(i,j)}$ defined as follows (Comber *et al.*, 2017):

$$Kappa_{(i,j)} = \frac{N \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_{i+} * x_{+i})}{N^2 - \sum_{i=1}^r (x_{i+} * x_{+i})} \quad (3)$$

where N denotes the total number of observations in the matrix, r is the number of rows, x_{ii} denotes the number of observations in row i and column i , while x_{i+} and x_{+i} are the marginal totals of row i and column i , respectively.

4 Results

4.1 Spatial distribution characteristics and area comparisons

Field surveys and related literature (CIG, 1988; Zheng, 1999; Ding *et al.*, 2015) show that while grasslands, deserts, and barren land are the main primary LC types on the Qiangtang Plateau, the former is the most widely distributed (Figure 2). The data presented in Figure 2 shows that grassland, barren land, deserts, and water bodies can be clearly differentiated.

Data from CCI-LC, GlobeLand30, and GLC2000 suggest that grassland covers the entire Qiangtang Plateau, while MCD12Q1 data imply a mainly southeastern distribution (Figure 2). In GlobCover, IGBPDIS, and UMD datasets, the spatial distribution of grassland is much smaller than suggested by the other four datasets. The spatial difference is most apparent in

terms of barren land which is mainly distributed in the northwestern transitional grassland in MCD12Q1, but appears in the middle of the GlobCover data. The proportion of barren land remains small and are scattered across the Qiangtang Plateau in other LC datasets, in contrast. The desert type in IGBPDIS is mainly distributed in the northeast, coincident with the spatial distribution seen in MCD12Q1, while the LC type covers almost the entire Qiangtang Plateau based on UMD data. The visual characteristics of water bodies are mainly distributed in the south, southeast, and northwest of the study area, and no significant regional differences are apparent between the seven LC datasets.

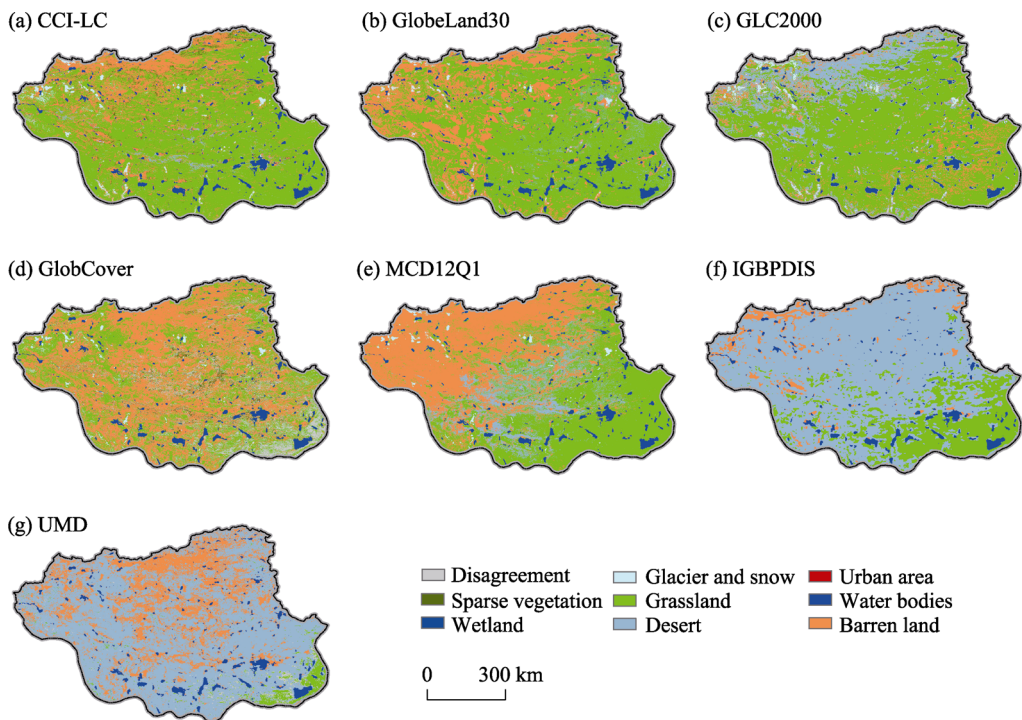


Figure 2 Map showing the spatial distribution of primary LC types within seven Qiangtang Plateau datasets

We compared the area occupied by each LC class in all datasets (Figure 3). The comparison shows that all LC types encompass areas less than $1 \times 10^4 \text{ km}^2$, with the exception of sparse vegetation and urban areas. The largest grassland area was recovered on the basis of GlobeLand30 data, as much as 21 times the smallest area recovered on the basis of UMD data. The discrepancy of grassland area is very similar between GlobeLand30, CCI-LC, and GLC2000 data sources as well as between GlobCover and MCD12Q1 compilations, respectively. In contrast, however, grassland areas reconstructed on the basis of GlobCover and MCD12Q1 data are much less than that estimated on the basis of CCI-LC, GlobeLand30, and GLC2000 compilations. However, the area of barren land estimated based on GlobCover and MCD12Q1 datasets is much higher than in other cases. The area of barren land is directly related to the vegetation canopy definition threshold for this LC type. Results show that the threshold vegetation canopy for barren land is less than 20% based on GlobCover

data, wider than that for other LC datasets which are less than 10%.

