UNCERTAINTY ASSESSMENT OF GLOBELAND30 LAND COVER DATA SET OVER CENTRAL ASIA

Bo Sun [a, *], Xi Chen [a, b], Qiming Zhou [a, c]

[a] Center for Geo-spatial Information, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, 1068 Xueyuan Blvd., Nanshan district, 518055, Shenzhen, Guangdong, P. R. China
[b] Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, 818 South Beijing Road, 830011, Urumqi, Xinjiang, P. R. China
[c] Department of Geography, Hong Kong Baptist University, Kowloon Tong, Hong Kong, P. R. China

KEY WORDS: GlobeLand30, land cover, remote sensing, accuracy, Central Asia

ABSTRACT:

GlobeLand30, the world's first 30m-resolution global land cover data set, has recently been issued for research on global change at a fine resolution. Given the accuracy of GlobeLand30 data may show significant variation in different parts of the world and data quality at continental scale has not been validated yet, this study aims to evaluate the uncertainty of the data over Central Asia. Since it is difficult to get long-term historical ground references, GlobeLand30 data at the most recent epoch (*i.e.*, GlobeLand30-2010) was assessed. In the test, a large sample size was adopted, and more than 25 thousand samples were selected by a random sampling scheme and interpreted manually as ground references based on higher resolution imagery at the same epoch, such as images from ZY-3 (China Resources Series) satellite and Google earth. Cross validation of image interpretation by three well-trained interpreters was adopted to make the references more reliable. Error matrix and *Kappa* coefficient were utilized to quantify data accuracies in terms of classification accuracy. Results show that the GlobeLand30-2010 data presents an overall accuracy of 46% in the study area. As for specific land cover types, bare land illustrates a high user's accuracy but a lower producer's accuracy. At the same time, the accuracies of grassland and forest are significantly lower than other types. The majority of misclassification types come from bare land. It implies a difficulty of distinguishing grassland or forest from bare land in the study area. In addition, the confusion between shrub land and grassland also results in the misclassification. The results serve as a useful reference of data accuracy for further analysis of land cover change in Central Asia as well as the applications of GlobeLand30 data at a regional or continental scale.

1. INTRODUCTION

Central Asia located in the hinterland of the Eurasian Continent is a typical continental inland arid region. It has a harsh arid environment and its eco-system is very fragile. Land cover change in the arid zone of Central Asia has been increasingly regarded as a sensitive indicator to echo global environmental changes (Sun & Zhou, 2016).

Mapping land cover change with remote sensing at a large spatial scale have been studied for the past decades. Due to the limitation of remote sensing technology on the balance between imagery coverage and spatial resolution, most land cover products at the global scale present coarse spatial resolutions, e.g., 300 m or much coarser. The most common products include Global Land Cover Classification (1981-1994) issued by The University of Maryland (UMD) (Hansen *et al.*, 2000), MODIS Land Cover Type Yearly L3 Global 500 m (MCD12Q1) issued by National Arial Space Agency (NASA) of the United States, GLC2000 Global issued by the European Commission's Joint Research Centre (JRC), and GlobCover Land Cover Maps (v2.2, v2.3) issued by European Space Agency (ESA).

GlobeLand30, the world's first 30-m-resolution global land cover data set, has been issued recently by National Geomatics Center of China (NGCC) for the research on global change at a finer spatial resolution (Chen *et al.*, 2014a; 2014b). Accurate land cover data set is important for a range of applications such as land cover change detection, environmental monitoring, and change process modelling (Wardlow and Callahan, 2014). Although a

satisfied overall accuracy is reported for the global-scale data set, the accuracy may show significant variation at different areas of the world when adopting different sampling schemes. Especially in arid zone, previous studies are seldom focused on the areas with small population. Deficiency of ground-based observations or ground reference data may lead to evaluation bias in accuracy assessment of land cover classification.

Since data quality has not been validated yet at continental scale, this study aims to evaluate GlobeLand30 data over the arid region of Central Asia. Accurate land cover information plays a key role in better understanding of environmental change process in the arid region. In addition, this is also a vital step for evaluating the result of land cover change detection.

2. METHODOLOGY

2.1 Study Area

Central Asia, referring to the central part of the Asia, is located in the hinterland of the Eurasian Continent. It is far away from oceans and generally has an arid and semi-arid environment. In this study, the five former Soviet Union countries are chosen as the study area, i.e., Kazakhstan, Tajikistan, Turkmenistan, Uzbekistan and Kyrgyzstan (Figure 1). The whole region covers an area of around 4 million km^2. Population size is 62.8 million in 2010 (Chen and Zhou, 2015).

