# Comparison and assessment of large-scale land cover datasets in China and adjacent regions

**YANG Yongke, XIAO Pengfeng, FENG Xuezhi, LI Haixing, CHANG Xiao, FENG Weiding**

1. *Department of Geographic Information Science, Nanjing University, Nanjing* 210023, *China*;
2. *Jiangsu Provincial Key Laboratory of Geographic Information Science and Technology, Nanjing University, Nanjing* 210023, *China*

**Abstract:** Large-scale land cover datasets comprise an important foundation of the research on land surface processes, ecosystem assessment, and environmental modeling. The evaluation of existing land cover datasets provides a guide to dataset use and new dataset production. Five kinds of global land cover datasets (IGBP DISCover, UMD, GLC2000, MOD12Q1, and GlobCover 2005) over China and adjacent regions are evaluated in this paper. First, the categories of five land cover datasets are translated into the International Geosphere-Biosphere Programme-IGBP scheme based on the correlation coefficients of the corresponding classes, which is computed according to the class definition in each land cover dataset. Second, the spatial agreements of the five land c over datasets are analyzed using visual comparison and per-pixel comparison. Finally, the classification accuracy of the five land cover datasets is evaluated based on validation samples collected through Google Earth high-resolution satellite images. The results show large areas of disagreement among the five land cover datasets, and the overall consistency among them is low. GLC2000 has the highest overall accuracy and Kappa coefficient, whereas GlobCover 2005 has the lowest overall accuracy and Kappa coefficient.

**Key words:** Land cover datasets, comparison analysis, accuracy assessment, Google Earth

**CLC number:** TP79 **Document code:** A

**Citation format:** **Yang Y K, Xiao P F, Feng X Z, Li H X, Chang X and Feng W D. 2014. Comparison and assessment of large-scale land cover datasets in China and adjacent regions. Journal of R emote Sensing, 1 8(2): 453 – 475 [DOI: 10.11834/jrs.20143055]**

## 1 INTRODUCTION

Earth surface is a complex synthesis comprising numerous land cover types. Thus, a precise and detailed description of land cover features and dynamics is crucial for many studies; such features include energy balance, carbon cycle, and biogeochemical cycle of the earth system (Chen, et al., 2011).

The development of remote sensing technology has made earth surface observation convenient, highly efficient, and low cost (Mei, et al., 2001). Satellite data have been an important source reference for land cover mapping with the improvement in spatial and temporal resolutions. Different types of satellite d ata-derived global land cover datasets exist, including the International Geosphere Biosphere Programme Data and Information System Cover—IGBP DISCover (Loveland, et al., 2000) produced by the United States Geological Survey, University of Maryland land cover product—UMD (Hansen, et al., 2000) d eveloped by the University of Maryland, MODIS Land Cover Type Product collection 4-MOD12Q1 (Firedl, et al., 2002) p roduced by Boston University, Global Land Cover 2000-GLC2000 (Bartholomé & Belward, 2005) produced by the European Community Joint Research Center, and the Global Land Cover Product—GlobCover 2005 (Bicheron, et al., 2008) p roduced by the European Space Agency. The classification scheme used by three United States global land cover products (IGBP DISCover, UMD, and MOD12Q1) is the IGBP scheme, which has 17 categories of land cover types, whereas the classification scheme used in two European global land cover products (GLC2000 and GlobCover 2005) is the Land Cover Classification System (LCCS) with 22 categories developed by Food and Agriculture Organization of the United Nations. A land cover map is a simulation and generalization of reality, although it can depict the properties of earth surface to some degree; the process of generalization results in some error and loss of information (Brown, et al., 1999). Thus, the evaluation of these existing global land cover products is meaningful for data usage.

**Received:** 2013-03-18; **Accepted:** 2013-09-30; **Version of record first published:** 2013-10-09

**Foundation:** National Basic Research Program of China (973 Program) (No. 2011CB952001)

**First author biography:** YANG Yongke (1987— ), male, master, he majors in remote sensing for natural resources and environment, remote sensing digital image processing. E-mail: yangyk198734@126.com

**Corresponding author biography:** XIAO Pengfeng (1979— ), male, Ph. D., associate professor, his research interests are remote sensing digital image p rocessing, remote sensing for natural resources and environment. E-mail: xiaopf@gmail.com

Existing methods for the accuracy assessment of global land cover datasets are the absolute and relative scale assessments. For absolute scale assessment, the accuracy of global land cover products will be measured based on independent field data. Ground truth data at a global or continent scale are unavailable, and the collection of field data is time consuming and expensive, particularly on such a large scale. Thus, evaluating the absolute accuracy of global land cover products is difficult and be limited by the quality of validation data (Foody, 2002, 2010). Except for the data producer, only Herold, et al. (2008) estimated the absolute accuracy of four 1 km resolution global land cover products with filed data. However, the overall accuracy derived from the error matrix at a global scale is inappropriate for specific continents or sub-regions (Comber, et al., 2012). For relative scale assessment, a comparative analysis among different global land cover products is conducted to determine agreement and disagreement. However, differences such as satellite data source, classification system, and methodologies used to develop different land cover datasets will limit comparability and compatibility (Giri, et al., 2005; McCallum, et al., 2006; Herold, et al., 2008). A comparative analysis between IGBP DISCover and UMD (Hansen & Reed, 2000) shows that the wide and varying sets of ancillary data sources used in different classification techniques are important variables influencing the classification result. Giri, et al. (2005) found that both total area and spatial a greement between GLC2000 and MOD12Q1 vary from region to region: (1) southern Siberia extending to the border of K azakhstan, Mongolia, and China; (2) Tibetan plateau are two of the five major zones of disagreement. McCallum, et al. (2006) analyzed the spatial agreement among four 1 km resolution global land cover datasets and argued that agreement is very low in Asia. Liang and Gong (2010) studied the mapping uncertainty of MODIS land cover products and highlighted that low producer accuracy was mainly observed in mountainous areas and translation zones. Ran, et al. (2009) estimated the accuracy of the four 1 km resolution global land cover datasets over China based on a large-scale (1∶ 100000) land use map of China in 2000 produced by the Chinese Academy of Sciences with an a ggregated classification scheme. Wu, et al. (2008) estimated the accuracy of croplands of four 1 km resolution global land cover datasets over China based on a large-scale (1∶ 100000) land use map of China in 2000 produced by the Chinese Academy of S ciences, they found that the cropland accuracy of four global land cover datasets over China varied from region to region. Niu, et al. (2012) assessed the accuracy of permanent wetlands of Glob Cover 2009 based on Chinese wetlands in 2008 produced through visual interpretation by the Chinese Academy of Sciences. Gao and Jia (2012) analyzed the spatial and quantitative agreement between MOD12Q1 and GLC2000 in China, and the result shows that the accuracy of water, grassland, cropland, and barren are high in both global land cover products.

In this paper, the accuracy of five global land cover products (IGBP DISCover, UMD, GLC2000, MOD12Q1, and GlobCover 2005) over China and adjacent regions was evaluated. First, categories of the five global land cover datasets were converted into the IGBP scheme based on the correlation coefficients of the corresponding classes, which were computed according to the class definition of each land cover dataset. Second, spatial agreements among five global land cover datasets were analyzed through visual comparison and per-pixel comparison. Finally, the classification accuracy of five land cover datasets was evaluated with independent validation samples c ollected through Google Earth high-resolution satellite images.

## 2 STUDY AREA AND DATA

### 2.1 Description of study area

The study area is located 0°N—70°N and 40°E—170°E, including China and adjacent regions. In the study area, the t emperature zones are tropical, sub-tropical, temperate, warm temperate, and cool temperate from south to north. The climates are monsoon climate and continental climate from east to west, the elevation increases evidently from east to west, and the terrain changes mainly because of the effect of mountains and plateaus. All these elements result in the complex land surface and rich landscape of the area. The population of study area is more than 50% of the total world population. This high population density resulted in destructive human activities, particularly in recent decades. Rapid industrialization and urbanization seriously influenced the earth surface. China has been one of the most frequent regions studied in terms of land use/cover change in the world.

### 2.2 Global land cover datasets

Five global land cover datasets were downloaded for free from a Web site and reprojected to Lambert conformal conic p rojection. Detailed information on the five land cover products is described as: (1) IGBP DISCover 1 km resolution global land cover product (1992—1993); (2) UMD 1 km resolution global land cover product (1992—1993); (3) GLC2000 1 km resolution global land cover product (2000); (4) MOD12Q1 1 km resolution global land cover product (2001); and (5) GlobCover 2005 300 m resolution global land cover product (2005—2006). Table 1 shows the characteristics of the five global land cover datasets.

Urban and water areas are too small to be well classified u sing coarse-resolution satellite data. Thus, methods used to e xtract urban and water information differ for each global land cover product. In IGBP DISCover, urban and water areas come from the Digital Chart of the World by the Defense Mapping A gency (Loveland, et al., 2000). The UMD product uses the same urban data as IGBP DISCover, but the water layer is a water mask made for the MODIS sensor (Hansen, et al., 2000). In GLC2000 over China, urban data are classified based on a visual interpretation of the SPOT vegetation data in August, whereas water data were extracted using unsupervised classification along with the remaining classes (Xu, et al., 2005). Urban and water areas in MOD12Q1 were classified using classification tree along with other classes (Friedl, et al., 2002). For Glob Cover 2005, water was masked with land/ocean boundary in a Medium Resolution Imaging Spectrometer, whereas urban areas were determined using supervised classification (Bicheron, et al., 2008).

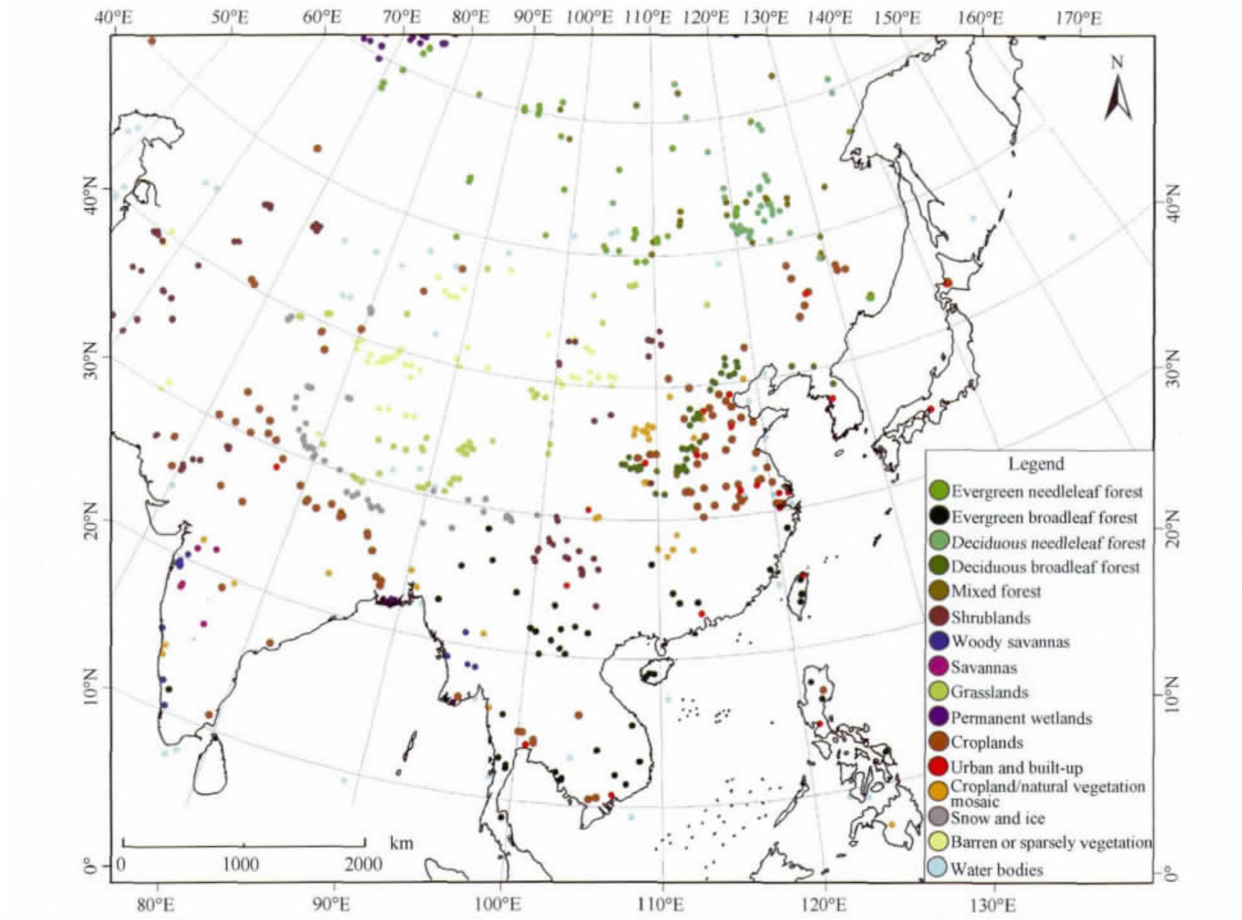


Fig. 1 Schematic diagram of the study area and the distribution situation of validation samples from 2000 to 2001

**Table 1 Characteristics of five global land cover products**

| | IGBP DISCover | UMD | GLC 2000 | MOD12Q1 2001 | GlobCover 2005 |
|---|---|---|---|---|---|
| Sensor | AVHRR | AVHRR | SPOT VEGETATION | Terra/MODIS | ENVISAT/MERIS |
| Input data | 1992-04—1993-03 12 monthly NDVI composites covering | 1992-04—1993-03 41 metrics derived from NDVI and five bands | 1999-11—2000-12 36 10-day NDVI, DEM and meteorological composites | 2001-01—2002-01 16-day Nadir BRDF adjusted Reflectance, seven spectral bands, 16-day EVI | 2004-10—2006-06 13 spectral bands (300 m resolution) |
| Classification scheme | IGBP (17 classes) | IGBP (14 classes) | LCCS (22 classes) | IGBP (17 classes) | LCCS (22 classes) |
| Classification model and classification method | Each continent is classified separately by K-means method and then stitched together. Water and urban were masked with existing data | Entire globe id classified using a classification tree algorithm. All pixels in terminal node will be labeled to class with large probability | The globe is divided into 19 regions. China is dived into 9 region based on climate condition, each region is classified separately by ISODATA method | Entire globe id classified using a classification tree algorithm. The pixels in terminal node will be labeled to class with large probability | The world is split into 22 regions, each region is classified independently by unsupervised clustering expect for urban and wetland |

The wide and varying sets of ancillary data used in classifying process are also important variables influencing the classification result (Hansen & Reed, 2000). In the unsupervised classification method, ancillary data can be used to assign the specific class labels to the cluster polygon. In the supervised classification method, ancillary data can be used to identify labels of the training sample. The ancillary data used in IGBP DISCover include digital and hardcopy land cover maps, atlases, and Landsat i mages. The ancillary data for UMD are Landsat Multispectral Scanning System images. For GLC2000 in China, the ancillary data are 1∶ 1000000 land use maps and 1∶ 1000000 vegetation maps. The main ancillary data for MOD12Q1 are Landsat T hematic Mapper images.

## 2.3 Reference data

Land cover maps (1992 and 2001) derived from Landsat TM image and a high-resolution image (2005) of Hangzhou City are used as reference data to reveal detailed characteristics of w ater and urban areas using five global land cover datasets at a local scale. Features of urban and water are too small to be c aptured through a comparative analysis of the entire study area. Thus, local scale comparative analysis is required to reveal the consistency among the five land cover datasets.