Largest areas of desert are found in the study area based on IGBPDIS data (Figure 2) even though the main LC class should actually be barren land. Based on IGBPDIS and UMD data, desert areas are four times larger than in other cases. This overestimation of desert based on IGBPDIS and UMD data is strongly related to the very weak spectral differences between barren land and desert LC types and is even less differentiation on coarser resolution images. Unsurprisingly, variance in the area of water bodies reconstructed from the seven LC datasets is much smaller than other classes as the LC type is easy to classify. Areas of glacier and snow are also close to one another in most datasets, with the exception of IGBPDIS and UMD. One reason is the fact that UMD was designed based on the Simple Biosphere (SiB) model which does not include snow and glacier types. The areas of glacier types in both IGBPDIS and UMD data are closely related to the over-mapping of barren land in other datasets.

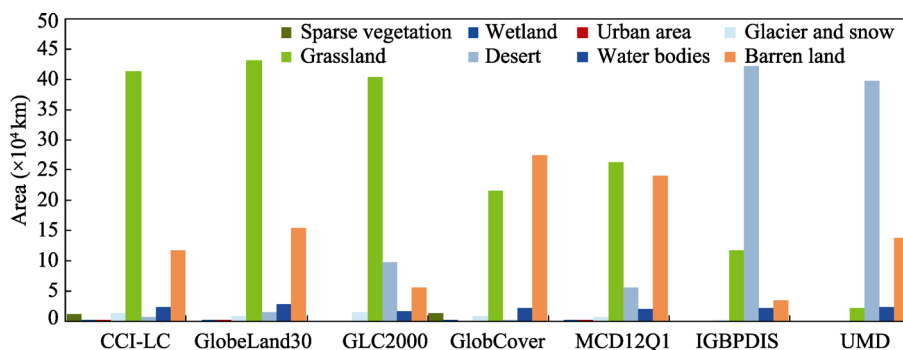


Figure 3 LC class areas in each of the seven Qiangtang Plateau datasets

Grassland area based on CCI-LC data generally exhibits an upward trend although an overall downward trend was seen in MCD12Q1 data. Results show that although both LC datasets are characterized by an upward trend between 2004 and 2010, rates are different in both cases. Grassland area based on CCI-LC data increased rapidly but remained stable and slightly decreased when based on MCD12Q1 data. These outcomes both contrast with recent research that suggests no obvious changes in the above ground net primary productivity of grassland between 1983 and 2014 on the Tibetan Plateau (Liu *et al.*, 2018).

At the same time, however, changes in barren land and wetland area based on CCI-LC data are in accordance with results from MCD12Q1 over the period between 2001 and 2010. These results both imply a rapidly decreasing trend, although the rate based on MCD12Q1 data was faster than CCI-LC. Changes in barren land within this area also conform to an inverse trend between 2010 and 2015. Specially, the area continues to decrease rapidly based on CCI-LC data while switching from a reduction to an increase when based on MCD12Q1 data with a turning point around 2010.

Trends and rates of change in water bodies between the two datasets are not obviously different, but the total area of water bodies in CCI-LC is higher than that in MCD12Q1.

Areas of glacier and snow based on CCI-LC data did not change between 2001 and 2015,

a result that is obviously in conflict with the result of general retreat of glaciers (Yao *et al.*, 2012). It is therefore clear that CCI-LC data exhibits weaker classification accuracy when used to reconstruct glacier areas. In contrast, MCD12Q1 data suggest that the areas of glaciers and snow rose rapidly within the study area between 2001 and 2009, another result which conflicts with our knowledge about glacial retreat (Yao *et al.*, 2012).

Urban land types based on MCD12Q1 data did not change between 2001 and 2015, in conflict with enhanced human activities on the Qiangtang Plateau (Venter *et al.*, 2016). The reason is that the distribution of urban land is not concentrated and the patch is broken. It remains difficult to identify smaller-scale construction land in coarse resolution images. It is the case that prior to 2013, the area of urban land reconstructed based on CCI-LC data was small and no distinct change was evident. Since 2013, however, the area of this LC type has risen dramatically.

MCD12Q1 data suggest the absence of sparse vegetation types across the study area while CCI-LC data suggests a rapid increase in this LC type between 2001 and 2008. Indeed, subsequent to 2008, the area of sparse vegetation across this region increased rapidly but decreased more and more rapid prior to this inflection point. The result is coincident with the accelerated reinforcement of the grazing pressure within this area as sparse vegetation is influenced to a greater extent by the factors (Figure 4).