Overall topography of Central Asia comprises highland in the southeast and lowland in the northwest. The Pamir Plateau in

---

* Corresponding author, E-mail: sun.bo@msn.com

eastern Tajikistan, connected with the Tianshan Mountains, is the highest point with an elevation of between 4,000 m to 7,500 m. The climate in Central Asia is primarily influenced by the Westerly. This area has a typical interior arid and semi-arid climate. The driest region is located in the southern part. Precipitation mainly occurs as relief rainfall in the eastern mountainous regions.
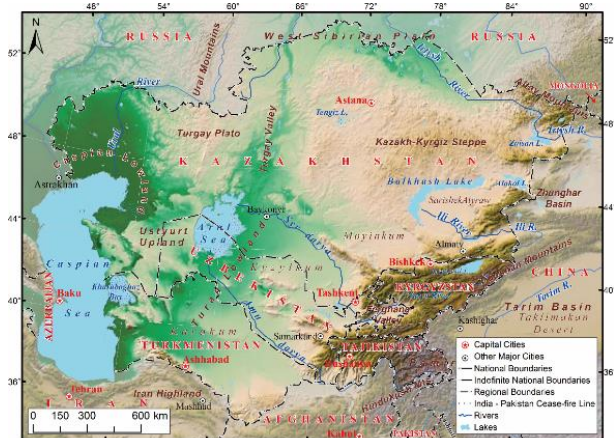

Figure 1. The location map of Central Asia (Source from: Chen and Zhou, 2015)

2.2 GlobeLand30 data

Given the difficulty of getting long-term historical ground reference data, GlobeLand30 at the most recent epoch (*i.e.* GlobeLand30-2010) is evaluated in this study. The GlobeLand30-2010 data is generalized from satellite-based images including Landsat TM and ETM+ multispectral images and Chinese Environmental Disaster Alleviation Satellite (HJ-1) CCD multispectral images. Cloudless images over vegetation growing seasons within the time frame from 2009 to 2011 are utilized. The classification system contains ten land cover types (refer to Table 1), and a high accuracy of 83% at the global scale has been declared officially (NGCC, 2014).

Ground reference data for accuracy assessment rely much on manual interpretation of high spatial resolution images when lacking ground survey data. However, some land cover types are difficult to distinguish by naked eyes. To reduce the confusion errors, seven merged classes from the original land cover types are adopted, namely cultivated land, forest, grassland, water body, artificial surfaces, bare land and permanent snow and ice. The translation table is listed as follows (Table 1). The experimental data covering the entire region of Central Asia is split out from the original GlobeLand30 data (Figure 2).

Table 1. Translation table for class merge

Original type/Code[*]	Translated type/Code
Cultivated land (10)	Cultivated land (1)
Forest (20)	Forest (2)
Grassland (30)	Grassland (3)
Shrub land (40)	Forest (2)
Wetland (50)	Water body (4)
Water body (60)	Water body (4)
Artificial s*urfaces (80)*	Artificial surfaces (5)
Bare land (90)	Bare land (6)
Permanent snow and ice (100)	Permanent snow and ice (7)

* Land cover type of *Tundra (70)* in the original classification system is not available for the study area.
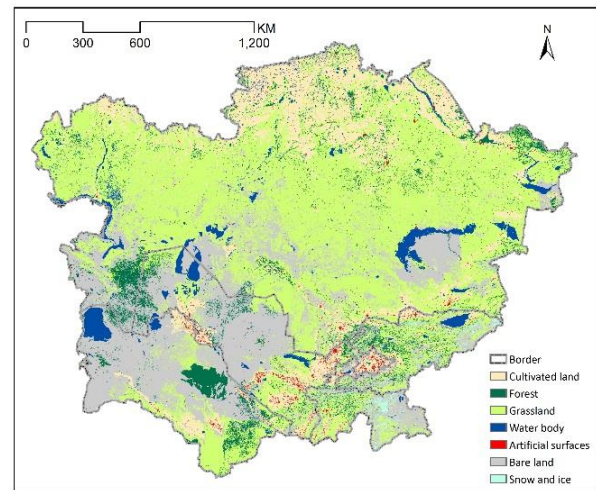

Figure 2. Land cover classification of Central Asia in 2010 (reproduced from the GlobeLand30-2010 data with map projection transform).