## 3 METHODOLOGY AND DATA PROCESSING

### 3.1 Transformation of classification scheme

The same classification scheme served as basis of the comparative analysis among five global land cover products. Except for IGBP DISCover and MOD12Q1, the classification schemes defined in each global land cover dataset differ in terms of categories, range of vegetation canopy, and height limitation used to distinguish tree and shrublands. In recent studies (Bartholomé & Belward, 2005; Giri, et al., 2005; McCallum, et al., 2006), no u niform translation rule governs all these classification schemes.

In this study, except for open and closed shrublands that were aggregated into shrublands, all other classes in the IGBP scheme remained unchanged. All five land cover datasets were then converted into the IGBP scheme. In the definition of each class, only two quantitative algorithms exist: the range of vegetation canopy and tree heigh limitation. Tree height limitation is used to distinguish shrublands and trees. The definition of vegetation canopy of different class is a range from 0—100% (Hansen, et al., 2000; Friedl, et al., 2002; Bartholomé & Belward, 2005; Bicheron, et al., 2008), which was selected to computed the be used to correlation coefficients and to establish the transformation rule in this study. If the corresponding class in two datasets has a high overlapping range of vegetation canopy, their correlation coefficients will also be high.

$$C = \frac{1}{2}\left(\frac{R_{1D}}{R_1} + \frac{R_{1D}}{R_D}\right) \tag{1}$$

where $R_1$ is the length of vegetation canopy range of a certain class defined in the IGBP scheme, $R_D$ is the length of vegetation canopy range of a corresponding class defined in the dataset $\boldsymbol{D}$, and $R_{1D}$ is the length of the overlapping range of vegetation c anopy of corresponding classes in the IGBP scheme and dataset $\boldsymbol{D}$. The correlation coefficient value is $[0, 1]$; if no overlapping range exists, the correlation coefficient is 0. However, the correlation coefficient of water, urban, and built-up areas, as well as that of permanent wetlands and cropland could not be computed because no similar algorithm exists in the definition. Mixed forest and cropland/natural vegetation mosaic are the mosaics of different land cover types, but proportions differ in five datasets, such that the correlation coefficient likewise could not be computed. Table 2 shows the corresponding relationship and correlation coefficients of different classes among five global land cover datasets.

**Table 2 Corresponding relationship and correlation coefficients of different classes among five global land cover datasets**

| IGBP DISCover/MOD12Q1 | UMD | GLC2000 | GlobCover 2005 |
|---|---|---|---|
| Evergreen needleleaf forest (>60%), 1 | Evergreen needleleaf forest (>60%), 1 | Tree cover, needle-leaved, evergreen (>15%), 0.74 | Closed needleleaved evergreen green forest (>40%), 0.83 |
| Evergreen broadleaf forest (>60%), 1 | Evergreen broadleaf forest (>60%), 1 | Tree cover, broad-leaved, evergreen (>15%), 0.74 | Closed to open broadleaved evergreen forest (>15%), 0.74 |
| Deciduous needleleaf forest (>60%), 1 | Deciduous needleleaf forest (>60%), 1 | Tree cover, needled-leaved, deciduous (>15%), 0.74 | Open needleleaved deciduous or evergreen forest (15%—40%), 0.74 |
| Deciduous broadleaf forest (>60%), 1 | Deciduous broadleaf forest (>60%), 1 | Tree cover, broadleaved, deciduous, closed (>40%); Tree cover broadleaved, deciduous, open (15%—40%), 0.74 | Closed broadleaved deciduous forest (>40%), Open broadleaved deciduous forest/woodland (15%—40%), 0.74 |
| Mixed forest (<60%), 1 | Mixed forest (<60%), 1 | Tree cover, mixed leaf type (>15%) | Closed to open mixed broadleaved and needleleaved forest (>15%) |
| Open and closed shrublands, 1 | Open and closed shrublands, 1 | Shrub cover, closed-open, evergreen<br>Shrub cover closed-open, deciduous (>15%), 0.97 | Closed to open shrubland (>15%), 0.97 |
| Woody savannas (tree cover 30%—60%), 1 | Woodland (tree cover 40%—60%), 0.83 | Mosaic, tree cover/other natural vegetation (20%—70%), 0.80 | Mosaic forest or shrubland (50%—70%) / grassland (20%—50%);<br>Mosaic grassland (50%—70%) / forest or shrubland (20%—50%),0.80 |
| Savannas (tree cover 10%—30%), 1 | Wooded grassland (tree cover 10%—30%), 0.83 | — | — |
| Grasslands (>10%),1 | Grasslands (>10%), 1 | Herbaceous cover, closed-open (>15%) 0.97 | Closed to open herbaceous vegetation (>10%), 0.97 |
| Permanent wetlands | — | Tree cover, regularly flooded, fresh water; Tree cover, regularly flooded, saline water regularly flooded; Regularly flooded shrub and/or herbaceous cover | Closed to open (>15%) broadleaved forest regularly flooded; Closed (>40%) broadleaved forest or shrublands permanently flooded; Closed to open (>15%) grassland on regularly flooded or waterlogged soil |
| Croplands | Croplands | Cultivated and managed areas | Post-flooding or irrigated croplands(or aquatic) Rainfed croplands |
| Urban and built-up | Urban and built | Artificial surfaces and associated areas | Artificial surfaces and associated areas |
| Cropland/natural vegetation | — | Mosaic cropland/tree cover/other natural vegetation; Mosaic cropland/shrub or grass cover | Mosaic cropland (50%—70%) / vegetation (20%—50%); Mosaic vegetation (50%—70%) / cropland (20%—50%) |
| Snow and ice | — | Snow and ice | Permanent snow and ice |
| Barren or sparsely vegetation(<10%), 1 | Bare ground(<10%), 1 | Sparse herbaceous or sparse shrubland; Bare areas (<15%), 0.97 | Sparse herbaceous or sparse shrubland Bare areas (<15%), 0.97 |
| Water bodies | Water | Water bodies | Water bodies |

Note: The content in "( )" is the correlation coefficients, "—" indicates no this class in datasets.

### 3.2 Spatial agreement

The aim of spatial agreement analysis is to reveal the differences and similarities in terms of the spatial distribution situation of different classes among different datasets. In this study, spatial agreements were implemented through visual comparison and per-pixel comparison. Visual comparison can provide the spatial variation among the five datasets from region to region subjectively. Per-pixel comparison can calculate the overall agreement ($A$) and per-class agreement ($B_i$) among different datasets. The e quations to compute $A$ and $B_i$ are as

$$A = \frac{\sum_{1}^{11} XY_i}{\sum_{1}^{11} (X_i + Y_i)/2} \times 100\% \qquad (2)$$

$$B_i = \frac{XY_i}{(X_i + Y_i)/2} \times 100\% \qquad (3)$$

where $X_i$ refers to the number of class $i$ pixels in dataset $\boldsymbol{X}$, $Y_i$ denotes the number of class $i$ pixels in dataset $\boldsymbol{Y}$, and $XY_i$ stands for the number of class $i$ pixels in both datasets $\boldsymbol{X}$ and $\boldsymbol{Y}$ with the same location. UMD has no permanent wetlands, cropland/natural vegetation mosaic, snow, and ice, whereas GLC2000 and GlobCover 2005 have no savannas. Thus, these classes are not involved in spatial agreement analysis. In the study area, the i nland water cannot be separated from ocean because of the lack of coastline data. Thus, water bodies are likewise not considered. Except for these classes, all the remaining 11 classes were c onsidered in spatial agreement analysis.

### 3.3 Collection of validation samples

Validation samples were collected through human interpretation of high-resolution images provided by Google Earth. The availability of images via Google Earth has been a crucial data source for land cover mapping or accuracy assessment (Bicheron, et al., 2008; Friedl, et al., 2010; Clark, et al., 2010). The advantages of Google Earth: (1) it provides free access to high-resolution images; (2) it provides synoptic views of the entire sample plot at different angles and spatial scales; (3) the spatial location error of its high-resolution images is low (approximately 15 ± 5 m) (Clark, et al., 2010); (4) the images that it provides are crucial ancillary data for the interpretation of high-resolution i mages; (5) the timeline in Google Earth can be used to look up land cover types at different times.

Errors are of two types: spatial and interpretation errors, which are expected in the validation samples collected from Google Earth. Spatial error is mainly caused by the use of different spatial coordinate systems or terrain displacements. The resolutions of the five global land cover datasets are 300 m and 1 000 m. Thus, the (15 ± 5) m spatial error (Clark, et al., 2010) has a negligible effect on coarse-resolution land cover datasets, and validation samples will be reprojected into same projection with a global land cover dataset. Interpretation errors are mainly caused subjective factors. For instance, different interpreters may have different perspectives on the same image based on their d iscipline or background. Four rules were formulated to reduce interpretation error and to ensure the homogeneity of the validation samples: (1) Validation samples must be selected from the center of a large homogeneous area, and the size of each sample should be approximately four pixels (approximate 2 km × 2 km or 600 m × 600 m). For example, one 1 km × 1 km region c omposed of 60% needleleaf forest and 40% broadleaf forest can be regarded as a fixed forest when the spatial resolution is 1 km. However, an error emerges if all cells contained by this r egion are regarded as fixed forest when the spatial resolution changes to 300 m. Thus, the size of each validation sample must be suitable for the resolution of land cover datasets. (2) Validation samples must be selected from an area that has a high-resolution image. (3) For deciduous and evergreen forests, the validation sample must be interpreted based on images in different periods. (4) Images of the corresponding location can be used to help identify the land cover type when high-resolution image interpretation is difficult.

The validation samples used in this paper consisted of two different time periods (2000—2001 and 2004—2006). Until 2000, high-resolution images have become available to the public. In addition, no ground truth data could be used to collect samples, making it impossible to collect validation samples from 1992 to 1993. Two main questions, that is, how much and where the land cover has changed from 1992 to 2001 in the study area, must be answered to confirm whether the validation samples in 2000 to 2001 can reflect the real earth surface and assess the a ccuracy of global land cover from 1992 to 1993. Some studies (Liu & Buheaosier, 2000; Wang, et al., 2001, 2002; Liu, et al., 2002, 2003; Tian, et al., 2003; Li, et al., 2005; Liu, et al., 2009) showed that land cover change in China mainly occurs in: (1) Traditional agricultural regions, such as Huang-Huai-Hai Plain, Yangtse Delta, and Sichuan Basin, whereas cropland decreased significantly because of urbanization. (2) Northeast and northwest regions, such as farming-pastoral or translation zones, whereas forest or grassland is cultivated into croplands. However, land cover change is not evident in western regions. The effect of the policy of cropland conversion to forest and grassland is only e vident in the local region. Moreover, the speed of reforestation is slower than the cultivation of cropland. The classes of land cover change in China in the 1990s are major cropland, grassland, forest, urban, and built-up areas. Grassland in the north underwent serious degradation from high coverage to low coverage, even undergoing desertification (Li, 1997). No land cover change occurred in most areas of China. Supposing that validation samples in 2000 to 2001 were obtained from areas with no large land cover change, the condition could well reflect the real property of earth surface from 1992 to 1993 and could thus be used to validate the classification accuracy of land cover products in 1992 to 1993.

Table 3 shows detailed information on the validation samples, where "—" indicates that this class is absent in the land cover dataset. Validation samples in 2000 to 2001 were used to assess the accuracy of IGBP DISCover, UMD, GLC 2000, and MOD12Q1. Validation samples in 2004 to 2006 were used to a ssess GlobCover 2005, whereas validation samples in 2004 to 2006 were collected based on validation samples in 2000 to 2001. If the land cover type of one certain sample exhibits no change in two periods, this sample will be retained and resized to

600 m × 600 m; otherwise, a new sample will be collected to r eplace it.

**Table 3 Shows detailed information on the validation**

| Class name | IGBP DISCover /MOD12Q1 | UMD | GLC2000 | Glob Cover 2005 |
|---|---|---|---|---|
| Evergreen needleleaf forest | 37 | 37 | 37 | 43 |
| Evergreen broadleaf forest | 44 | 44 | 44 | 45 |
| Deciduous needleleaf forest | 37 | 37 | 37 | 31 |
| Deciduous broadleaf forest | 37 | 37 | 37 | 39 |
| Mixed forest | 28 | 28 | 28 | 25 |
| Shrublands | 69 | 69 | 20 | 22 |
| Woody savannas | 13 | 13 | 13 | 14 |
| Savannas | 5 | 5 | — | — |
| Grasslands | 44 | 44 | 44 | 53 |
| Permanent wetlands | 25 | — | 25 | 25 |
| Croplands | 101 | 101 | 101 | 103 |
| Urban and built-up | 22 | 22 | 22 | 38 |
| Cropland/natural vegetation mosaic | 34 | — | 34 | 32 |
| Snow and ice | 43 | — | 43 | 35 |
| Barren or sparsely vegetation | 50 | 50 | 50 | 49 |
| Water bodies | 49 | 49 | 49 | 52 |
| Total number | 638 | 536 | 584 | 606 |

## 3.4 Confusion matrix

Confusion matrix is the most widely used method for accuracy assessment (Foody, 2002; Herold, et al. 2008, Ran, et al., 2009; Clark, et al., 2010). This approach is a cross-tabulation of the map class against the field data or reference data and can provide numerous accuracy measures, the most commonly used measures are overall accuracy, producer′s accuracy, user′s accuracy, and Kappa coefficients. These accuracy indices are defined as

$$\text{Overall accuracy} = \frac{\sum_{i=1}^{17} X_{ii}}{N^2} \times 100\% \tag{4}$$

$$\text{Kappa coefficient} = \frac{N\sum_{i=1}^{17} X_{ii} - \sum_{i=1}^{17}(X_{i+}X_{+1})}{N^2 - \sum_{i=1}^{17}(X_{i+}X_{+1})} \tag{5}$$

$$\text{User′s accuracy} = \frac{X_{ii}}{X_{i+}} \times 100\% \tag{6}$$

$$\text{Producer′s accuracy} = \frac{X_{ii}}{X_{+i}} \times 100\% \tag{7}$$

where $X_{ii}$ refers to the number of class $i$ pixels that were correctly classified, $X_{i+}$ denotes the number of class $i$ pixels in the classification result, $X_{+i}$ stands for the number of class $i$ pixels in the reference data, and $N$ is the total number of all the pixels.

# 4 RESULTS AND DISCUSSIONS

## 4.1 Spatial agreement analysis

Land use/cover changes rapidly in the study area, and the timeframe of the five global land cover datasets ranges from 1992 to 2005. Thus, in theory, the real change in land cover will significantly affect spatial agreement analysis. However, some studies (Friedl, et al., 2010; McCallum, et al., 2006) show that land cover change that stems from classification method is well above the actual change level. Liu, et al. (2003) argued that the land cover area that changed in China from 1990 to 2000 c omprised only approximately 0.5% of the total land areas area of the country. In this study, a comparative analysis shows that the agreement among different land cover datasets is very low. For example, the agreement between IGBP DISCover and UMD is only 37%, whereas the disagreement level is larger than e xpected. Consequently, although the timeframes of the five land cover datasets differ, the agreement analysis among these land cover datasets remains meaningful.

### *4.1.1 Visual comparison*

Visual comparison was performed both at the regional and local scales. The regional scale comparison mainly shows the a greement and difference among five datasets from region to r egion. The local scale comparison mainly shows the d etailed characteristics of water and urban areas of the five datasets.

The comparative analysis at the regional scale was divided into two types of content according to the IGBP and generalized schemes. In the generalized scheme, 16 classes in the IGBP scheme were aggregated into eight classes: forest, woody savannas, grasslands, croplands, barren or sparse vegetation, water bodies, urban, and built-up areas. Forest was aggregated by evergreen needleleaf forest, evergreen broadleaf forest, deciduous needleleaf forest, deciduous broadleaf forest, mixed forest, and shrublands.