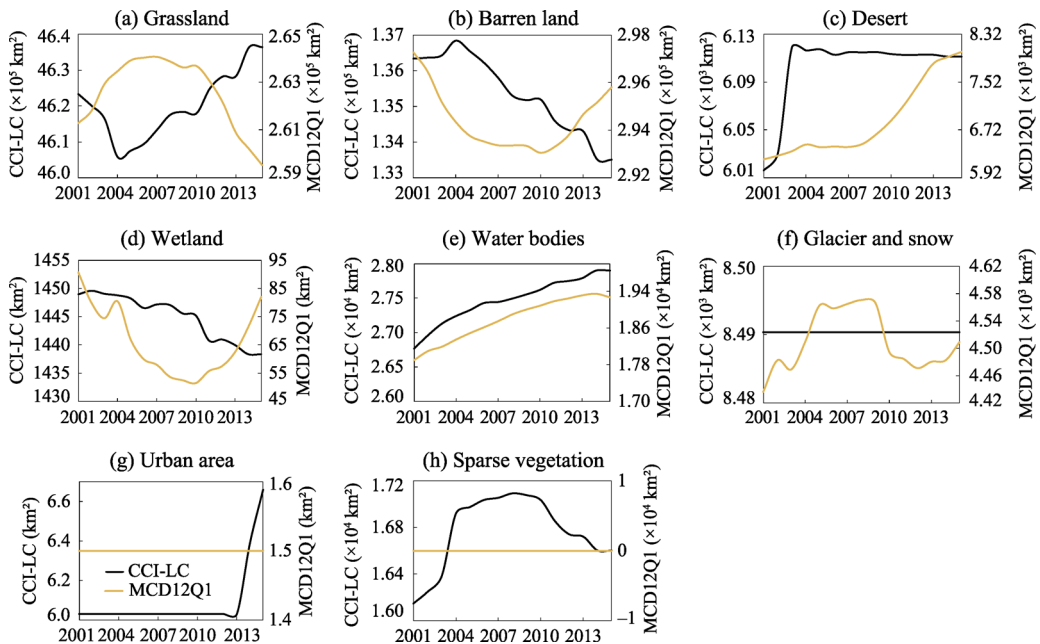


Figure 4 Time series changes of eight LC classes based on CCI-LC and MCD12Q1 data over the Qiangtang Plateau

Discrepancies between MCD12Q1 with CCI-LC LC data can also be explained by differences in classification definitions in thematic legends. It is the case that legend class descriptions are not always unique or totally clear and can therefore result in classification un-

certainties. Unclear LC class descriptions can also severely hamper the conversion and joining of legends between the seven LC products and therefore further result in map uncertainties and losses in correspondence.

4.2 Spatial agreements and disagreements

Results show that high levels of spatial agreement occur mainly in the homogeneous south-eastern region of the Qiangtang Plateau. Reconstructions based on MCD12Q1 (2009 and 2010), CCI-LC (2009 and 2010), MCD12Q1/CCI-LC, and GlobeLand30 data all exhibit high levels of agreement within this region as it mainly contains grasslands and lakes that are easy to differentiate compared with other LC types. However, results also show that GlobCover data exhibit a low level of accuracy in this region, and that levels of consistency between MCD12Q1/CCI-LC and GlobCover reconstructions are similarly low.

Results also show that low levels of spatial agreement are widely distributed across heterogeneous northwestern and northeastern parts of the Qiangtang Plateau. These regions mainly contain barren land and grassland. Indeed, grasslands within alpine areas are sparsely distributed and are usually withered in the autumn and winter. These features are very hard to distinguish from images and are often easily confused with barren land.

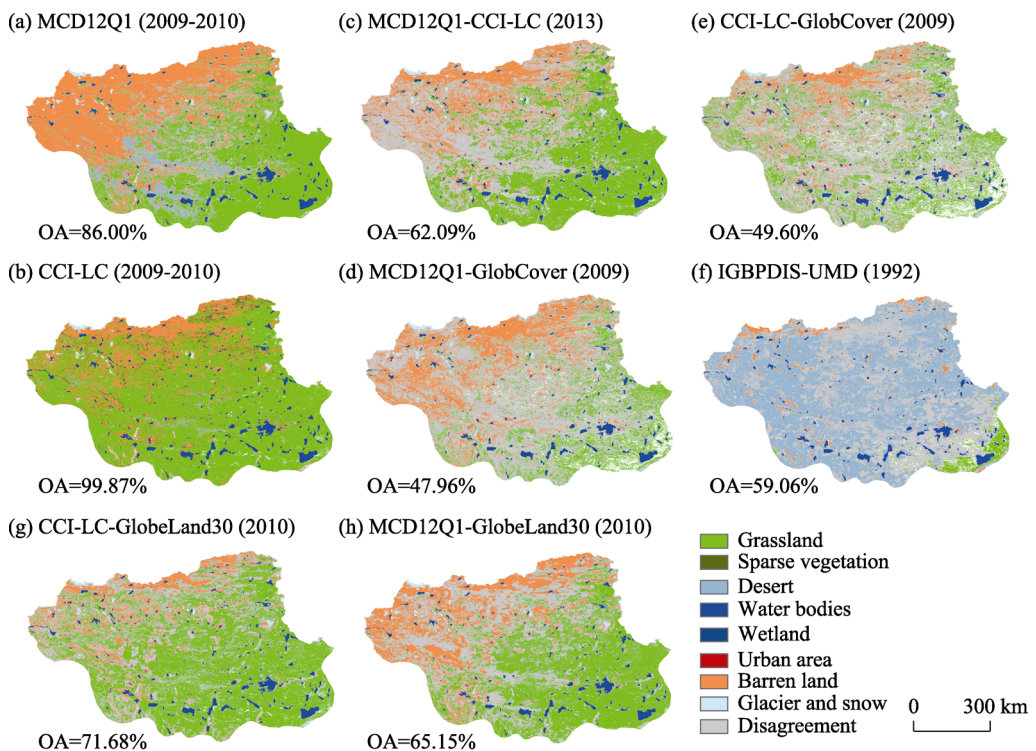


Figure 5 Maps showing spatial agreements and disagreements between different LC datasets over the Qiangtang Plateau