2.3 Sampling of references

The collection of reference data is generally based on a stratified random sampling method. A two-tier sampling scheme is adopted. For the first tier, samples are randomly selected according to land cover in the study area. Apart from that, a so called high-density sampling is employed to get dense samples at the areas covered by complex ground features.

More than 27 thousand sampling points over the study area are collected (Figure 3) and interpreted manually based on high spatial resolution imagery acquired at the same epoch as the GlobeLand30-2010 data. High-resolution images include China resources series (ZY-3) satellite and Google earth imagery. Cross validation of image interpretation by three well-trained interpreters is conducted so as to make the references more reliable. A software package for data validation is developed to assist in the process of image interpretation (refer to Appendix). Validated ground references with an agreement of at least two interpreters are utilized.
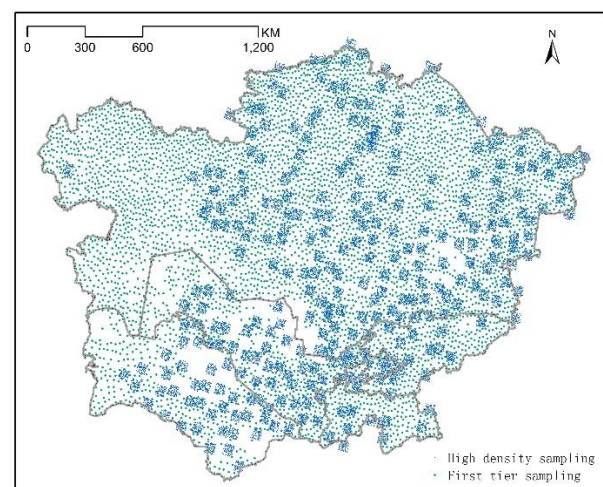

Figure 3. Sampling scheme for accuracy assessment

2.4 Accuracy assessment

Error matrix and *Kappa* coefficient are adopted to quantify data quality in terms of classification accuracy at the pixel level. An error matrix is prepared, on a category-by-category basis, to provide a basic description of land cover classification accuracy. Evaluation indicators include overall accuracy, producer's and user's accuracies. The *Kappa* coefficient is calculated based on the error matrix by the following equation (1) (Lillesand *et al.*, 2015). Besides, area statistics are compared with current major land cover products.

$$Kappa = \frac{N \sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r}(x_{i+} \cdot x_{+i})}{N^2 - \sum_{i=1}^{r}(x_{i+} \cdot x_{+i})} \qquad (1)$$

where
r = number of rows in error matrix
x_{ii} = number of observations in row i and column i
x_{i+} = total of observations in row i
x_{+i} = total of observations in column i
N = total number of observations included in matrix

3. RESULTS AND ANALYSIS

3.1 Test samples

Since ground references are cross-validated from individual judgements by three persons, the consistency of three-time interpretation results are tested (*Kappa* = 0.65, 0.85, 0.80, respectively, comparing to the validated result). With the agreement of at least two interpreters, more than 25 thousand ground references are finally confirmed for the assessment. The proportions of various land cover types of test samples are illustrated in Figure 4.
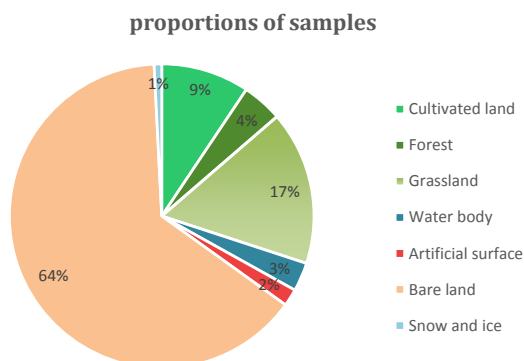
proportions of samples

Figure 4. Proportions of cross-validated samples by land cover type.

3.2 Pixel-based accuracies

Table 2 illustrates the error matrix based on validated ground references versus the GlobeLand30-2010 data. The result indicates an overall accuracy of 46% with *Kappa* coefficient of 0.283 for the study area.

As for specific land cover types, accuracy varies obviously. Bare land shows a high user's accuracy (93%) but a low producer's accuracy (35%). In contrast, cultivated land shows a lower user's accuracy (48%) but a high producer's accuracy (92%). Besides, accuracies of grassland (22%) and forest cover (26%) are significantly lower than other types.