A visual comparison analysis with the generalized scheme is shown in Fig.2. Six major areas of disagreement among the five global land cover datasets are in (1) Russia; (2) Central Asia (Kazakhstan extending to the border of Pakistan); (3) India; (4) West Tibetan Plateau of China; (5) Southeast area of China; (6) Mongolia. The forest area in UMD is less than that in other datasets in Russia, whereas croplands are less than that in other datasets in India. UMD and IGBP DISCover have more woody s avannas than other land cover datasets.

A visual comparison analysis with the IGBP scheme is shown in Fig.3. Shrublands in GLC2000 and GlobCover2005 are less than the other three land cover datasets. Deciduous needleleaf forest in GLC2000 and GlobCover 2005 are more than that of the other three land cover datasets. Spatial agreement among the five land cover datasets has a strong relationship with the original classification scheme used in each dataset. Datasets with the same original classification scheme have better spatial a greement than those with a classification scheme that is different from the original. Misclassification b etween grasslands with barren or sparse vegetation mainly o ccurs in Kazakhstan and Mongolia. Misclassification among forests with croplands mainly occurs in Southeast China and I ndia. Misclassification of the five kinds of forest mainly occurs in Russia. Misclassification between shrublands and other kinds of forest is minimal.

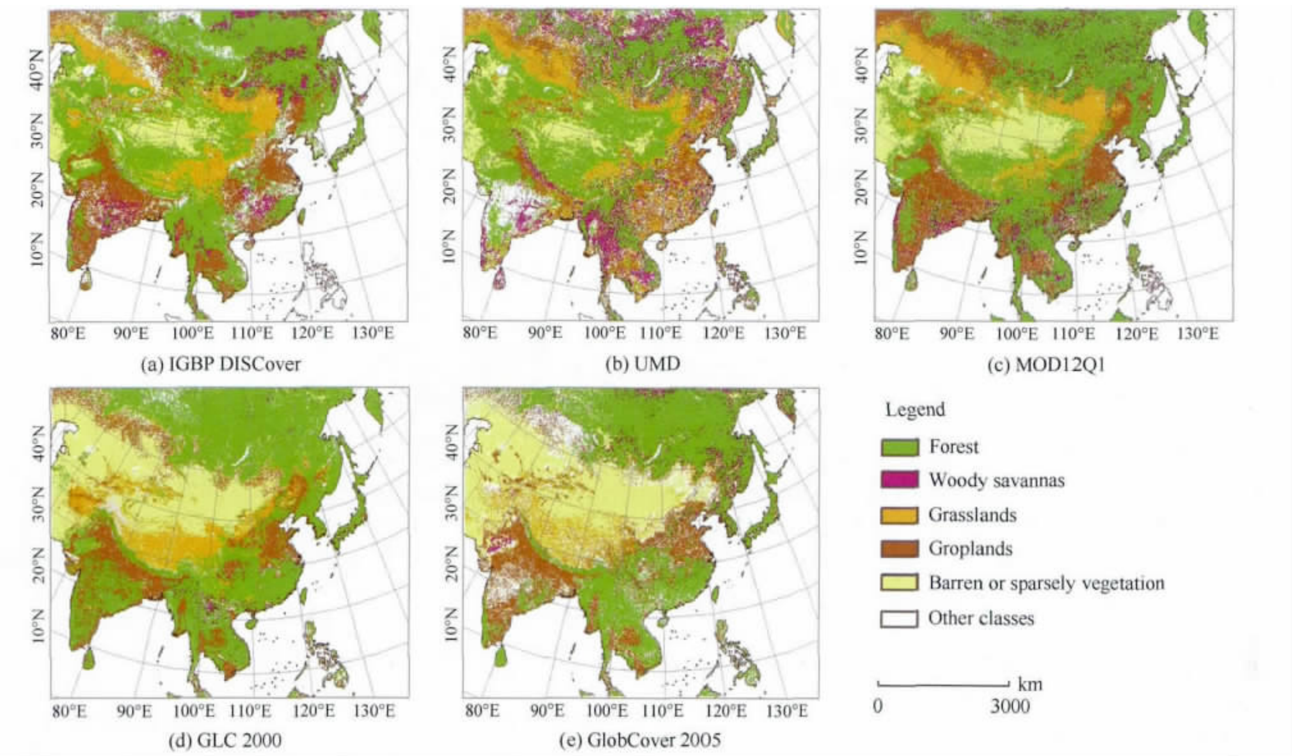


Fig. 2 Comparative analysis of five land cover products with generalized scheme at the regional scale

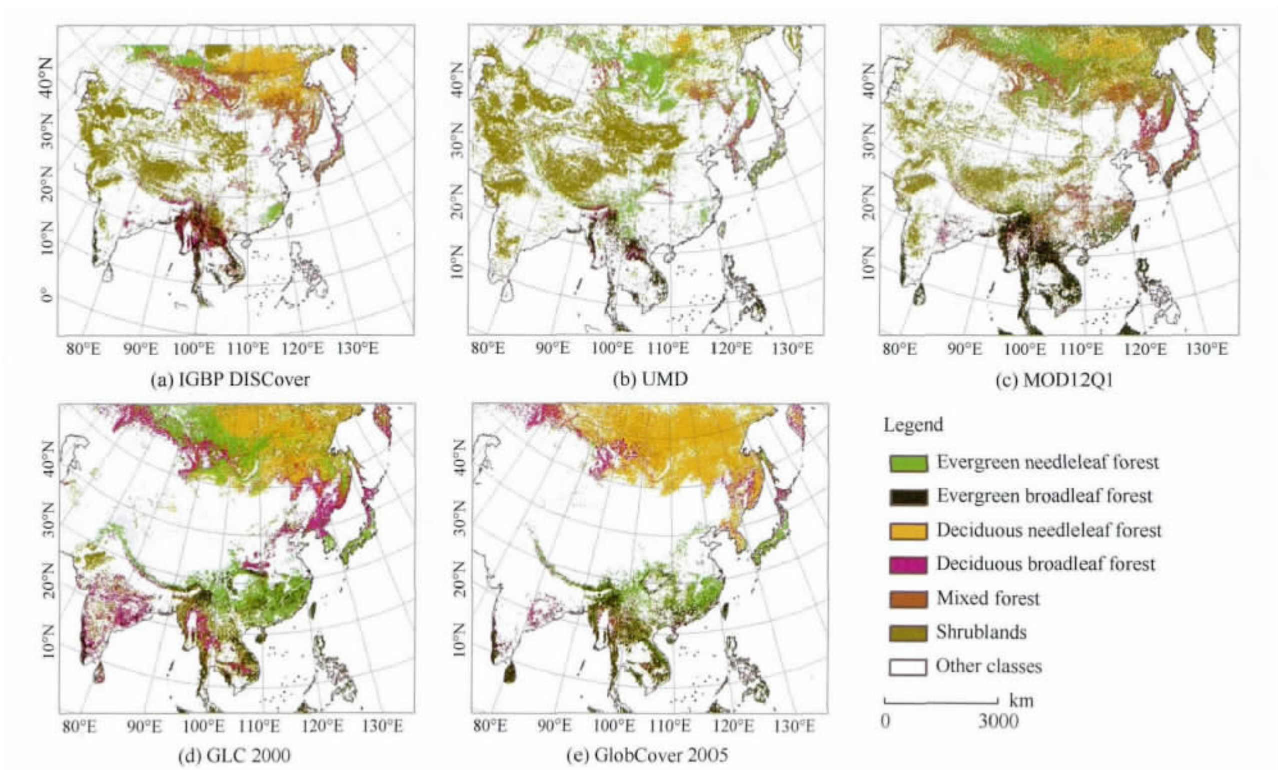


Fig. 3 Comparative analysis of five kinds of forest in the IGBP scheme at the regional scale

Fig. 4 shows a visual comparison of water bodies, urban, and built-up areas at a local scale with reference data. Compared with land cover map (TM, 1992), the urban and built-up area in Hangzhou City is underestimated, but the area of water bodies is overestimated. Large areas of croplands were misclassified into grassland in UMD. In addition, the area of woody savannas is larger than expected. Compared with land cover map (TM, 2001), the area of the west lake was exaggerated in GLC2000, and west lake should be located southwest of Hangzhou rather than where it was shown in GLC2000. Urban and built-up areas were seriously overestimated in MOD12Q1. Urban areas and water bodies in GlobCover 2005 exhibit good agreement with the high-resolution image of 2005, except that some forests southwest of west lake were misclassified as urban and built-up areas.

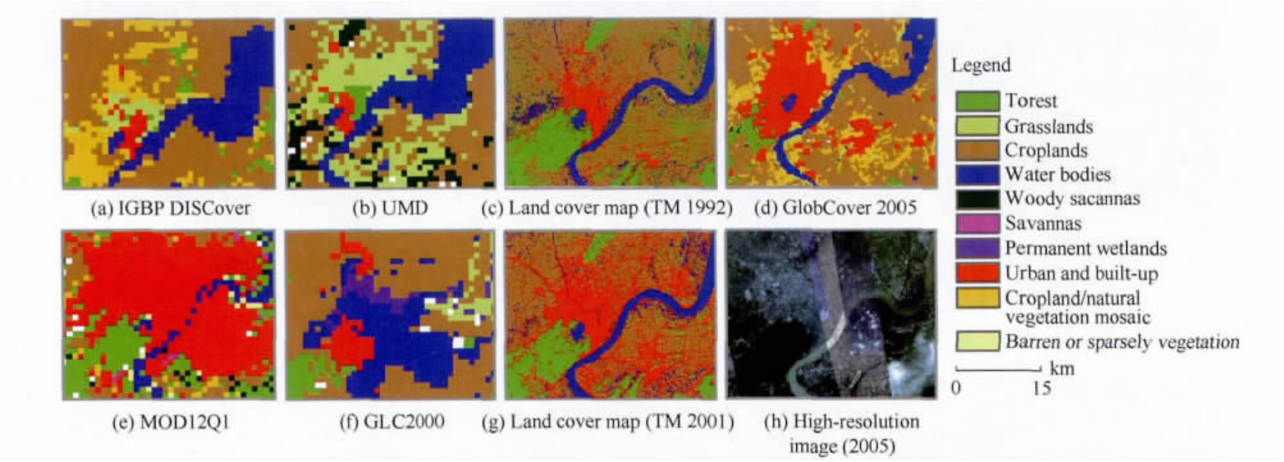


Fig. 4 Comparative analysis with reference data at the local scale

*4.1.2 Per-pixel comparison*

Per-pixel comparison is a supplement to visual comparison. Instead of a subjective description, per-pixel comparison can d epict the spatial agreement among different land cover datasets quantitatively. Overall agreement and per-class agreement were computed between IGBP DISCover and UMD, GLC2000, and MOD12Q1. As described above, land cover change caused by classification is well above the actual change level, thus achieving overall agreement and per-class agreement among all four 1 km resolution land cover datasets (IGBP DISCover, UMD, GLC2000, and MOD12Q1). The resolution of Glob Cover 2005 differs from that of other land cover datasets and was thus excluded from per-pixel comparison.

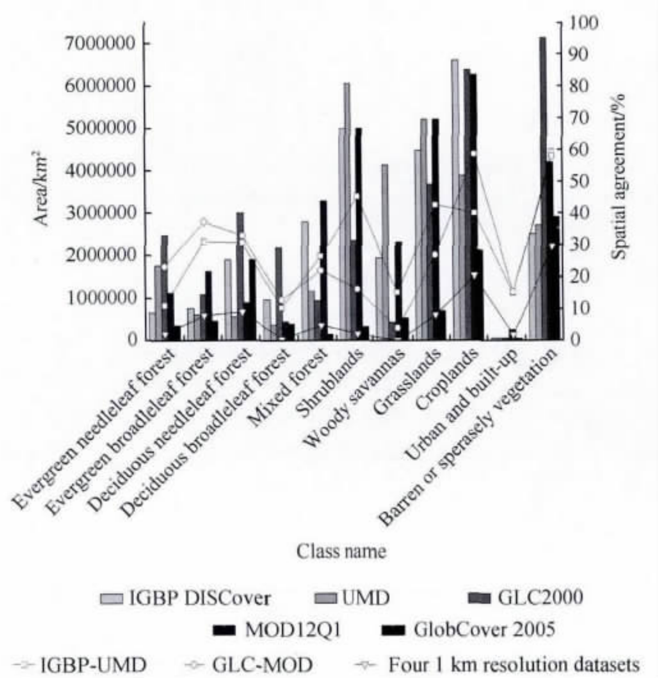


Fig. 5 Area and spatial agreement of 11 classes

Fig. 5 shows the area and spatial agreement of 11 classes a mong different land cover datasets. The difference in area of 11 classes in five land cover datasets is large. Per-class agreement a mong different datasets varied significantly from class to class. Overall agreement is 36.93% between IGBP DISCover and UMD. Barren or sparse vegetation had the highest agreement (59.35%), whereas all remaining classes exhibited an agreement below 50%. Although urban and built-up areas used in UMD were taken directly from IGBP DISCover, the agreement is only 15.14% and is mainly affected by the spatial error caused by the reprojection process. Overall agreement is 36.67% between GLC2000 and MOD12Q1. Agreements of croplands and barren or sparsely vegetation are 58.66% and 58.04%, respectively, whereas the agreement of the remaining classes is lower than 5 0.00%. Overall agreement of IGBP DISCover, UMD, GLC2000, and MOD12Q1 is only 11.30%, with barren or sparse vegetation having the highest agreement at only 29.54%.

## 4.2 Accuracy evaluation

Five confusion matrices of the five global land cover datasets were established with the validation samples. Limited by the length of the paper, five confusion matrices are not shown here. Four accuracy measures: producer's accuracy, user's accuracy, overall accuracy, and Kappa coefficient, which were computed from confusion matrices, are shown in Fig. 6, Fig. 7 and Table 4.

Fig. 6 shows the user's accuracy of 11 classes in five land cover datasets. User's accuracy of urban and built-up areas in five land cover datasets are all above 90%. User's accuracy of evergreen broadleaf forest, croplands, and water bodies are all a bove 60%. User's accuracy of deciduous broadleaf forest, mixed forest, shrublands, and savannas are very low, and the difference is large between different land cover datasets.

Fig. 7 shows the producer's accuracy of 11 classes in five datasets. Except for the producer's accuracy of water bodies, the producer's accuracy in the five land cover datasets is above 85%. The producer's accuracy of other classes in five land cover datasets exhibited a large variation. For example, the producer's accuracy of croplands ranged from 73.35% to 89.39%, and that of evergreen broadleaf forest from 30% to 73.57%. Evergreen needleleaf forest and woody savannas had the lowest producer's accuracy.