The results of this study show that the consistency of data for adjacent years within the same LC dataset series tends to be higher, for example, CCI-LC (2009 and 2010) and

MCD12Q1 (2009 and 2010). Indeed, between the two LC datasets, the consistency of CCI-LC (99.87%) is higher than that based on MCD12Q1 (86.00%). It is the case that regions characterized by low MCD12Q1 data consistency tend to mainly contain barren land and grassland. As these two LC types are easily mixed, the fact that MCD12Q1 data is based on a single MODIS image means that this source is less accurate than CCI-LC which is developed based on multiple image sources. Data show that CCI-LC and GlobeLand30 data in their consistency are relatively highly in agreement (71.68%) within different series, even though these values are generally low amongst in other compilations.

4.3 Analysis of spatial variation accuracy

4.3.1 Assessing the accuracy of land cover datasets

Taking the seven LC datasets as a set, an assessment value greater than equal to five implies a high data assessment (H), but when this value is less than or equal to four, data assessment is low (L). The accuracy value interval was defined the same as for assessment.

The data presented in Figure 6 shows that amongst the 923 sample points, the ratio between low assessment and low accuracy (LL) is 77.14% while the ratio between high assessment and high accuracy (HH) is just 5.63%, less than the ratio between high assessment and low accuracy (HL) (17.23%). This means that from the perspective of different HL types, the ratio between desert and barren land is 4.12%, greater than either water bodies (3.79%) or grassland (2.71%). These results are indicative of a more widespread phenomenon that assessment is high across the field area but that accuracy is very low. Therefore, we can know that there is no consensus between assessment and accuracy.

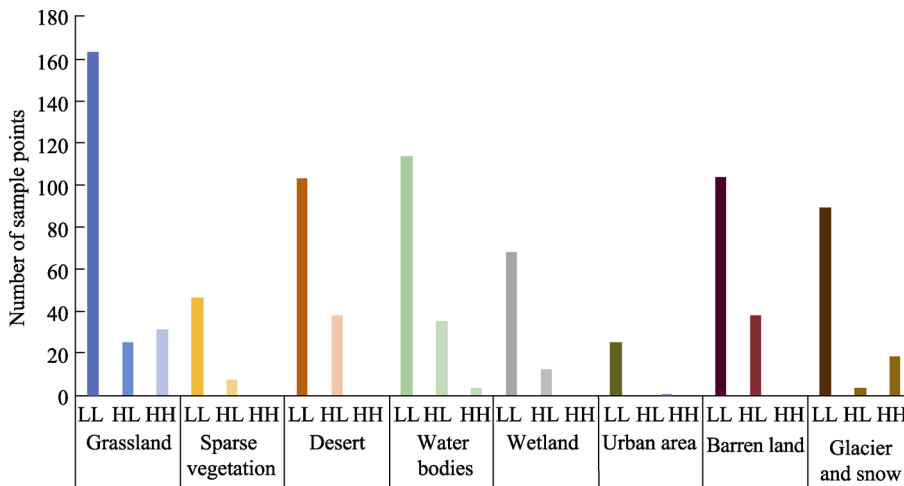


Figure 6 Assessment and accuracy frequencies of different LC categories over the Qiangtang Plateau

4.3.2 The spatial accuracy of boolean classification datasets

Local accuracy derived from our use of the GWR varies markedly across the study area. Data show that OA values for the seven LC datasets considered here range between 20% and 60% while their local accuracy values range from zero to 100%. Mean OA suggests that the

highest local accuracy value seen within the seven LC datasets is for GlobeLand30 (42.08%), followed by CCI-LC (34.92%). The order of other sets is GLC2000 (34.55%), IGBPDIS (31.42%), MCD12Q1 (28.12%), GlobCover (27.62%), and UMD (26.47%). Maximum OA values also vary between the seven LC datasets, and the order of maximum OA values for datasets is GlobeLand30 (77.39%), GlobCover (74.11%), GLC2000 (66.12%), CCI-LC (61.43%), MCD12Q1 (59.76%), IGBPDI (50.76%), and UMD (47.07%). Data show that discrepancies between the seven LC datasets in terms of both mean and maximum kappa coefficients are almost the same as OA values (Table 3).

The highest OA values recovered in this analysis occur in the center (H1) and the south-east (H2) of the Qiangtang Plateau, both areas are characterized by higher proportions of lakes and grassland. It is also generally the case that OA values exhibit an apparent relationship with degree of landscape heterogeneity. This region is generally characterized by relatively low landscape complexity as it is adjacent to an alpine area and human influence remains relatively weak. At the same time, data accuracy in the north (L1), northeast (L2), and west (L3) of the Qiangtang Plateau is relatively low across our datasets because of the higher landscape complexity of these regions. Large areas of desert, barren land, and grassland are also found within the L2 region. Results show that desert and barren LC types are often misclassified as grassland and that this is one key factor influencing classification accuracy. Indeed, accuracy values recovered for the southern (U) region of the Qiangtang Plateau are also quite uncharacteristic as both local maximum and minimum precision record are distributed within this area. Numerous lakes, marshy wetlands, and grasslands occur within the U region and the first two of these are easily confused.