Table 2. Error matrix and accuracy measures

		C	F	G	W	A	B	S	total
	C	2166	83	958	23	94	1167	2	4493
	F	8	363	239	32	6	710	14	1372
	G	113	539	2600	118	50	8468	66	11954
GlobeLand30-2010	W	1	15	77	540	2	94	0	729
	A	57	8	16	1	296	27	0	405
	B	7	67	220	52	10	5649	68	6073
	S	0	0	1	0	0	16	52	69
	total	2352	1075	4111	766	458	16131	202	

The header row above the columns reads "Ground references*".

Producer's accuracy:
C = 2166/2352 = 92%
F = 363/1075 = 34%
G = 2600/4111 = 63%
W = 540/766 = 70%
A = 296/458 = 65%
B = 5649/16131 = 35%
S = 52/202 = 26%

User's accuracy:
C = 2166/4493 = 48%
F = 363/1372 = 26%
G = 2600/11954 = 22%
W = 540/729 = 74%
A = 296/405 = 73%
B = 5649/6073 = 93%
S = 52/69 = 75%

Overall accuracy = 11666/25095 = 46%
Kappa coefficient = 0.283

* Land cover types: C, cultivated land; F, forest; G, grassland; W, water body; A, artificial surface; B, bare land; S, permanent snow and ice.

3.3 Area comparison with other land cover products

Area comparison with current major land cover products is shown in Figure 5. Reference products include GlobCover-2009 issued by ESA (GC-2009 for short) and Global Land Cover Classification issued by UMD (GLC-90s for short), whose spatial resolutions are 300 m and 1,000 m, respectively. Given different classification systems are utilized for different products, land cover types are unified in the analysis so as to make area statistics comparable.
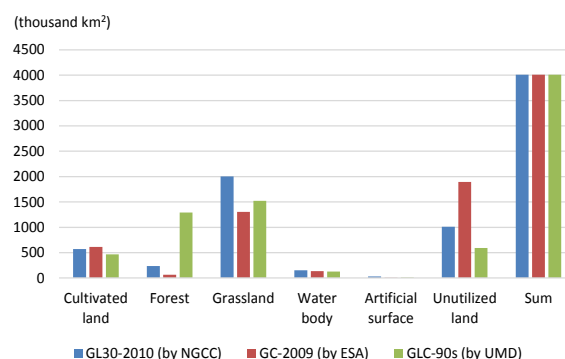

Figure 5. Area statistics and comparison with other major land cover products in Central Asia

According to the results, area statistics of different types from the three products do not agree well except cultivated land and water body. Obvious differences can be seen in land cover types of forest, grassland, and unutilized land. It should be pointed out that the UMD data describes land cover situation at a different period (i.e., from 1980s to the early 1990s) from the other two products (i.e., 2009 or 2010).

4. DISCUSSION

4.1 Sampling considerations

For getting an unbiased result, random sampling has been used to collect reference data in accuracy assessment. It is simple, and ensures all classes are adequately represented (Foody, 2002). Meanwhile, sample size is important. A large sample size allows more accurate statistics especially in a large area. Generally following a random sampling scheme, dense samples are also taken from high resolution images in the areas with complex landscape. Besides, some scholars also indicates that the number of samples for each categories might be adjusted to highlight the relative important categories in particular applications (Lillesand, 2015).

4.2 Accuracy at a continental scale

Although a satisfied accuracy has been reported for the GlobeLand30-2010 data at the global scale, the result illustrates a low overall accuracy at the continental scale in Central Asia. Moreover, classification accuracies show a large variation by land cover type.

From the error matrix in Table 2, bare land including sandy and rocky deserts and unutilized lands presents a high user's accuracy; while its producer's accuracy is rather low. It means that the classification of GlobeLand30 data may underestimate the area of bare land in the study area. On the other side, very low accuracies come from grassland and forest cover types. To take the misclassification into account, bare land is the major error source. It implies that distinguishing grassland and forest from bare land can be a hard task in the arid lands of Central Asia. In addition, misclassification also comes from the difficulty of distinguishing shrub land from grassland, even on the ground for this area, let along on images with the sight from space.

4.3 Uncertainties in accuracy assessment

A lower accuracy in this study indicates the consistency between ground references and land cover product is not good. In accuracy assessment, we assume ground reference data as the ground truth. However, one should realize that uncertainties still exist in reference data. Potential sources of uncertainties in the accuracy assessment come from various aspects.