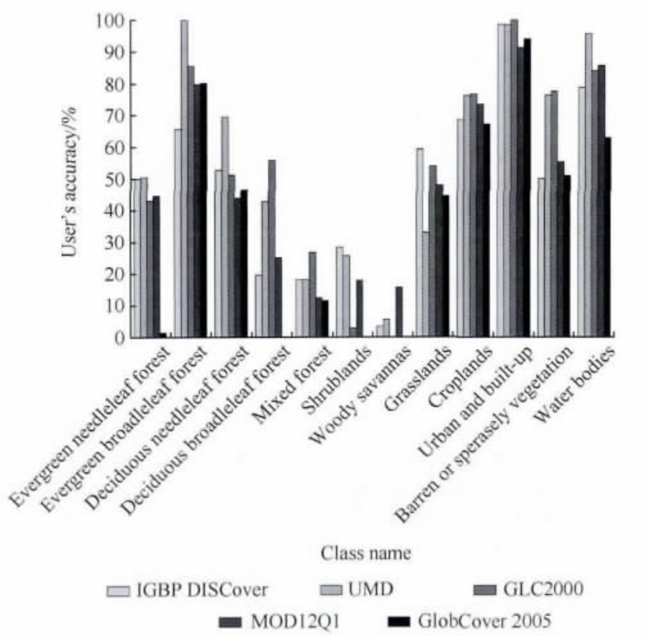


Fig. 6 User's accuracy of different classes

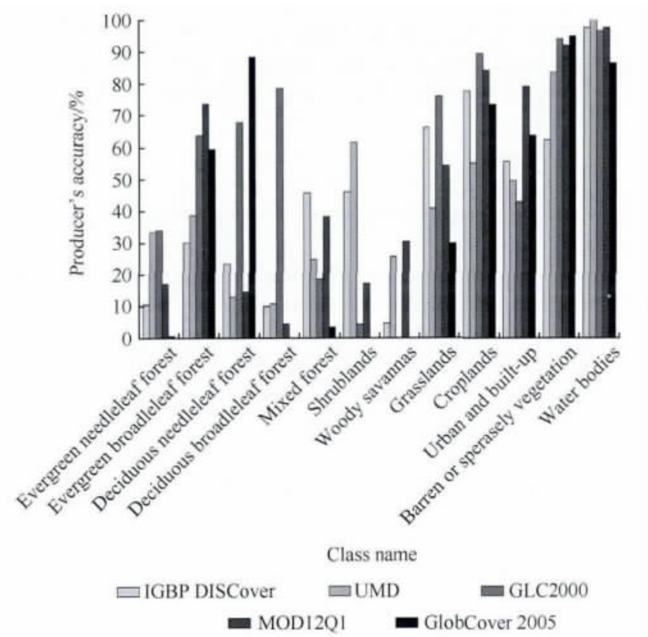


Fig. 7 Producer's accuracy of different classes

For urban and built-up areas in five land cover datasets, user's accuracy is above 90.00%, but producer's accuracy ranged from 42.86% to 78.95% because the urban and built-up areas in IGBP DISCover and UMD complied with existing maps range from 1960 to 1980, making them outdated and thus cannot represent the urban areas during the period of rapid urbanization. The confusion matrix shows that 14.29% urban and built-up areas were misclassified into cropland, and 8.27% of urban and built-up areas were misclassified into cropland/nature vegetation in IGBP DISCover, whereas 15.04% of urban and built-up areas were misclassified into croplands in UMD. Producer's accuracy of urban and built-up areas in GLC2000 is the lowest among five land cover datasets because urban and built-up areas in GLC2000 were classified through visual interpretation, such that numerous subjective factors influenced the classification result. The confusion matrix shows that 23.30% of urban and built-up areas were misclassified into croplands. MO12Q1 had the highest producer's accuracy in terms of urban and built-up areas because all pixels in the terminal node were labeled as the dominant class, which can result in the overestimation of urban and built-up areas, particularly for classes around these areas that have with similar spectral features (Fig. 3). Producer's accuracy of shrublands for IGBP DISCover and UMD was higher than that of other land cover datasets because some shrublands were m isclassified into grasslands in GLC2000 and GlobCover 2005 (Fig. 3). Producer's accuracy of grasslands in GlobCover 2005 was the lowest among five land cover datasets because 24.62% of grassland was misclassified into barren or sparse vegetation.

Accuracy of snow and ice, permanent wetlands, cropland/natural vegetation mosaic, and savannas are shown in Table 4. User's accuracy of snow and ice and permanent wetlands in five land cover datasets were all above 82.35%; by contrast, their producer's accuracy was very low, except in GLC2000. Both producer's accuracy and user's accuracy of woody savannas in UMD were the highest among five land cover datasets, because the collection of training samples was not limited to tropical regions. Areas outside of the tropical regions, which have land covers fitting the description of woody savannas, were also added to the validation sample database.

The confusion matrices show misclassification mainly occurs between shrublands and grassland, croplands and cropland/nature vegetation mosaic, woody savannas and savannas, and among five kinds of forests. The misclassification of five kinds of forests significantly affects the overall accuracy and Kappa coefficients. In some cases, however, determining the accuracy of a land cover dataset with a generalized scheme is sufficient for data use. Thus, the accuracy of forest aggregated with evergreen needleleaf forest, evergreen broadleaf forest, and deciduous needleleaf forest, as well as overall accuracy and Kappa coefficient with the aggregated scheme were also estimated in this study (Table 5).

**Table 4 Accuracy of savannas, permanent wetlands, cropland/natural vegetation mosaic and snow and ice in five datasets**

/%

| | IGBP DISCover | | UMD | | GLC2000 | | MOD12Q1 | | GlobCover 2005 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Producer's accuracy | User's accuracy | Producer's accuracy | User's accuracy | Producer's accuracy | User's accuracy | Producer's accuracy | User's accuracy | Producer's accuracy | User's accuracy |
| Savannas | 0.00 | 0.00 | 57.81 | 35.34 | — | — | 5.26 | 3.13 | — | — |
| Permanent wetlands | 45.45 | 90.91 | — | — | 63.64 | 82.35 | 31.82 | 93.33 | 14.73 | 95.56 |
| Cropland/natural vegetation mosaic | 30.30 | 68.66 | — | — | 3.79 | 15.63 | 1.52 | 8.70 | 19.54 | 8.92 |
| Snow and ice | 37.65 | 98.67 | — | — | 87.65 | 97.93 | 14.81 | 100.00 | 53.54 | 95.50 |

**Table 5 Accuracy of five datasets with IGBP scheme and generalized scheme**

/%

| | Overall accuracy | | Kappa coefficients | | Forest | |
|---|---|---|---|---|---|---|
| | IGBP scheme | Generalized scheme | IGBP scheme | Generalized scheme | Producer's accuracy | User's accuracy |
| IGBP DISCover | 51.58 | 60.75 | 0.47 | 0.55 | 60.77 | 86.01 |
| UMD | 56.71 | 58.94 | 0.52 | 0.53 | 50.09 | 77.93 |
| GLC2000 | 67.72 | 76.06 | 0.64 | 0.73 | 86.83 | 87.13 |
| MOD12Q1 | 53.19 | 59.44 | 0.48 | 0.54 | 69.43 | 82.89 |
| GlobCover 2005 | 51.12 | 58.33 | 0.45 | 0.51 | 71.83 | 76.89 |

Overall accuracy and Kappa coefficient of five land cover datasets with aggregated scheme exhibited an evident increase compared with that with the IGBP scheme. The increases in o verall accuracy and Kappa coefficient in GLC2000 were the highest. By contrast, the increases in overall accuracy and Kappa coefficient in UMD were the lowest. Regardless of whether the IGBP scheme or generalized scheme was used, GLC2000 always had the highest overall accuracy and highest Kappa coefficient a mong the five land cover datasets, whereas GlobCover 2005 a lways had the lowest overall accuracy and Kappa coefficient. User's accuracy of forest is above 75.00% in the five land cover datasets, but the difference in producer's accuracy of forest among five land cover datasets remained large. User's accuracy of forest among five datasets was above 75.00%, but user's a ccuracy varied greatly among the five datasets.

## 5 CONCLUSION

Spatial agreement among the five datasets is significantly related to the original classification scheme used in five land cover datasets. Land cover datasets with the same original classification scheme exhibited better agreement than those with d ifferent original classification schemes. GLC2000 had the highest overall accuracy, whereas GlobCover 2005 had the lowest overall accuracy among five land cover datasets. In all classes, only barren, croplands, and water bodies had high producer's a ccuracy as well as high user's accuracy. The accuracy of other classes differed significantly among five land cover datasets.

The accuracy of five land cover datasets with both the IGBP scheme and aggregated scheme were estimated in this paper to meet the demands of different users. All classes in the IGBP scheme were retained with no change, except that open and closed shrublands were aggregated into shrublands, the evaluation of classification accuracy of five land cover datasets with IGBP scheme can provide the detailed accuracy of all classes in this scheme.

Satellite data, classification scheme, and classification method used in the five global land cover datasets differ, thus r esulting in evident differences in the classification result. For data users, selecting an appropriate dataset or integrating the a dvantages of different datasets according to their application purpose is very important. For data producers, selecting the optimal satellite data, classification scheme, and methodologies a ccording to the scientific goals of developing land cover dataset is crucial so that the land cover dataset can better meet the d emands of data users.

## REFERENCES

Bartholomé E and 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 [DOI: 10.1080/01431160412331291297]

Bicheron P, Defourny P, Brockmann C, Schouten L, Vancutsem C, Huc M, Bontemps S, Leroy M, Achard F, Herold M, Ranera F and Arino O. 2008. GlobCover Products Description and Validation Report 31401 1. Toulouse: Medias-France

Brown J F, Loveland T R, Ohlen D O and Zhu Z L. 1999. The global land-cover characteristics database: the users' perspective. Photogrammetric Engineering and Remote Sensing, 65: 1069 - 1074

Chen J, Chen J, Gong P, Liao A P and He C Y. 2011. Higher resolution global land cover mapping. Geomatics World, 9(2): 12 - 14

Clark M L, Aide T M, Grau H R and Riner G. 2010. A scalable approach to mapping annual land cover at 250 m using modis time series data: a case study in the Dry Chaco ecoregion of South America. Remote Sensing of Environment, 114(11): 2816 - 2832 [DOI: 10.1016/j.rse.2010.07.001]

Comber A, Fisher P, Brunsdon C and Khmag A. 2012. Spatial analysis of remote sensing image classification accuracy. Remote Sensing of Environment, 127: 237 - 246 [DOI: 10.1016/j.rse.2012.09.005]

Foody G M. 2002. Status of land cover classification accuracy assessment. Remote Sensing of Environment, 80(1): 185 - 201 [DOI: 10.1016/S0034 - 4257(01)00295 - 4]

Foody G M. 2010. Assessing the accuracy of land cover change with imperfect ground reference data. Remote Sensing of Environment, 114(10): 2271 - 2285 [DOI: 10.1016/j.rse.2010.05.003]

Friedl M A, McIver D K, Hodges J C F, Zhang X Y, Muchoney D, Strahler A H, Woodcock C E, Gopal S, Schneider A, Cooper A, Baccini A, Gao F and Schaaf C. 2002. Global land cover mapping from modis: algorithms and early results. Remote Sensing of Environment, 83(1/2): 287 - 302 [DOI: 10.1016/S0034 - 4257(02)00078 - 0]

Friedl M A, Sulla-Menashe D, Tan B, Schneider A, Ramankutty N, Sibley A and Huang X M. 2010. Modis collection 5 global land cover: algorithm refinements and characterization of new datasets. Remote Sensing of Environment, 114(1): 168 - 182 [DOI: 10.1016/j.rse.2009.08.016]

Gao H and Jia G S. 2012. Spatial and quantitative comparison of satellite-derived land cover products over China. Atmospheric and Oceanic Science Letter, 5(5): 426 - 434

Giri C, Zhu Z L and 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 [DOI: 10.1016/j.rse.2004.09.005]

Hansen M C and Reed B. 2000. A comparison of the IGBP DISCover and University Maryland 1 km global land cover products. International Journal of Remote Sensing. 21(6/7): 1365 - 1373 [10.

1080/014311600210209]

Hansen M C, Defries R S, Townshend J R G and Sohlberg R. 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 [DOI: 10.1080/014311600210209]

Herold M, Mayaux P, Woodcock C E, Baccini A and Schmullius C. 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 [DOI: 10.1016/j.rse.2007.11.013]

Li B. 1997. The rangeland degradation in north China and its preventive strategy. Scientia Agricutura Sinica, 30(6): 1 - 9

Li Y C, Gong P, Chen J, Liu C X and He C Y. 2005. Landscape pattern and its dynamical change in north China during 1989—1999. Journal of Soil and Water Conservation, 19(5): 143 - 146

Liang L and Gong P. 2010. An assessment of MODIS collection 5 global land cover product for biological conservation studies // Proceedings of the 2010 18th International Conference on Geoinformatics. Beijing, China: IEEE: 1 - 6 [DOI: 10.1109/GEOINFORMATICS.2010.5567991]

Liu J Y and Buheaosier. 2000. Study on spatial-temporal feature of modern land-use change in China: using remote sensing technology. Quaternary Sciences, 20(3): 229 - 239

Liu J Y, Liu M L, Zhuang D F, Zhang Z X and Deng X Z. 2002. Study on the spatial-temporal feature of land-use change in China. Science in China (Series D), 32(12): 1031 - 1043

Liu J Y, Zhang Z X, Zhuang D F, Wang Y M, Zhou W C, Zhang S W, Li R D, Jiang N and Wu S X. 2003. A study on the spatial-temporal dynamic changes of land-use and driving forces analyses of China in the 1990s. Geographical Research, 22(1): 1 - 12

Liu J Y, Zhang Z X, Xu X L, Kuang W H, Zhou W C, Zhang S W, Li R D, Yan C Q, Yu D S, Wu S X and Jiang N. 2009. Spatial patterns and driving forces of land use change in China in the early 21st century. Acta Geographica Sinica, 64(12): 1411 - 1420

Loveland T R, Reed B C, Brown J F, Ohlen D O, Zhu Z, Yang L and Merchant J W. 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 [DOI: 10.1080/014311600210191]

McCallum I, Obersteiner M, Nilsson S and Shvidenko A. 2006. A spatial comparison of four satellite derived 1km global land cover datasets. International Journal of Applied Earth Observation and Geoinformation, 8(4): 246 - 255 [DOI: 10.1016/j.jag.2005.12.002]

Mei A X, Peng W L, Qin Q M and Liu H P. 2001. Introduction to Remote Sensing. Beijing: High Education Press: 5 - 6

Niu Z G, Shan Y X and Zhang H Y. 2012. Accuracy assessment of wetland categories from the GlobCover2009 data over China. Wetland Science, 10(4): 289 - 395

Ran Y H, Li X and Lu L. 2009. Accuracy evaluation of the four remote sensing based land cover products over China. Journal of Glaciology and Geocryology, 31(3): 490 - 499

Tian G J, Zhuang D F and Liu M L. 2003. The spatial-temporal dynamic change of cultivated land in China in 1990s. Advance in Earth Sciences, 18(1): 30 - 36

Wang S Y, Liu J Y, Zhang Z X, Zhou Q B and Zhao X L. 2001. Analysis on spatial-temporal features of land use in China. Acta Geographica Sinica, 56(5): 631 - 639

Wang S Y, Zhang Z X, Zhou Q B, Zhao X L and Zhang Z K. 2002. Dynamic change of forest land and grassland based on RS and GIS. Resources Science, 24(5): 64 - 69

Wu W B, Shibasaki R, Yang P, Ongaro L, Zhou Q and Tang H J. 2008. Validation and comparison of 1 km global land cover products in China. International Journal of Remote Sensing, 29(13): 3769 - 3785 [DOI: 10.1080/01431160701881897]

Xu W T, Wu B F, Yan C Z and Huang H P. 2005. China land cover 2000 using SPOT VGT S10 data. Journal of Remote Sensing, 9(2): 204 - 214