It is not surprising that landscape heterogeneity has been regarded as the major factor contributing to the accuracy of LC datasets. This variable can significantly complicate our understanding of ground surface spectral characteristics and make RS classifications increasingly problematic. Landscape heterogeneity explains the marked discordance between data products reported in this analysis. The results show that CCI-LC, GlobeLand30, GLC2000, and GlobCover data are all highly locally accurate for the H1 region, but that MCD12Q1, IGBPDIS, and UMD data have low accuracy values. Precision of GlobCover and UMD data is low across the H2, while the other five LC datasets considered here yield high local accuracy values. Our results show that GLC2000 and IGBPDIS data are highly precise for the L1 region while the other five LC datasets yield low local accuracy values. Barren lands and desert are the dominant LC class in this region but were seriously misclassified as grassland based on CCI-LC and MCD12Q1 data. Results show that CCI-LC, GlobeLand30, and GlobCover data can be used to classify a small area of the L3 region with high precision as the altitude is high in L3 area and so most LC is labeled as glaciers. Classifications of remaining areas based on the other four LC datasets generate low accuracy values.

Our synthesis of OA and kappa coefficients for the seven LC datasets shows that areas characterized by higher data classification accuracy on the Qiangtang Plateau are mainly located centrally alongside a small region in the east. Kappa coefficients show that the LC area classified with a high degree of accuracy is small and mainly occurs in the center of the Qiangtang Plateau (Figure 8).

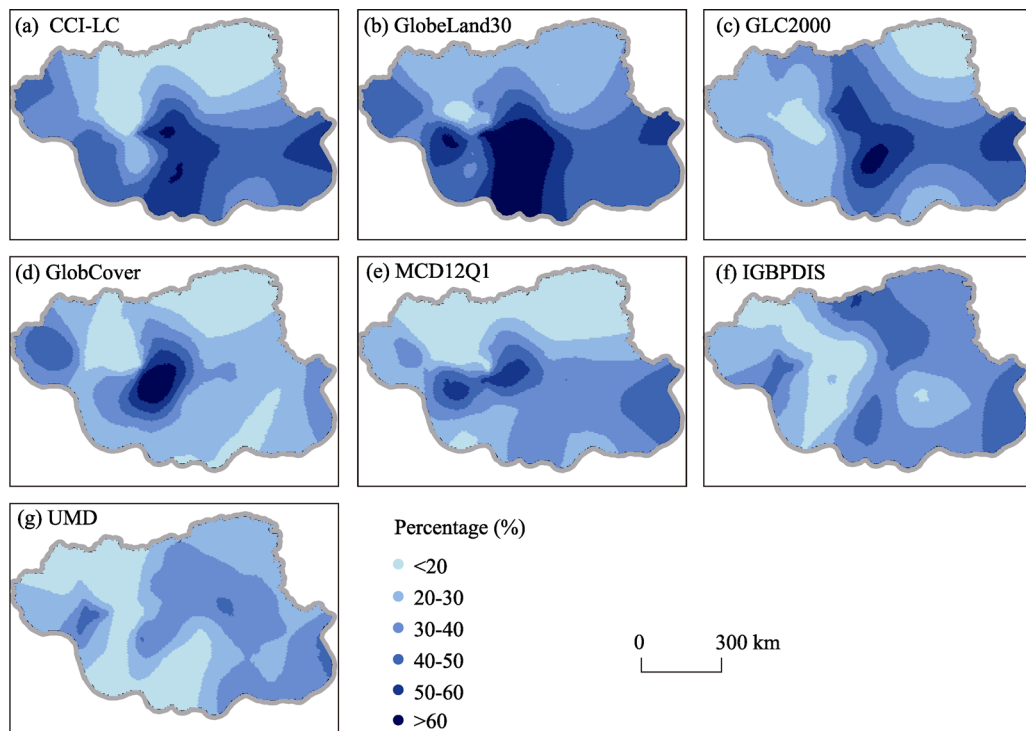


Figure 7 Maps showing OA spatial variation within the seven LC datasets over the Qiangtang Plateau

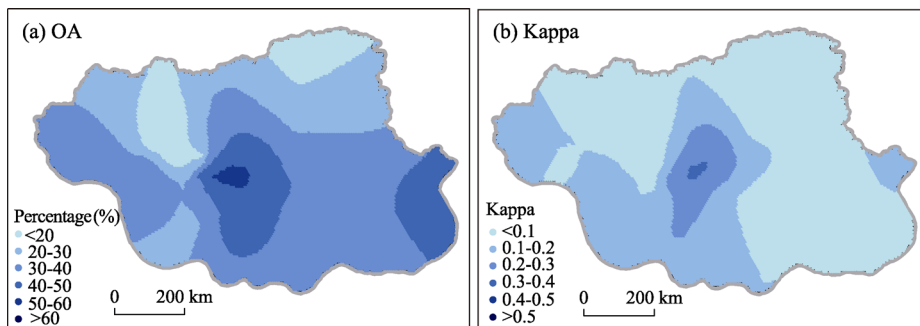


Figure 8 Maps showing the spatial distributions of OA values (a) and kappa coefficients (b) over the Qiangtang Plateau