4.3.1 Positional accuracy: Map projection transform or geometric correction may bring positional errors at pixel level. For example, an acceptable error range is within one pixel.

4.3.2 Type definition of references: Ambiguous definition or description of land cover type may cause a wrong classification of ground reference data. For example, determining an area with "vegetation cover lower than 10%" or "higher than 5%" is confused.

4.3.3 Seasonal effect: Judgements of land cover types may be different when data acquired at different periods. For example, an ephemeral stream in arid areas might be identified as "water body" or "bare land" in different seasons.

4.3.4 Scale issues: Mixed-pixel problem is very common in remote sensing image classification. Land cover classification of GlobeLand30 data is based on 30 m resolution images. While, determination of land cover type for the reference data is based on high-resolution images (6 m or above).

5. CONCLUSIONS

Assessing the accuracy of land cover data is a vital step before using them in various applications. Although existing global-scale land cover products have been declared a satisfied classification accuracy, variation may still exist at different areas of the world. This study has evaluated the accuracy of the world's first 30-m resolution land cove product - GlobelLand30 in Central Asia. A large sample size from a two-tier sampling scheme is adopted. Cross validation from three well-trained image interpreters is employed to make ground references more reliable. Result shows an overall accuracy of 46 present for the GlobeLand30-2010 data in the study area. The major error comes from the confusion between bare land and grassland. The research results can serve as a useful reference of data accuracy for further improvement of land cover classification as well as the applications of GlobeLand30 data at a regional or continental scale.

REFERENCES

Chen, J., Ban, Y., Li, S., 2014a. China: Open access to Earth land-cover map, *Nature*, 514(7523), pp. 434.

Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M., Zhang, W., Tong, X., Mills, J., 2014b. Global land cover mapping at 30m resolution: A POK-based operational approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, 103, pp. 7-27.

Chen, X., & Zhou, Q. (Eds.), 2015. *Ecological and Environmental Remote Sensing in Arid Zone – a case study in Central Asia*. Science Press, Beijing, pp. 1-2.

Foody, G.M., 2002. Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 80, pp. 185-201.

Hansen, M., DeFries, R., Townshend, J.R.G. & Sohlberg, R., 2000. Global land cover classification at 1km resolution using a decision tree classifier, *International Journal of Remote Sensing*, 21, pp. 1331-1365.

Lillesand, T.M., Kiefer, R.W., & Chipman, J.W., 2015. *Remote Sensing and Image Interpretation (Seventh edition)*. John Wiley & Sons, New Jersey.

National Geomatics Center of China (NGCC), 2014. 30-meter Global Land Cover Dataset (GlobeLand30) – Product

Description. Data available in: http://www.globallandcover.com/GLC30Download/index.aspx (1 Sept. 2015)

Olofsson, P., Foody, G.M., Stehman, S.V., Woodcock, C.E., 2012. Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sensing of Environment*, 129, pp. 122-131.

Sun, B., & Zhou, Q., 2016. Expressing the spatio-temporal pattern of farmland expansion in arid regions. *Journal of Arid Environments*, 124, pp. 118-127.

Wardlow, B.D., & Callahan, K., 2014. A multi-scale accuracy assessment of the MODIS irrigated agriculture data-set (MIrAD) for the state of Nebraska, USA. *GIScience & Remote Sensing*, 51(5), pp. 575-592.

APPENDIX

Illustration of ground reference validation.

A.   Validation with Google earth imagery
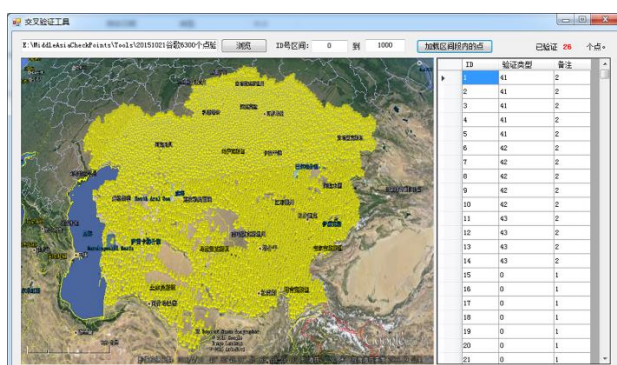
B.   Validation with high-resolution imagery (ZY-3)