# 大尺度土地覆盖数据集在中国及周边区域的精度评价

杨永可，肖鹏峰，冯学智，李海星，常潇，冯威丁

1. 南京大学 地理信息科学系，江苏 南京 210023;
2. 南京大学 江苏省地理信息技术重点实验室，江苏 南京 210023

**摘　要**：大尺度土地覆盖数据是全球陆地表层过程研究、生态系统评估、环境建模等科学研究的重要基础，研究现有数据集的特点对数据使用者及生产新的数据集都具有指导意义。本研究以中国及周边区域为研究区，根据不同分类体系对地物的定义，研究不同分类体系中对应地物的相关系数，并将所有分类体系转换为 IGBP 分类体系；然后，从定性和定量两方面分析现有 5 种土地覆盖数据集(IGBP DISCover、UMD、GLC2000、MOD12Q1 和 GlobCover 2005) 的空间一致性；并利用 Google Earth 高分影像选取两期验证样本评价 5 种土地覆盖数据集的精度。结果表明：同种地物在不同土地覆盖数据集之间的空间分布格局差异较大，且不同土地覆盖数据集之间的总体一致性系数较低；5 种土地覆盖数据集中，GLC2000 的总体精度和 Kappa 系数均最高，GlobCover 2005 的总体精度和 Kappa 系数均最低。

**关键词**：土地覆盖数据集，比较分析，精度评价，Google Earth

**中图分类号**：TP79　　**文献标志码**：A

**引用格式**：杨永可，肖鹏峰，冯学智，李海星，常潇，冯威丁. 2014. 大尺度土地覆盖数据集在中国及周边区域的精度评价. 遥感学报，18(2)：453－475

Yang Y K，Xiao P F，Feng X Z，Li H X，Chang X and Feng W D. 2014. Comparison and assessment of large-scale land cover datasets in China and adjacent regions. Journal of Remote Sensing，18(2)：453－475 [DOI：10.11834/jrs.20143055]

## 1 引　言

地表覆盖是地球表面各种物质类型及属性的综合体，准确测定全球地表覆盖的空间分布和动态变化，对研究地球系统的能量平衡、碳循环，及生物地球化学循环和气候变化等有重要意义(陈军 等，2011)。

遥感技术具有可大面积同步观测、方便快捷和经济高效等特点(梅安新 等，2001)，随着时间分辨率和空间分辨率的提高，遥感数据逐渐成为土地覆盖制图的重要数据源。目前常用的全球土地覆盖数据集有 5 种，分别为美国地质调查局(USGS) 研发的全球土地覆盖数据集 IGBP DISCover(International Geosphere-Biosphere Programme Data and Information System Cover)(Loveland 等，2000)；马里兰大学研发的全球土地覆盖数据集 UMD(University of Maryland)(Hansen 等，2000)；波士顿大学研发的全球土地覆盖数据集 MOD12Q1(MODIS Land Cover Type Product collection 4)(Friedl 等，2002)；欧盟联合研究中心研发的全球土地覆盖数据集 GLC2000(Global Land Cover 2000)(Bartholomé 和 Belward，2005)；以及欧洲空间局(ESA) 研发的 300 m 分辨率全球土地覆盖数据集 GlobCover 2005(Global Land Cover Product)(Bicheron 等，2008)。美国研发的 3 种土地覆盖数据集使用的是 IGBP 分类体系，而欧洲研发的两种土地覆盖数据集则使用的是联合国粮农组织制定的 LCCS(Land Cover Classification System) 分类系统。土地覆盖数据是对地表覆盖特征的模拟和概括，能在一定程度上反映真实地表覆盖信息(Brown 等，1999)，但对地表覆盖特征的模拟和概括

收稿日期：2013-03-18；修订日期：2013-09-30；优先数字出版日期：2013-10-09
基金项目：国家重点基础研究发展计划(973 计划)(编号：2011CB952001)
第一作者简介：杨永可(1987—)，男，硕士研究生，专业为遥感数字图像处理、资源环境遥感。E-mail：yangyk198734@126.com
通信作者简介：肖鹏峰(1979—　)，男，博士，副教授，主要研究方向为遥感数字图像处理、资源环境遥感。E-mail：xiaopf@gmail.com

会不可避免的造成信息丢失或错误，故评价土地覆盖数据质量具有重要意义。

全球土地覆盖数据集质量评价方法主要有两种: (1) 利用验证样本计算土覆盖数据集的精度，此时样本数量、质量和抽样设计是制约评价结果客观性的重要因素(Foody,2002,2010)，由于在大尺度上没有真实地表信息可供使用，验证样本选取较为困难。除数据生产方之外，只有 Herold 等人(2008) 使用独立的验证样本(每类选取 25 个样本，每个样本包含 1 个像元) 评价了 4 种 1 km 分辨率土地覆盖数据集在全球尺度上的精度，但全球范围的总体精度无法客观反映局部地区的精度信息(Comber 等，2012)。(2) 利用已有的全球土地覆盖数据集进行比较分析和相互验证，此时由于不同土地覆盖数据集使用的分类数据源、分类体系、分类模型与方法各不相同，造成不同土地覆盖数据集差异较大(Giri 等,2005; McCallum 等,2006)，可比性和兼容性受到限制(Herold 等,2008)。Hansen 和 Reed (2000) 比较了 IGBP DISCover 和 UMD 数据集，指出对于相同的输入数据，不同的分类方法及分类过程中使用的辅助数据是造成分类结果差异的重要因素。Giri 等人(2005) 指出 GLC2000 和 MODIS 产品之间不同地物的面积和空间一致性都会随着区域变化而变化，且在西伯利亚南部至哈萨克斯坦、蒙古及中国边界，以及青藏高原等地区出现大范围不一致现象。M cCallum等人(2006) 指出 4 种 1 km 分辨率全球土地覆盖数据集在亚洲区域内的完全一致性非常低。Liang 和 Gong(2010) 分析了 MODIS 产品制图不确定性，结果显示在大面积均质区域内制图精度较高，而不同地物类型的过渡地带或边缘地区的制图精度则较低。冉有华等人(2009) 以 2000 年中国土地利用数据为参考，以合并后的分类体系(7 类) 为基础，评价了中国区域内的 4 种 1 km 分辨率全球土地覆盖数据集的精度。Wu 等人(2008) 以 2000 年中国土地覆盖数据为基础，研究了 4 种 1 km 分辨率全球土地覆盖数据在中国不同区域的耕地精度，发现 MODIS 和 GLC2000 数据集中耕地精度相对较高; 华北和东部区域的耕地制图精度较高，西北和东南区域的耕地制图精度较低。牛振国等人(2012) 以目视解译的中国 2008 年湿地遥感制图数据为基础，从湿地面积、类型和空间一致性 3 个方面研究了 GlobCover 产品在中国区域内的湿地精度，结果表明 GlobCover 产品在中国区域内湿地的分类精度较低。Gao 和 Jia(2012) 以 2000 年中国土地利用数据为基础，利用合并后的分类体系比较了 MODIS 和 GLC2000 数据集的空间一致性及模糊一致性，并利用独立的验证样本评价 GLC2000 和 MODIS 的精度，认为两种数据集中水体、草地、耕地和裸地精度都较高。

本研究选取中国及周边区域为研究区，以 IGBP 分类体系为基础，根据不同分类体系中的地物定义，分析分类体系的转换关系，并将 5 种土地覆盖数据集转换成 IGBP 分类体系，在分析不同土地覆盖数据集特点的基础上，从定性和定量比较两方面分析不同土地覆盖数据集的空间一致性; 结合 Google Earth 高分影像和其他辅助信息选取验证样本进行精度评价，对精度评价结果进行比较分析。

## 2 研究区与数据

### 2.1 研究区

研究区涵盖中国及周边区域(图 1)，从南向北为热带、亚热带、暖温带、温带、寒温带; 从东向西为海洋性和大陆性气候，东部为季风区; 受高原或高山影响，地形起伏大，垂直结构明显; 地表覆盖特征复杂，景观多样性丰富。研究区内孕育着世界 50% 以上的人口，人类活动强烈，近几十年高速工业化、城市化对该区域土地利用/覆盖类型产生巨大影响，使其成为世界上土地利用/覆盖变化最剧烈地区之一。

### 2.2 土地覆盖数据集

5 种土地覆盖数据集由数据发布网站免费下载，并转成 Lambert 等角投影: (1) IGBP DISCover 1992 年—1993 年 1 km 分辨率土地覆盖数据集; (2) UMD 1992 年—1993 年 1 km 分辨率土地覆盖数据集; (3) GLC 2000 年 1 km 分辨率土地覆盖数据集; (4) MOD12Q1 2001 年 1 km 分辨率土地覆盖数据集; (5) GlobCover 2005 年 300 m 分辨率全球土地覆盖数据集; 5 种土地覆盖数据集的特点见表 1。

大尺度土地覆盖数据集中的城镇和陆地水体图斑面积较小，精确提取城镇和水体信息比较困难，不同数据集使用的方法各不相同。IGBP DISCover的城镇和水体直接使用 DCW(Digital Chart Of World) 中的水文和城镇数据进行掩膜得到(Loveland 等,2000)；UMD 直接使用 IGBP DISCover 的城镇数据，水体则使用为 MODIS 传感器制作的水体数据进行掩膜得到(Hansen 等,2000)；GLC2000

的城镇信息是根据8月下旬每10天最大值合成数据,通过目视解译得到(徐文婷 等,2005);MO12Q1中城镇和水体都使用监督分类树方法进行提取(Friedl 等,2002)。GlobCover 2005的水体数据由ENVISAT卫星携带的MERIS(Medium Resolution Imaging Spectrometer)传感器自带的水/陆边界进行掩膜得到,并结合SRTM(Shuttle Radar Topography Mission)得到的水体数据进行改善,城镇则通过单独的监督分类方法进行提取(Bicheron 等,2008)。

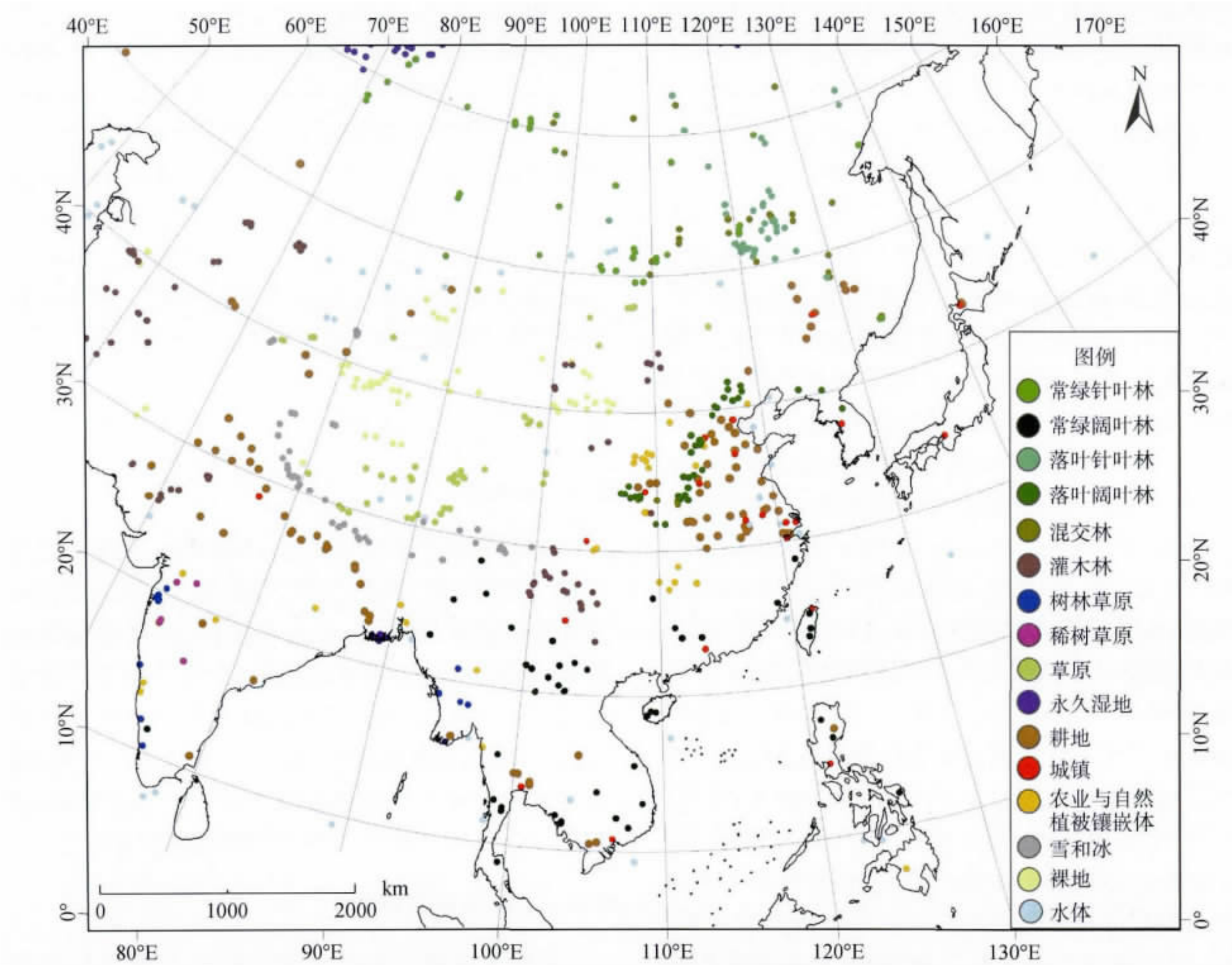


图1 研究区范围与2000年—2001年验证样本分布情况

表1 5种土地覆盖数据集的特点

| | IGBP DISCover | UMD | GLC2000 | MOD12Q1 | GlobCover 2005 |
|---|---|---|---|---|---|
| | AVHRR | AVHRR | SPOT vegetation | MODIS | ENVISAT/MIRIS |
| 输入数据 | 1992-04至1993-04的12期NDVI时间序列数据 | 1992-04至1993-03,由NDVI和5个波段得到的41个规则矩阵 | 中国区域为2000年36期NDVI时间序列数据、加权后的自然因子数据 | 2001年每16天的经BRDF调整后的反射率数据、7个光谱波段,增强型NDVI、雪和冰 | 2004-10至2006-06的13个光谱波段(300 m分辨率) |
| 分类体系 | IGBP 17类 | IGBP 14类 | LCCS 22类 | IGBP 17类 | LCCS 22类 |
| 分类模型与分类方法 | 将全球分为5个区,每个区单独分类。对水体、城镇掩膜后使用K-means分类,再人工解译,解译结果根据Olson分类体系转换为IGBP分类体系 | 以全球为整体进行分类。对水体、城镇掩膜后用监督分类树分类,分类树输出的叶子节点中所有像元赋值为占比例最大类别 | 将全球分为19个区,每个区单独分类。中国按气候分9个区,用ISODATA方法进行分类,并进行人工解译;城镇由目视解译得到 | 以全球为整体进行分类。分类方法为监督的神经网络和分类树方法,监督分类树与UMD使用的方法相似 | 将全球分为22个气候区,每个区单独分类。先进行水体掩膜,然后用监督分类提取湿地和城镇,对剩余像元进行多维聚类,并自动赋值为LCCS分类系统 |

生产不同土地覆盖数据集使用的解译数据(辅助数据)也是造成不同土地覆盖数据集差异的重要因(Hansen 和 Reed,2000)。对于监督分类而言,解译数据用于辅助选取训练样本;对非监督分类,解译数据用来确定聚类图斑属于哪种地物类别。IGBP DISCover 使用的解译数据包括 Landsat 影像、其他已有的土地覆盖产品和地图册等;UMD 使用的解译数据主要是 Landsat MSS(Multispectral Scanning System)影像;GLC2000 中国区域的解译数据包括中国 1∶100 万土地利用图、1∶100 万中国植被图集、中国植被气候分区图和农业气候区划图;MOD12Q1 使用的解译数据主要为 Landsat TM(Thematic Mapper)影像。