Table 3 OA values and kappa coefficients for the seven LC datasets over the Qiangtang Plateau

Datasets	Mean OA	Max OA	Min OA	Mean kappa	Max kappa	Min kappa
GLC2000	34.55%	66.12%	12.27%	0.11	0.32	0.00
IGBPDIS	31.42%	50.76%	9.07%	0.04	0.26	0.00
UMD	26.47%	47.07%	6.53%	0.04	0.18	0.00
MCD12Q1	28.12%	59.76%	3.69%	0.05	0.37	0.00
GlobCover	27.62%	74.11%	3.52%	0.08	0.40	0.00
CCI-LC	34.92%	61.43%	2.9%	0.15	0.45	0.00
GlobeLand30	42.08%	77.39%	9.83%	0.21	0.61	0.00
Summary	32.17%	52.05%	12.82%	0.10	0.31	0.00

5 Discussion

The results of this analysis show that data accuracy was influenced by the conversion between different classification systems as well as the principles and connotations of these categories which result in different interpretations of RS data across the study area (Zhang *et al.*, 2017). Comparing our results with those previously reported by Yang *et al.* (2017), we agree that spatial accuracy values for CCI-LC and GLC2000 Qiangtang Plateau datasets are higher than those derived from either UMD or IGBPDIS compilations (Yang *et al.*, 2017). At the same time, the OA values reported by Yang *et al.* (2017) are much higher than our results. This discrepancy can be explained by the fact that we used eight LC types in this analysis while Yang *et al.* (2017) utilized just three, that is, grasslands, wetlands, and waterbodies. It is also noteworthy that the bandwidth parameter used for the spatial accuracy evaluation method applied here is based on the adaptive value of a connected area, which means that the rougher the classification, the larger the bandwidth, and therefore more sample points of the same LC type will lead to a higher value. Classification accuracy shows large deviations are also the result of limited field measurement data. Future investigations and corrections will be needed in order to more precisely design classification systems.

The interpolation method used here to predict accuracy at unsampled locations should also take more factors into account in future work, especially terrain characteristics instead of just spatial distance. Additional research will therefore be necessary to examine whether, or not, error maps can be used directly to improve classifications. Indeed, as these comprise classification errors, it might be possible to subtract them from corresponding LC maps and generate enhancements. This will require ensuring that the error maps utilized are of acceptable quality.

The vegetation types and environmental characteristics of the Qiangtang Plateau significantly influence LC data classifications. The results of this analysis show that the vegetation types in this region that are particular to Qiangtang Plateau (i.e., grasslands, barren lands, and desert) exhibit limited differences in spectra, image color, and texture which can easily result in confusions over their classification. The Qiangtang Plateau environment itself also exerts a significant impact on LC data classifications as the vegetation growing season here is mostly restricted to the summer and is much shorter than in low altitude areas (Oteros *et al.*, 2013; Zhang *et al.*, 2013). This means that it is difficult to monitor vegetation types outside of the short growing season and images within LC datasets may have been captured at different times of the year. An image acquired outside of the growing season is likely to be much less accurately interpreted.

The fact that the sample points evaluated in this analysis were collected subsequent to RS data might also influence accuracy assessments. The sample data used here were collected in field investigations between 2013 and 2016, while the seven LC datasets were acquired over 20 years between 1992/1993 and 2013. Field investigations show that the Qiangtang Plateau mostly comprises uninhabited regions and is mainly influenced by natural factors such as precipitation, temperature, evaporation, radiation, and soil moisture, which means that overall variation in LC types is actually quite small. Indeed, because of extreme environmental conditions, the sample point data fields collected for this study are mainly from areas where there is acceptable ground transportation and are thus not evenly distributed spatially. The

data presented here are nevertheless an important component of the first basic investigation on this region and therefore do provide an important reference for LC classification accuracy.

6 Conclusions

We report a number of clear spatial discrepancies amongst LC datasets in this study. High spatial agreements within the field area are mainly seen within the homogeneous southeastern region of the Qiangtang Plateau that is covered by grassland and lakes, while low spatial agreements are widely distributed across heterogeneous northwestern and northeastern areas covered with deserts, barren land, and grasslands. Data show that deserts and barren lands are more difficult to identify correctly, and that water bodies are easily confused with grasslands. We demonstrate clear time series changes discrepancy in LC data (with the exception of waterbodies) over the period between 2001 and 2015 comparing CCI-LC and MCD12Q1 compilations and highlight the existence of opposite trends in barren lands and wetlands. These results mean that LC changes across this region might not be the result of external factors in all cases, but rather due to data errors.

Calculating ratios of assessment and accuracy based on sample points and LC datasets shows that the area proportion of HL (17.23%) is higher than HH (5.63%). Therefore, it is unreasonable to use assessment method to determine classification accuracy even though low corresponding values are synchronous.