### 2.3 参考数据

采用 Landsat TM 影像分类结果(1992 年、2001 年)和 Google Earth 高分影像(2005 年)作为局部空间一致性比较的参考数据。因城镇和陆地水体图斑面积较小,全局尺度上不能反映其在不同土地覆盖数据集之间的差异,故选取杭州及周边区域为典型区进行局部尺度比较分析,并结合参考数据进行验证。

## 3 评价方法与数据预处理

### 3.1 分类体系转换

建立相同的分类体系是 5 种土地覆盖数据集比较分析的基础。IGBP DISCover 和 MOD12Q1 使用的分类体系相同,其他数据集则各有差异,包括地物类别、地物定义和乔木与灌木的树高界限 3 个方面。IGBP DISCover 和 MOD12Q1 中乔木与灌木的树高界限为 5 m;UMD 为 2 m;GLC2000 中国区域为 5 m;GlobCover 2005 为 5 m。已有研究中,不同分类体系的转换关系各不相同(Bartholomé 和 Belward 2005;Giri 等,2005;McCallum 等,2006)。

本研究以 IGBP 分类体系为基础,分析不同分类体系的转换关系。因为 GLC2000 和 GlobCover2005 中将灌木林分为常绿灌木林和落叶灌木林两种,而 IGBP DISCover、UMD 和 MOD12Q1 中却将灌木林分为郁闭灌木林和稀疏灌木林,为了便于进行不同分类体系的转换,将 IGBP 分类体系中的郁闭灌木林和稀疏灌木林合并成灌木林,其他地物类别不变。不同分类体系对地物的定义只有植被覆盖率和树高界限两个量化参数,而树高界限仅用来区分乔木和灌木。5 种土地覆盖数据集中定义的植被覆盖率是一个区间(Hansen 等,2000;Friedl 等,2002;Bartholomé 和 Belward,2005;Bicheron 等,2008),可根据不同土地覆盖数据集定义的植被覆盖率重叠范围分析定义相关性,从而为分类体系转换提供依据。如果两种数据集中植被覆盖率重叠范围越大,则其定义相关性会越强,如果两者定义的植被覆盖率没有重叠区间,则两者之间没有任何关系。本研究根据植被覆盖率来分析不同分类体系的转换关系,并将 5 种土地覆盖数据集转换成 IGBP 分类体系,计算不同土地覆盖数据集与 IGBP 分类体系中对应地物的相关系数 $C$。

$$C = \frac{1}{2}\left(\frac{R_{ID}}{R_I} + \frac{R_{ID}}{R_D}\right) \tag{1}$$

式中,$R_I$ 是 IGBP 分类体系中某类地物的植被覆盖率区间长度,区间长度指植被覆盖率区间上限减去植被覆盖率区间下限,是无单位的量;$R_D$ 是数据集 $D$ 中对应地物的植被覆盖率区间长度,$R_{ID}$ 是数据集 $D$ 和 IGBP 分类体系中对应地物的植被覆盖率重叠区间的长度。相关系数的值域为 0—1,相关系数与重叠区间长度成正比;若无重叠区间则相关系数为 0。水体、城镇、永久湿地和耕地中不涉及植被覆盖率,不能计算相关系数,混交林、农业与自然植被镶嵌体的相关系数无法计算。5 种土地覆盖数据集中地物对应关系及相关系数见表 2。

### 3.2 空间一致性

空间一致性是对不同土地覆盖数据集在相同空间位置上地物类别一致性的描述,用于不同数据集的对比和相互验证,分定性比较和定量比较两方面。定性比较指视觉上的空间分布格局比较;定量比较可计算不同数据集的总体一致性系数 $A$ 和同种地物在不同数据集之间的一致性系数 $B_i$。

$$A = \frac{\sum_{1}^{11} XY_i}{\sum_{1}^{11} (X_i + Y_i)/2} \times 100\% \tag{2}$$

$$B_i = \frac{XY_i}{(X_i + Y_i)/2} \times 100\% \tag{3}$$

式中,$X_i$ 为数据集 $X$ 中 $i$ 类地物总面积,$Y_i$ 为数据集 $Y$ 中 $i$ 类地物总面积,$XY_i$ 为在相同空间位置上数据集 $X$ 和 $Y$ 都为 $i$ 类地物的总面积。由于 UMD 数据集中没有雪和冰、永久湿地、农业与自然植被镶嵌体,而 GLC2000 和 GlobCover 2005 中没有稀疏草原,所以此 4 种地物类型不参与一致性比较。另外,研究区包含大面积的海洋水体,而 5 种土地覆盖

数据集中的水体提取方法不同,使用的水体掩膜数据不同,没有统一的海陆边界,无法将 5 种数据集的陆地水体分离出来进行一致性比较,因而水体也不参与比较。故共有 11 种类型地物参与比较。

表 2 5 种土地覆盖数据集的地物类别对应关系及相关系数

| IGBP DISCover/MOD12Q1 | UMD | GLC2000 | GlobCover 2005 |
|---|---|---|---|
| 常绿针叶林(>60%),1 | 常绿针叶林(>60%),1 | 常绿针叶林(>15%),0.74 | 常绿针叶林(>40%),0.83 |
| 常绿阔叶林(>60%),1 | 常绿阔叶林(>60%),1 | 常绿阔叶林(>15%),0.74 | 常绿阔叶林(>15%),0.74 |
| 落叶针叶林(>60%),1 | 落叶针叶林(>60%),1 | 落叶针叶林(>15%),0.74 | 落叶针叶林(15%—40%),0.74 |
| 落叶阔叶林(>60%),1 | 落叶阔叶林(>60%),1 | 郁闭落叶阔叶林(>40%)、稀疏落叶阔叶(15%—40%),0.74 | 郁闭落叶阔叶林(>40%)、稀疏落叶阔叶林(15%—40%),0.74 |
| 混交林(≤60%),1 | 混交林(≤60%),1 | 混交林(>15%) | 混交林(>15%) |
| 郁闭、稀疏灌木林,1 | 郁闭、稀疏灌木林,1 | 常绿、落叶灌木林(>15%),0.97 | 郁闭灌木林(>15%),0.97 |
| 树林草原(林地 30%—60%),1 | 稀疏林地(林地 40%—60%),0.83 | 林地/其他植被镶嵌体(林地 20%—70%),0.80 | 林地/灌木(50%—70%)与草原(20%—50%)、草原(50%—70%)/林地与灌木(20%—50%)镶嵌体,0.80 |
| 稀树草原(林地 10%—30%),1 | 有林草地(林地 10%—30%),0.83 | — | — |
| 草原(>10%),1 | 草原(>10%),1 | 草原(>15%),0.97 | 草原(>15%),0.97 |
| 永久湿地 | — | 定期淹没林(淡水)、滨海湿地、沼泽 | 定期淹没阔叶林(淡水)、海滨湿地、沼泽 |
| 耕地 | 耕地 | 耕地 | 水田、旱地 |
| 城镇 | 城镇 | 城镇 | 城镇 |
| 农业与自然植被镶嵌体 | — | 农业与自然植被镶嵌体 | 耕地(50%—70%)与自然植被(20%—50%)、自然植被(50%—70%)与耕地(20%—50%)镶嵌体, |
| 雪和冰 | — | 雪和冰 | 雪和冰 |
| 裸地(<10%),1 | 裸地(<10%),1 | 稀疏草地或灌木、裸地(<15%),0.97 | 稀疏植被、裸地(<15%),0.97 |
| 水体 | 水体 | 水体 | 水体 |

注:括号内数字表示植被覆盖率范围;逗号后面的数字表示相应地物在不同分类体系中的相关系数 "无"表示无此类型地物。

### 3.3 验证样本获取

随着 Google Earth 高分影像可用性增强,它逐渐成为一种重要数据源,很多学者开始使用其进行土地覆盖制图或精度评价研究(Bicheron 等,2008;Friedl 等,2010;Clark 等,2010)。以 Google Earth 为工具选取样本优势如下:(1)免费查看高分影像,可浏览全球范围的虚拟视图;(2)样本整体视图,可在不同视角、高度和比例尺条件下查看样本整体情况;(3)定位精度高,定位误差约为(15±5) m,满足粗分辨率精度要求(Clark 等,2010);(4)局部区域照片,为样本判定提供重要依据;(5)样本时效性,根据时间轴确定样本时间的有效区间。

利用 Google Earth 获取的验证样本主要包含空间位置误差和解译误差。空间位置误差主要由地理参考系统和地形因素造成,比如高分影像未进行正射校正可能造成的定位误差,以及地理参考系统不同造成的空间位置偏移。本研究中使用的 5 种土地覆盖数据集包含 1 km 和 300 m 两种分辨率,而 Google Earth 的定位误差约为(15±5) m(Clark 等,2010),可满足 1 km 和 300 m 分辨率土地覆盖数据的空间位置精度要求;样本选好之后,将其转投影成 Lambert 等面积投影,使验证样本和 5 种土地覆盖数据集的地理参考系统相同,消除由不同的地理参考系统引起的空间位置偏移。解译误差主要由主观因素造成,如不同的学科背景对相同的影像会有不同的认识。影像在样本范围内的地物均质性也是影响样本精度的重要因素。为了降低解译误

差和保证样本均质性，本研究以不同类型地物的影像特征和空间分布格局为先验知识，制定如下抽样规则：（1）样本须在大面积均质区域中心选取，每个样本大小为4个像素（约2 km×2 km或600 m×600 m），5种土地覆盖数据集的空间分辨率不同，在选取验证样本时需考虑空间分辨率差异引起的误差。假设在1 km×1 km格网内有60%的针叶林和40%的阔叶林，则在1 km分辨率尺度下可将其土地覆盖类型定义为混交林；但是，在300 m分辨率尺度下则不能将其包含的所有像元的土地覆盖类型定义为混交林，否则会出现空间分辨率引起的误差。因而，选择验证样本时应充分考虑空间分辨率的差异，而不能用相同大小的验证样本评价不同空间分辨率的土地覆盖产品。（2）样本须在有高分影像的区域选取。（3）对常绿和落叶林地，需结合Google Earth时间轴查看不同时相的高分影像，例如，常绿针叶林、常绿阔叶林的判读依据主要是冬季的高分影像；落叶针叶林、落叶阔叶林则需要结合冬季和其他季节的高分影像。（4）对于部分难以确定地物类型的样本，需利用Google Earth提供的照片信息进行确定。

高分辨率卫星遥感影像从2000年才开始走向民用，且没有可使用的大尺度真实地表数据，故无法获取1992年—1993年的样本，只能选取两期（2000年—2001年，2004年—2006年）验证样本。为了使2000年—2001年的验证样本能反映1992年—1993年真实的地表信息，选取2000年—2001年样本时，要求每个样本必须在大面积均质区域中心选取，避开土地利用变化明显的区域，如不同地物类型的过度地带和边缘地区，以及农林、农牧交错带。因为已有研究显示，中国土地利用变化主要在传统的农作区（包括黄淮海平原、长江三角洲地区和四川盆地）内建筑用地扩张占用大面积的耕地，北方的农牧、农林交错地带与西北绿洲地区大量林地、草地被开垦为耕地，西部地区则变化相对缓慢，退耕还林还草政策实施效果在局部区域有所体现，但退耕还林还草的速度低于林地草地被开开垦的速度；主要的土地利用变化类型是耕地—城镇、林草—耕地之间的转换（刘纪远和布和敖斯尔，2000；王思远 等，2001；王思远 等，2002；刘纪远 等，2002；刘纪远 等，2003；田光进 等，2003；李月臣 等，2005；刘纪远 等，2009）。且北方的草地退化比较严重，在气候和人为因素影响下草地由高密度草地向低密度草地和荒漠化转变（李博，1997）。故避开20世纪90年代土地利用变化明显区域，就可使2000年—2001年的验证样本最大程度地反映1992年—1993年真实地表信息，并用其评价1992年—1993年土地覆盖数据精度。

2000年—2001年的验证样本用于评价4种1 km分辨率产品，选样时间轴在2000年—2001年间；2004年—2006年的验证样本用于评价300 m分辨率产品（GlobCover 2005），时间轴在2004年—2006年间，且是以2000年—2001年验证样本为基础进行修改的，当2000年的样本满足时效性时保留该样本，将其范围调整至600 m×600 m，否则重新选取。不同数据集使用的验证样本类别及数量如表3。

表3 5种土地覆盖数据集使用的验证样本情况

| 类别名称 | IGBP DISCover /MOD12Q1 | UMD | GLC2000 | GlobCover 2005 |
|---|---|---|---|---|
| 常绿针叶林 | 37 | 37 | 37 | 43 |
| 常绿阔叶林 | 44 | 44 | 44 | 45 |
| 落叶针叶林 | 37 | 37 | 37 | 31 |
| 落叶阔叶林 | 37 | 37 | 37 | 39 |
| 混交林 | 28 | 28 | 28 | 25 |
| 灌木林 | 69 | 69 | 20 | 22 |
| 树林草原 | 13 | 13 | 13 | 14 |
| 稀树草原 | 5 | 5 | — | — |
| 草原 | 44 | 44 | 44 | 53 |
| 永久湿地 | 25 | — | 25 | 25 |
| 耕地 | 101 | 101 | 101 | 103 |
| 城镇 | 22 | 22 | 22 | 38 |
| 农业与自然植被镶嵌体 | 34 | — | 34 | 32 |
| 雪和冰 | 43 | — | 43 | 35 |
| 裸地 | 50 | 50 | 50 | 49 |
| 水体 | 49 | 49 | 49 | 52 |
| 总样本数 | 638 | 536 | 584 | 606 |

### 3.4 混淆矩阵分析

混淆矩阵分析是进行土地覆盖数据集精度评价常用的方法，在精度评价中至关重要（Foody，2002；Herold 等，2008；冉有华 等 2009；Clark 等，2010），是地表真实信息与分类结果之间的一张交叉2维表，可提供总体精度、生产者精度和用户精度等精度信息，且通过混淆矩阵生成的Kappa系数可衡量地表真实信息和分类结果的总体一致性程度。

$$总体精度 = \frac{\sum_{i=1}^{17} X_{ii}}{N^2} \times 100\% \tag{4}$$

$$\text{Kappa 系数} = \frac{N\sum_{i=1}^{17} X_{ii} - \sum_{i=1}^{17}(X_{i+}X_{+1})}{N^2 - \sum_{i=1}^{17}(X_{i+}X_{+1})} \tag{5}$$

$$用户精度 = \frac{X_{ii}}{X_{i+}} \times 100\% \tag{6}$$

$$生产者精度 = \frac{X_{ii}}{X_{+i}} \times 100\% \tag{7}$$

式中,$X_{ii}$是 $i$ 类地物正确分类的像元数,$N$ 表示所有地物的总像元数,$X_{i+}$ 分类图中 $i$ 类地物的总像元数,$X_{+i}$表示验证样本中 $i$ 类地物的总像元数。

# 4 结果与讨论

## 4.1 空间一致性分析

一致性比较是不同数据集之间的相互验证,因为5种数据集的时间相差十几年,且研究区内土地利用/覆盖变化非常快,从理论上讲,时间差异对不同数据集之间的一致性比较会有一定影响。然而,事实上数据集之间由分类引起差异远远大于地表真实变化信息(Friedl 等,2010; McCallum 等,2006)。研究(刘纪远 等,2003)显示20世纪90年代中国区域内主要有13个土地利用动态区,在13个动态区内土地利用变化总面积约为49114.6 km$^2$,仅为全国总面积的0.5%左右。另外,一致性比较显示,即使两种数据集的时间相同,其一致性也非常低。例如,IGBP DISCover 和 UMD 的总体一致性仅为37%左右。即不同数据集之间的差异主要是由分类引起的,虽然不同数据集的时间不同,进行一致性分析仍然有一定的意义。

### 4.1.1 定性比较

定性比较可从全局比较和局部比较两方面进行。全局比较用于分析相同地物在5种土地覆盖数据集中的空间分布格局特征及差异,局部比较用于选取典型区分析城镇和水体在不同土地覆盖数据集中的细节特征及差异。

全局比较分两级进行,首先将常绿针叶林、常绿阔叶林、落叶针叶林、落叶阔叶林、混交林和灌木林合并为林地,比较林地、树林草原、草原、耕地和裸地在5种土地覆盖数据集中的空间分布格局(图2)。然后单独比较常绿针叶林、常绿阔叶林、落叶针叶林、落叶阔叶林、混交林和灌木林的空间分布格局(图3)。

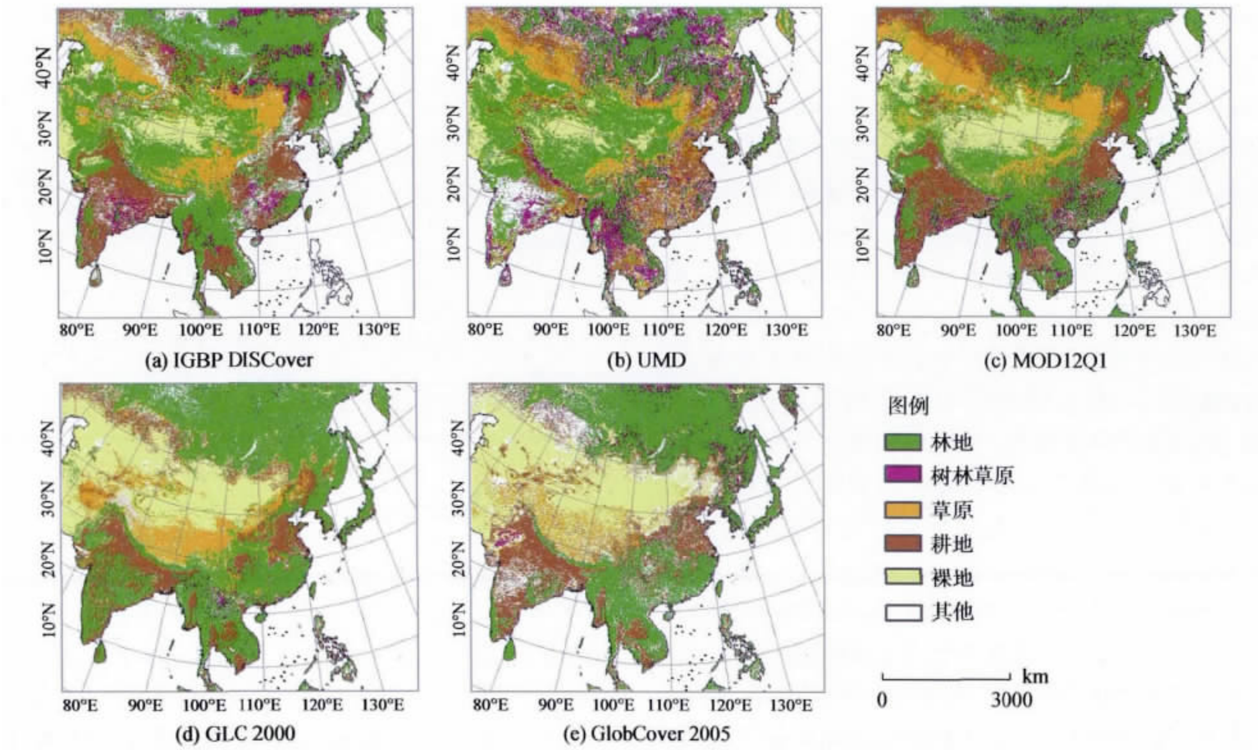


图2 林地、树林草原、草原、耕地和裸地的空间分布格局比较

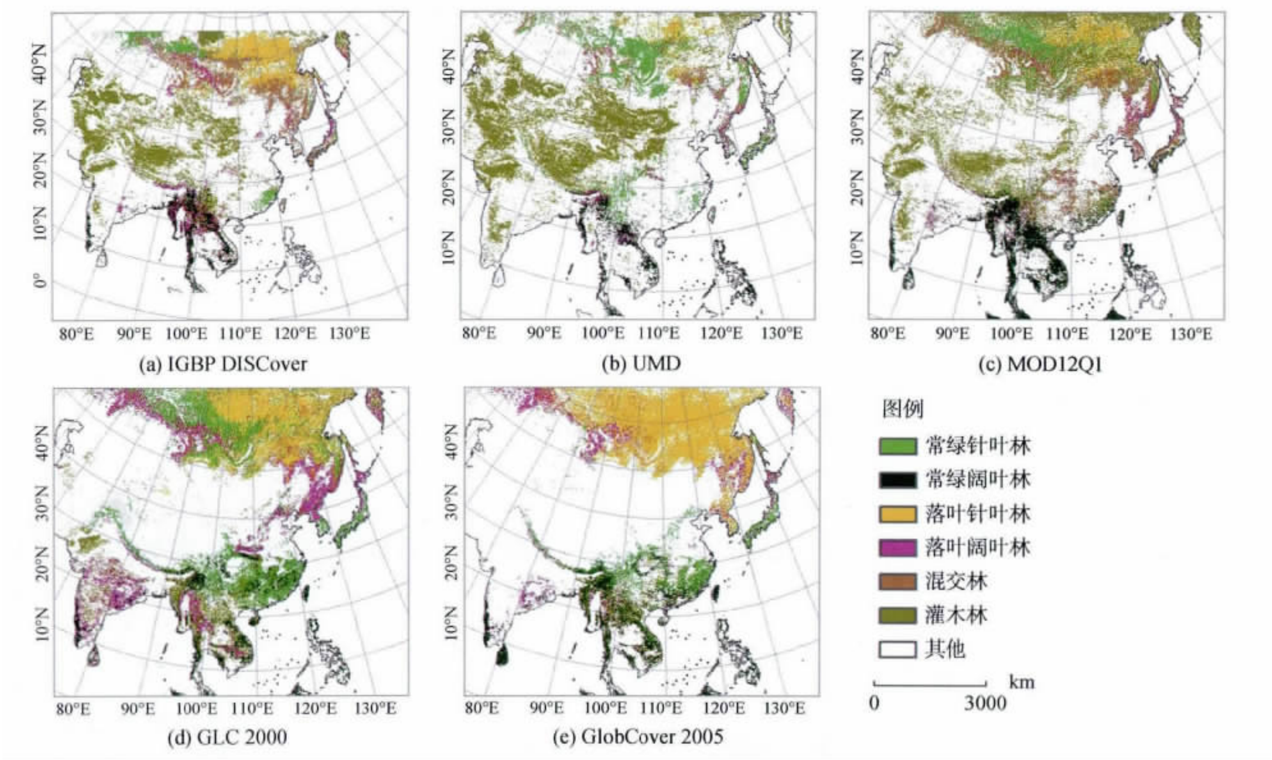


图3 5 种乔木林地的空间分布格局比较

图2 显示,5 种数据集在俄罗斯、中亚地区(包含哈萨克斯坦及其以南至巴基斯坦地区)、印度、中国青藏高原西部、中国东南沿海地区和蒙古地区存在大面积不一致现象。UMD 数据集在俄罗斯地区的林地偏少,在印度地区的耕地偏少;且 UMD 和 IGBP DISCover 中的树林草原均较多。另外,根据不同土地覆盖数据集使用的原始分类体系进行比较发现,在中亚地区、蒙古和青藏高原地区,IGBP DISCover、UMD 和 MOD12Q1(原始分类体系都为 IGBP) 3 种数据集的一致性很高,而 GLC2000 和 GlobCover 之间的一致性很高,而两种分类体系之间的差异较大。图 3 显示,6 种林地在 5 种土地覆盖数据集之间的差异同样跟土地覆盖数据集使用的原始分类体系密切相关,如 GLC2000 和 GlobCover 2005 中灌木林相对偏少,落叶针叶林相对偏多;而 IGBP DISCover、UMD 和 MOD12Q1 中灌木林相对较多,而落叶针叶林则相对较少。结合图 2 和图 3 发现,在上述 6 个大面积不一致区域内,5 种土地覆盖数据集之间不同地物的错分现象主要表现为:草原与裸地错分(哈萨克斯坦、蒙古北部)、草原与灌木林错分(中国青藏高原西部),裸地与灌木林错分(中亚地区、蒙古南部),常绿针叶林、常绿阔叶林与耕地错分(中国东南沿海地区),落叶阔叶林与耕地错分(印度地区),5 种乔木林地之间的错分(俄罗斯地区)。同时还发现,灌木林与 5 种乔木林地之间的错分现象较少,林地之间的错分现象主要是 5 种乔木林地之间的相互错分。

5 种土地覆盖数据集与参考数据在杭州地区的局部比较如图 4 所示。与 1992 年 TM 分类结果相比,IGBP DISCover 和 UMD 中杭州地区的城镇面积相对偏小,而水体面积均偏大;另外,UMD 数据集中将杭州地区大面积耕地错分为草原,且在杭州地区出现较多的树林草原。与 2001 年 TM 分类结果相比,GLC2000 在杭州地区的水体面积严重夸大(西湖面积偏大),而钱塘江却消失,且城镇(杭州市区)与水体(西湖)的位置出现错误;MOD12Q1 在杭州地区的城镇面积严重高估,水体信息比较准确。GlobCover 2005 城镇和水体信息与 Google Earth 高分影像较吻合,但是在西湖西南角的植被错分为城镇。

### 4.1.2 定量比较

定量比较是对不同土地覆盖数据集整体区域空间一致性的量化描述,按时间节点计算不同土地覆盖数据集的总体一致性和具体地物一致性系数。同地物在不同土地覆盖数据集之间的一致性系数及面积比较如图 5 所示。

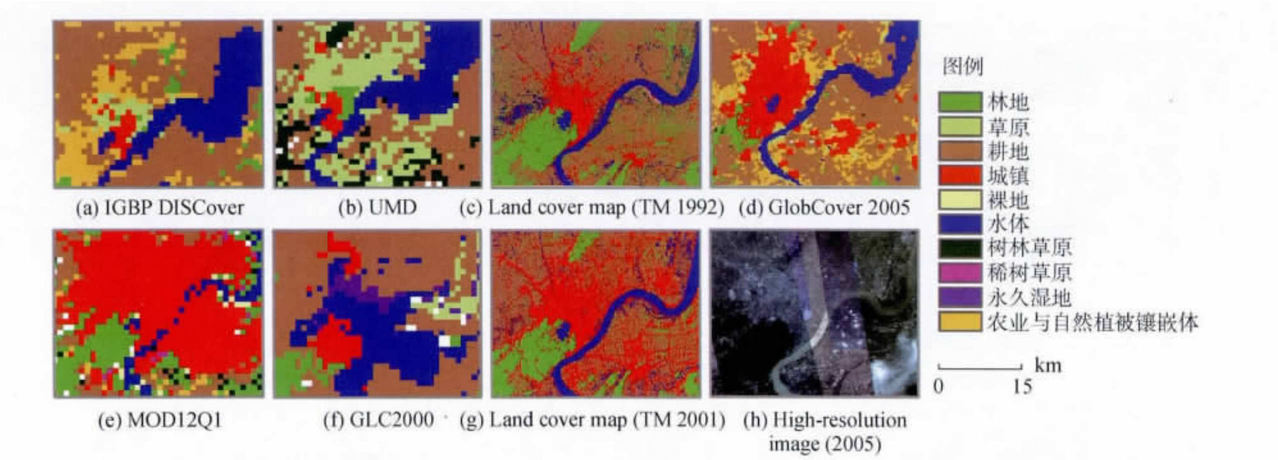


图4 5种土地覆盖数据集与参考数据在杭州地区的局部比较

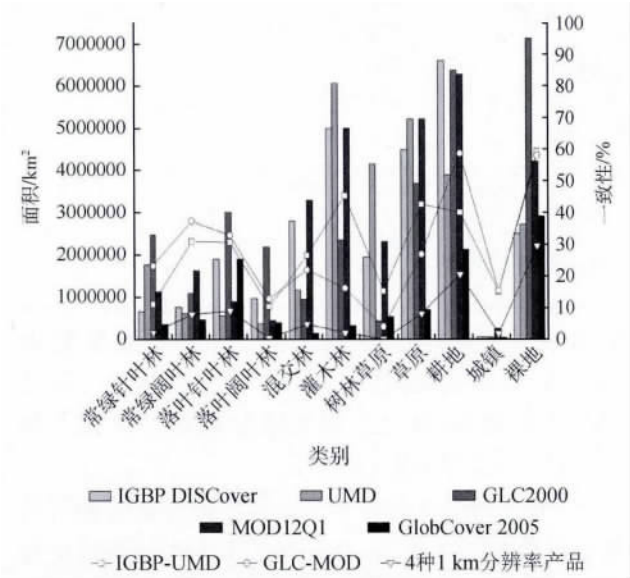


图5 不同土地覆盖数据集的面积与空间一致性

图5显示同种地物在不同土地覆盖数据集之间的面积差异较大,相同地物在不同土地覆盖数据集之间的一致性系数差异较大,而且不同土地覆盖数据集之间的一致性曲线具有相似的变化趋势。IGBP DISCover和UMD总体一致性为36.93%;单类别一致性最高的是裸地,达到59.35%;灌木林、草原和耕地分别为45.32%、42.65%和40.05%;尽管UMD的城镇数据由IGBP DISCover中的城镇数据掩膜得到,但两者的城镇一致性系数仅为15.14%,主要是因为两种数据集的原始投影不同,经过多次投影转换后城镇的位置发生偏移。GLC2000和MOD12Q1总体一致性为36.67%;单类别一致性最高的是耕地,达到58.66%;裸地为58.04%;常绿阔叶林、落叶针叶林、常绿针叶林和草原则依次降低,而城镇、灌木林和树林草原的空间一致性非常低。

4种土地覆盖数据集的完全一致性分析结果显示,4种土地覆盖数据集的总体一致性仅有11.30%;单类别一致性最高的是裸地,为29.54%;其次是耕地,为20.41%;其他类别的一致性则都低于10%。另外,相同地物在不同土地覆盖数据之间的面差异较大,如常绿针叶林、落叶阔叶林、混交林和树林草原等。

### 4.2 精度评价

利用验证样本分别建立5种土地覆盖数据集在整体区域的混淆矩阵,得到不同地物的用户精度和生产者精度,及总体精度和Kappa系数,受篇幅限制,未将混淆矩阵放入文中,不同地物的用户精度和生产者精度如图6和图7所示,树林草原、雪和冰、永久湿地、农业与自然植被镶嵌体4种地物的生产者精度和用户精度见表4。