Spatial local accuracy prediction results show that OA values for the seven LC datasets range between 20% and 60%. Specifically, local accuracy values for GlobeLand30 (mean OA: 42.08%; max OA: 77.39%; mean kappa: 0.21; max kappa: 0.61) and CCI-LC (mean: OA 34.92%; max OA: 61.43%; mean kappa: 0.15; max kappa: 0.45) are higher compared with other datasets. High OA values are mainly located in the center and southeastern regions of the Qiangtang Plateau where grasslands and lakes occur, while low OA values tend to be observed in the northern, northeastern, and western regions where a mixture of deserts, barren land, and grassland are seen.

Acknowledgements

We wish to thank Zhao Zhilong, Gu Changjun, Zheng Haipeng, and Wang Yukun of the Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, for helping us to collect field sample points over the course of this study. We also thank Xiao Pengfeng and Yang Yongke of Nanjing University for providing us with access to their sample point data.

References

- Bai Y, Feng M, Jiang H *et al.*, 2015. Validation of land cover maps in China using a sampling-based labeling approach. *Remote Sensing*, 7(8): 10589–10606.
- Bartholomé E, Belward A S, 2005. GLC2000: A new approach to global land cover mapping from Earth observation data. *International Journal of Remote Sensing*, 26(9): 1959–1977.
- Belgium U. Land Cover CCI Product User Guide Version 2, Report, leuven. University Catholique de Louvain.

- Bontemps S, Defourny P, Bogaert E V *et al.*, 2011. GLOBCOVER 2009-Products description and validation report, Report. Leuven: University Catholique de Louvain.
- Changchun Institute of Geography, CAS (CIG, CAS), 1988. Tibetan Vegetation. Beijing: Science Press. (in Chinese)
- Chen J, Chen J, Liao A *et al.*, 2015. Global land cover mapping at 30m resolution: A POK-based operational approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, 103: 7–27.
- Comber A, Brunsdon C, Charlton M *et al.*, 2017. Geographically weighted correspondence matrices for local error reporting and change analyses: Mapping the spatial distribution of errors and change. *Remote Sensing Letters*, 8(3): 234–243.
- Comber A, Fisher P, Brunsdon C *et al.*, 2012. Spatial analysis of remote sensing image classification accuracy. *Remote Sensing of Environment*, 127: 237–246.
- Congalton R G, Gu J, Yadav K *et al.*, 2014. Global land cover mapping: A review and uncertainty analysis. *Remote Sensing*, 6(12): 12070–12093.
- Ding M, Li L, Zhang Y *et al.*, 2015. Start of vegetation growing season on the Tibetan Plateau inferred from multiple methods based on GIMMS and SPOT NDVI data. *Journal of Geographical Sciences*, 25(2): 131–148.
- Estes L, Chen P, Debats S *et al.*, 2018. A large-area, spatially continuous assessment of land cover map error and its impact on downstream analyses. *Global Change Biology*, 24(1): 322–337.
- Foody G M, 2002. Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 80(1): 185–201.
- Friedl M A, Sulla Menashe D, 2011. Note to users of MODIS Land Cover (MCD12Q1) Products, Report, Washington District of Columbia: NASA.
- Friedl M A, Sulla-Menashe D, Tan B *et al.*, 2010. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, 114(1): 168–182.
- Fritz S, See L, McCallum I *et al.*, 2011. Highlighting continued uncertainty in global land cover maps for the user community. *Environmental Research Letters*, 6(4): 44005.
- Giri C, Zhu Z, Reed B, 2005. A comparative analysis of the Global Land Cover 2000 and MODIS land cover data sets. *Remote Sensing of Environment*, 94(1): 123–132.
- Gollini I, Lu B, Charlton M *et al.*, 2015. GWmodel: An R package for exploring spatial heterogeneity using geographically weighted models. *Journal of Statistical Software*, 63(17): 1–50.
- Gong P, Wang J, Yu L *et al.*, 2013. Finer resolution observation and monitoring of global land cover: First mapping results with Landsat TM and ETM+ data. *International Journal of Remote Sensing*, 34(7): 2607–2654.
- Grekousis G, Mountrakis G, Kavouras M, 2015. An overview of 21 global and 43 regional land-cover mapping products. *International Journal of Remote Sensing*, 36(21): 5309–5335.
- Guo W Q, Xu J L, Liu S Y *et al.*, 2014. The Second Glacier Inventory Dataset of China (version 1.0). Lanzhou, China: Cold and Arid Regions Science Data Center.
- Hansen M C, Defries R S, Townshend J R *et al.*, 2000. Global land cover classification at 1 km spatial resolution using a classification tree approach. *International Journal of Remote Sensing*, 21(6/7): 1331–1364.
- Herold M, Mayaux P, Woodcock C E *et al.*, 2008. Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets. *Remote Sensing of Environment*, 112(5): 2538–2556.
- Kaptué Tchuenté A T, Roujean J, De Jong S M, 2011. Comparison and relative quality assessment of the GLC2000, GLOBCOVER, MODIS and ECOCLIMAP land cover data sets at the African continental scale. *International Journal of Applied Earth Observation and Geoinformation*, 13(2): 207–219.
- Khatami R, Mountrakis G, Stehman S V, 2017a. Mapping per-pixel predicted accuracy of classified remote sensing images. *Remote Sensing of Environment*, 191: 156–167.
- Khatami R, Mountrakis G, Stehman S V, 2017b. Predicting individual pixel error in remote sensing soft classification. *Remote Sensing of Environment*, 199: 401–414.
- Lei G, Li A, Bian J *et al.*, 2016. Land cover mapping in southwestern China using the HC-MMK approach. *Remote Sensing*, 8(3054): 1–22.
- Li L, Liu Q, Zhang Y *et al.*, 2017. Spatial distribution and variation of precipitation in the Qiangtang Plateau.