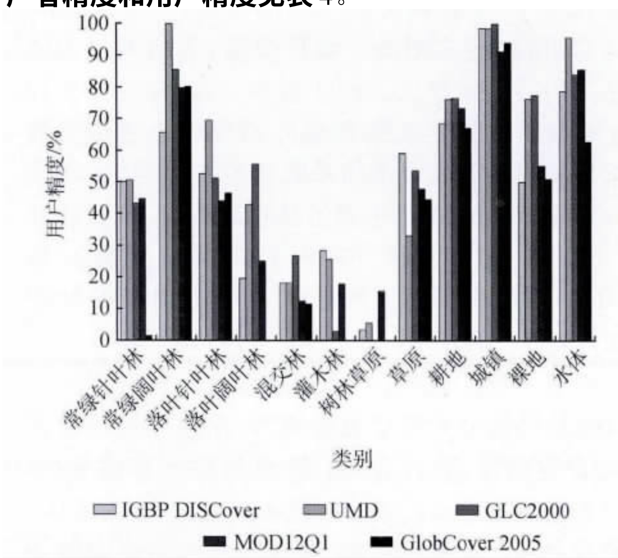


图6 不同地物的用户精度比较

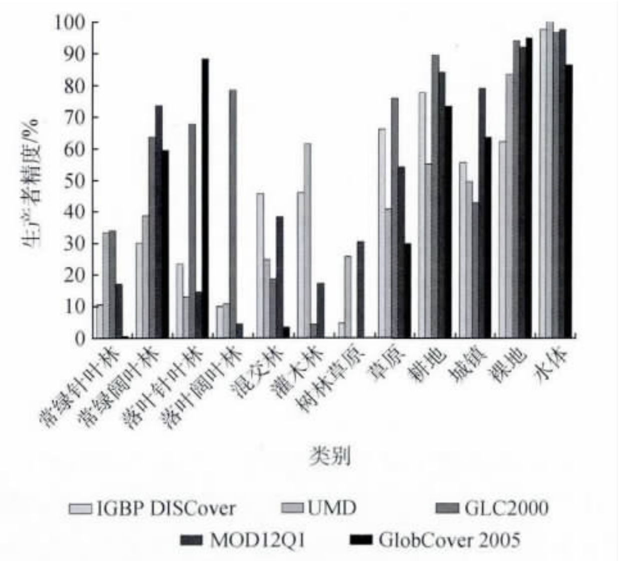


图7 不同地物的生产者精度比较

图6 显示,所有地物类型中,常绿阔叶林、耕地、城镇和水体的用户精度相对最高,且差异较小; 常绿针叶林、落叶针叶林的用户精度次之,在 50% 左右; 混交林、灌木林和树林草原的用户精度相对最低,且差异较大; 裸地在 UMD 和 GLC2000 中用户精度约为 80%,而在 IGBP DISCover、MOD12Q1 和 GlobCover 2005 中只有 50% 左右。

图7 显示,所有地物类型中,耕地、裸地和水体的生产者精度相对最高,且差异较小; 常绿针叶林和树林草原的生产者精度相对最低,且差异较大; 而落叶针叶林、落叶阔叶林、混交林、灌木林和草原等的生产者精度差异较大。

5 种土地覆盖数据集中城镇的用户精度都在 90% 之上,但其生产者精度差异却较大。IGBP DISCover 与 UMD 的中城镇的生产者精度较低,主要是因为其使用的 DCW 中的城镇数据是由 1960 年—1980 年的多源地图数据融合得到的,在城市化高速发展的时代,这些过时的数据不能准确的反映城镇的变化信息,IGBP DISCover 中有 14.29% 的城镇被分成了耕地,8.27% 的城镇被分成了农业与自然植被镶嵌体; IGBP DISCover 和 UMD 中灌木林的生产者精度较高,而其他土地覆盖数据集中灌木林的生产者精度却很低,与图 3 中其他土地覆盖数据集中将灌木林错分到裸地和草原有关。UMD 中则有 15.04% 的城镇被分为耕地。GLC2000 的城镇生产者精度最低,与中国区域内的城镇是由目视解译得到的,受主观因素影响较大有关,因为不同的解译者对相同的影像有不同的认识,有 23.30% 的城镇被分为耕地。MOD12Q1 的城镇生产者精度最高,主要是因为利用监督分类树方法进行分类时,最终输出的叶子节点是按不同地物所占比例赋值的,若城镇所占的比例最大,则所有像元全部赋值为城镇,故城镇面积会有一定程度的夸大,图 3 显示,城镇周围的耕地与其他地物会被分为城镇,使生产者精度被高估。Globover 2005 的草原生产者精度较低,主要是因为有 24.62% 的草原被分为裸地。

表 4 显示,5 种土地覆盖数据集中,永久湿地及雪和冰的用户精度均较高,生产者精度差异却较大; UMD 的稀树草原用户精度和生产者精度最高,其他数据集则较低,主要因为对稀树草原训练样本的采样不仅仅局限于热带区域,其他地区具有类似地表覆盖特征的地物也按稀树草原进行采样。

表 4 5 种土地覆盖数据集中稀树草原、永久湿地、农业与自然植被镶嵌体及雪和冰的精度信息

/%

| | IGBP DISCover | | UMD | | GLC2000 | | MOD12Q1 | | GlobCover 2005 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 生产者精度 | 用户精度 | 生产者精度 | 用户精度 | 生产者精度 | 用户精度 | 生产者精度 | 用户精度 | 生产者精度 | 用户精度 |
| 稀树草原 | 0.00 | 0.00 | 57.81 | 35.34 | — | — | 5.26 | 3.13 | — | — |
| 永久湿地 | 45.45 | 90.91 | — | — | 63.64 | 82.35 | 31.82 | 93.33 | 14.73 | 95.56 |
| 农业与自然植被镶嵌体 | 30.30 | 68.66 | — | — | 3.79 | 15.63 | 1.52 | 8.70 | 19.54 | 8.92 |
| 雪和冰 | 37.65 | 98.67 | — | — | 87.65 | 97.93 | 14.81 | 100.00 | 53.54 | 95.50 |

5 种土地覆盖数据集的混淆矩阵显示,灌木林和草原、农业与自然植被镶嵌体和耕地、稀树草原和树林草原、耕地和草原、耕地与城镇、裸地与草原、5 种乔木林地之间的错分现象较多,对不同地物的生产者精度和用户精度影响较大。部分实际需求只需一级林地的精度,而非林地二级分类的精度,故还计算了 5 种乔木林地合并成林地后 5 种数据集的总体精度、Kappa 系数,林地的生产者精度和用户精度(表 5)。

表5 5种林地合并前后5种土地覆盖数据集的精度信息

| | 总体精度/% | | Kappa 系数 | | 林地精度/% | |
|---|---|---|---|---|---|---|
| | 合并前 | 合并后 | 合并前 | 合并后 | 生产者精度 | 用户精度 |
| IGBP DISCover | 51.58 | 60.75 | 0.47 | 0.55 | 60.77 | 86.01 |
| UMD | 56.71 | 58.94 | 0.52 | 0.53 | 50.09 | 77.93 |
| GLC2000 | 67.72 | 76.06 | 0.64 | 0.73 | 86.83 | 87.13 |
| MOD12Q1 | 53.19 | 59.44 | 0.48 | 0.54 | 69.43 | 82.89 |
| GlobCover 2005 | 51.12 | 58.33 | 0.45 | 0.51 | 71.83 | 76.89 |

表5显示,5种林地合并后,5种土地覆盖数据集中林地用户精度都在75%以上,且差异不大;而林地的生产者精度差异却较大;5种土地覆盖数据集的总体精度和Kappa系数都有不同程度的提高。IGBP DISCover的总体精度提高的最多,达到9.17%;GLC2000、MOD12Q1和GlobCover 2005分别提高了8.34%、6.25%和7.21%;UMD仅提高了2.23%。IGBP DISCover和GLC2000的Kappa系数均提高0.8,MOD12Q1、GlobCover 2005均提高0.6,UMD仅提高了0.1。5种林地合并前后GLC2000的总体精度和Kappa系数均为最高,而GlobCover 2005的总体精度和Kappa系数均为最低。

## 5 结 论

5种土地覆盖数据集的比较分析和精度评价结果显示,相同地物在不同土地覆盖数据集之间的空间分布格局差异和一致性系数差异均较大;相同地物在原始分类体系相同的土地覆盖数据集之间的空间分布格局较为吻合;而在原始分类体系不同的土地覆盖数据集之间的空间分布格局差异和一致性系数差异均较大;5种土地覆盖数据集中,GLC2000的总体精度最高,GlobCover 2005的总体精度最低,5种土地覆盖数据集在局部区域都有明显错分现象;所有地物类型中,只有耕地、裸地和水体3类地物同时具有相对较高的生产者精度和用户精度;其他地物的生产者精度与用户精度差异较大。

本研究中只将IGBP分类体系中的郁闭灌木林和稀疏灌木林合并为灌木林,其他地物不变,最大限度保持了IGBP分类体系的完整性,且通过Google Earth可获得IGBP分类体系中所有地物类别的样本,故精度评价结果提供了5种土地覆盖数据集中丰富地物的精度信息,如林地二级分类精度,并计算了一级林地(5种乔木林地合并后)分类结果精度,可满足不同的用户需求。

受分类数据源、分类模型与方法、分类体系等因素影响,5种土地覆盖数据集之间的差异较大,且各有优缺点。用户可根据实际需求选择合适的数据集,或者综合使用现有数据集;对数据生产者而言,可根据数据集的应用目的,选择合适的分类数据源、分类模型与方法和分类体系,以便生产出满足应用需要的土地覆盖数据。

## 参考文献(References)

Bartholomé E and 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 [DOI: 10.1080/01431160412331291297]

Bicheron P, Defourny P, Brockmann C, Schouten L, Vancutsem C, Huc M, Bontemps S, Leroy M, Achard F, Herold M, Ranera F and Arino O. 2008. GlobCover Products Description and Validation Report 31401 1. Toulouse: Medias-France

Brown J F, Loveland T R, Ohlen D O and Zhu Z L. 1999. The global land-cover characteristics database: the users' perspective. Photogrammetric Engineering and Remote Sensing, 65: 1069-1074

陈军,陈晋,宫鹏,廖安平,何超英. 2011. 全球地表覆盖高分辨率遥感制图. 地理信息世界,9(2): 12-14

Clark M L, Aide T M, Grau H R and Riner G. 2010. A scalable approach to mapping annual land cover at 250 m using modis time series data: a case study in the Dry Chaco ecoregion of South America. Remote Sensing of Environment, 114(11): 2816-2832 [DOI: 10.1016/j.rse.2010.07.001]

Comber A, Fisher P, Brunsdon C and Khmag A. 2012. Spatial analysis of remote sensing image classification accuracy. Remote Sensing of Environment, 127: 237-246 [DOI: 10.1016/j.rse.2012.09.005]

Foody G M. 2002. Status of land cover classification accuracy assessment. Remote Sensing of Environment, 80(1): 185-201 [DOI: 10.1016/S0034-4257(01)00295-4]

Foody G M. 2010. Assessing the accuracy of land cover change with imperfect ground reference data. Remote Sensing of Environment, 114(10): 2271-2285 [DOI: 10.1016/j.rse.2010.05.003]

Friedl M A, McIver D K, Hodges J C F, Zhang X Y, Muchoney D, Strahler A H, Woodcock C E, Gopal S, Schneider A, Cooper A, Baccini A, Gao F and Schaaf C. 2002. Global land cover mapping from modis: algorithms and early results. Remote Sensing of Environment, 83(1/2): 287 - 302 [DOI: 10.1016/S0034 - 4257(02)00078 - 0]

Friedl M A, Sulla-Menashe D, Tan B, Schneider A, Ramankutty N, Sibley A and Huang X M. 2010. Modis collection 5 global land cover: algorithm refinements and characterization of new datasets. Remote Sensing of Environment, 114(1): 168 - 182 [DOI: 10.1016/j.rse.2009.08.016]

Gao H and Jia G S. 2012. Spatial and quantitative comparison of satellite-derived land cover products over China. Atmospheric and Oceanic Science Letter, 5(5): 426 - 434

Giri C, Zhu Z L and 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 [DOI: 10.1016/j.rse.2004.09.005]

Hansen M C and Reed B. 2000. A comparison of the IGBP DISCover and University Maryland 1 km global land cover products. International Journal of Remote Sensing. 21(6/7): 1365 - 1373 [10.1080/014311600210209]

Hansen M C, Defries R S, Townshend J R G and Sohlberg R. 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 [DOI: 10.1080/014311600210209]

Herold M, Mayaux P, Woodcock C E, Baccini A and Schmullius C. 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 [DOI: 10.1016/j.rse.2007.11.013]

**李博**. 1997. **中国北方草地退化及其防治对策. 中国农业科学**, 30(6): 1 - 9

**李月臣, 宫鹏, 陈晋, 刘春霞, 何春阳**. 2005. **中国北方 13 省土地利用景观格局变化分析. 水体保持学报**, 19(5): 143 - 146

Liang L and Gong P. 2010. An assessment of MODIS collection 5 global land cover product for biological conservation studies // Proceedings of the 2010 18th International Conference on Geoinformatics. Beijing, China: IEEE: 1 - 6 [DOI: 10.1109/GEOINFORMATICS.2010.5567991]

**刘纪远, 布和敖斯尔**. 2000. **中国土地利用变化现代过程时空特征的研究——基于卫星遥感数据. 第四纪研究**, 20(3): 229 - 239

**刘纪远, 刘明亮, 庄大方, 张增祥, 邓祥征**. 2002. **中国近期土地利用变化的空间格局分析. 中国科学**(D 辑), 32(12): 1031 - 1043

**刘纪远, 张增祥, 庄大方, 王一谋, 周万村, 张树文, 李仁东, 江南, 吴世新**. 2003. **20 世纪 90 年代中国土地利用变化时空特征及其成因分析. 地理研究**, 22(1): 1 - 12

**刘纪远, 张增祥, 徐新良, 匡文慧, 周万村, 张树文, 李仁东, 颜长珍, 于东升, 吴世新, 江南**. 2009. **21 世纪初中国土地利用变化的空间格局与驱动力分析. 地理学报**, 64(12): 1411 - 1420

Loveland T R, Reed B C, Brown J F, Ohlen D O, Zhu Z, Yang L and Merchant J W. 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 [DOI: 10.1080/014311600210191]

McCallum I, Obersteiner M, Nilsson S and Shvidenko A. 2006. A spatial comparison of four satellite derived 1km global land cover datasets. International Journal of Applied Earth Observation and Geoinformation, 8(4): 246 - 255 [DOI: 10.1016/j.jag.2005.12.002]

**梅安新, 彭望琭, 秦其明, 刘慧平**. 2001. **遥感导论. 北京: 高等教育出版社**: 5 - 6

**牛振国, 单玉秀, 张海英**. 2012. **全球土地覆盖 GlobCover2009 数据中的中国区域湿地数据精度评价. 湿地科学**, 10(4): 289 - 395

**冉有华, 李新, 卢玲**. 2009. **四种常用的全球 1km 土地覆盖数据中国区域的精度评价. 冰川冻土**, 31(3): 490 - 499

**田光进, 庄大方, 刘明亮**. 2003. **近 10 年中国耕地资源时空变化特征. 地球科学进展**, 18(1): 30 - 36

**王思远, 刘纪远, 张增祥, 周全斌, 赵晓丽**. 2001. **中国土地利用时空特征分析. 地理学报**, 56(5): 631 - 639

**王思远, 张增祥, 周全斌, 赵晓丽, 张宗科**. 2002. **遥感与 GIS 支持下的中国森林植被动态变化分析. 资源科学**, 24(5): 64 - 69

Wu W B, Shibasaki R, Yang P, Ongaro L, Zhou Q and Tang H J. 2008. Validation and comparison of 1 km global land cover products in China. International Journal of Remote Sensing, 29(13): 3769 - 3785 [DOI: 10.1080/01431160701881897]

**徐文婷, 吴炳方, 颜长珍, 黄慧萍**. 2005. **用 SPOT-VGT 数据制作中国 2000 年度土地覆盖数据. 遥感学报**, 9(2): 204 - 214