- Geographical Research*, 36(11): 2047–2060. (in Chinese)
- Liu H, Mi Z, Lin L *et al.*, 2018. Shifting plant species composition in response to climate change stabilizes grassland primary production. *Proceedings of the National Academy of Sciences*, 115(16): 4051–4056.
- Liu J, Xu X, Shao Q, 2008. Grassland degradation in the "Three-River Headwaters" region, Qinghai Province. *Journal of Geographical Sciences*, 18(3): 259–273.
- Liu Q, Zhang Y, Liu L *et al.*, 2017. Accuracy evaluation of the seven land cover data in Qiangtang Plateau. *Geographical Research*, 36(11): 2061–2074. (in Chinese)
- Loveland T R, Reed B C, Brown J F *et al.*, 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing*, 21(6/7): 1303–1330.
- McCallum I, Obersteiner M, Nilsson S *et al.*, 2006. A spatial comparison of four satellite derived 1 km global land cover datasets. *International Journal of Applied Earth Observation and Geoinformation*, 8(4): 246–255.
- Morissette J T, Privette J L, Christopher O *et al.*, 2002. A framework for the validation of MODIS land cover products. *Remote Sensing of Environment*, 83(1/2): 77–96.
- Nie Y, Sheng Y, Liu Q *et al.*, 2017. A regional-scale assessment of Himalayan glacial lake changes using satellite observations from 1990 to 2015. *Remote Sensing of Environment*, 189: 1–13.
- Oteros J, García-Mozo H, Vázquez L *et al.*, 2013. Modelling olive phenological response to weather and topography. *Agriculture, Ecosystems & Environment*, 179: 62–68.
- Ran Y, Li X, Lu L, 2010. Evaluation of four remote sensing based land cover products over China. *International Journal of Remote Sensing*, 31(2): 391–401.
- Schultz M, Tsensdbazaz N E, Herold M *et al.*, 2015. Utilizing the Global Land Cover 2000 reference dataset for a comparative accuracy assessment of 1 km global land cover maps. *ISPRS – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40(7): 503–510.
- Stehman S V, 2009. Sampling designs for accuracy assessment of land cover. *International Journal of Remote Sensing*, 30(20): 5243–5272.
- Stehman S V, 2014. Estimating area and map accuracy for stratified random sampling when the strata are different from the map classes. *International Journal of Remote Sensing*, 35(13): 4923–4939.
- Sutherland W J, Adams W M, Aronson R B *et al.*, 2009. One hundred questions of importance to the conservation of global biological diversity. *Conservation Biology*, 23(3): 557–567.
- Tsensdbazar N E, De Bruin S, Herold M, 2015. Assessing global land cover reference datasets for different user communities. *Isprs Journal of Photogrammetry and Remote Sensing*, 103: 93–114.
- Venter O, Sanderson E W, Magrach A *et al.*, 2016. Sixteen years of change in the global terrestrial human footprint and implications for biodiversity conservation. *Nature Communications*, 7: 12558.
- Wickham J, Stehman S V, Gass L *et al.*, 2017. Thematic accuracy assessment of the 2011 National Land Cover Database (NLCD). *Remote Sensing of Environment*, 191: 328–341.
- Wulder M A, Coops N C, 2014. Make Earth observations open access. *Nature*, 513(7516): 30–31.
- Yang Y, Xiao P, Feng X *et al.*, 2017. Accuracy assessment of seven global land cover datasets over China. *Isprs Journal of Photogrammetry and Remote Sensing*, 125: 156–173.
- Yao T, Thompson L, Yang W *et al.*, 2012. Different glacier status with atmospheric circulations in Tibetan Plateau and surroundings. *Nature Climate Change*, 2(9): 663–667.
- Yao Y, Zhang B, 2015. The spatial pattern of monthly air temperature of the Tibetan Plateau and its implications for the geo-ecology pattern of the Plateau. *Geographical Research*, 34(11): 2084–2094. (in Chinese)
- Zhang G, Zhang Y, Dong J *et al.*, 2013. Green-up dates in the Tibetan Plateau have continuously advanced from 1982 to 2011. *Proceedings of the National Academy of Sciences*, 110(11): 4309–4314.
- Zhang M, Ma M, De Maeyer P *et al.*, 2017. Uncertainties in classification system conversion and an analysis of inconsistencies in global land cover products. *ISPRS International Journal of Geo-Information*, 6(4): 112.
- Zhang Y, Wang Z, Wang X *et al.*, 2013. Land cover changes in the key regions and self reflection on ecological construction of the Tibetan Plateau. *Natural Journal*, (3): 187–192.
- Zheng D, 1999. Physical Geography of Karakorum-Kunlun Mountains. Beijing: Science Press. (in Chinese)