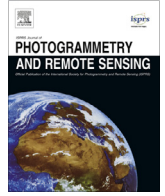




Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Global land cover mapping at 30 m resolution: A POK-based operational approach



Jun Chen^{a,*}, Jin Chen^b, Anping Liao^a, Xin Cao^b, Lijun Chen^a, Xuehong Chen^b, Chaoying He^a, Gang Han^a, Shu Peng^a, Miao Lu^a, Weiwei Zhang^a, Xiaohua Tong^c, Jon Mills^d

^a National Geomatics Center of China, Beijing 100830, China

^b State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing Normal University, Beijing 100875, China

^c College of Surveying and Geo-Informatics, Tongji University, Shanghai 200092, China

^d School of Civil Engineering and Geosciences, Newcastle University, Newcastle upon Tyne NE1 7RU, United Kingdom

ARTICLE INFO

Article history:

Received 6 August 2014

Received in revised form 30 August 2014

Accepted 3 September 2014

Available online 19 October 2014

Keywords:

Global land cover

Operational approach

Pixel-object-knowledge-based classification

GlobeLand30

ABSTRACT

Global Land Cover (GLC) information is fundamental for environmental change studies, land resource management, sustainable development, and many other societal benefits. Although GLC data exists at spatial resolutions of 300 m and 1000 m, a 30 m resolution mapping approach is now a feasible option for the next generation of GLC products. Since most significant human impacts on the land system can be captured at this scale, a number of researchers are focusing on such products. This paper reports the operational approach used in such a project, which aims to deliver reliable data products.

Over 10,000 Landsat-like satellite images are required to cover the entire Earth at 30 m resolution. To derive a GLC map from such a large volume of data necessitates the development of effective, efficient, economic and operational approaches. Automated approaches usually provide higher efficiency and thus more economic solutions, yet existing automated classification has been deemed ineffective because of the low classification accuracy achievable (typically below 65%) at global scale at 30 m resolution. As a result, an approach based on the integration of pixel- and object-based methods with knowledge (POK-based) has been developed. To handle the classification process of 10 land cover types, a split-and-merge strategy was employed, i.e. firstly each class identified in a prioritized sequence and then results are merged together. For the identification of each class, a robust integration of pixel- and object-based classification was developed. To improve the quality of the classification results, a knowledge-based interactive verification procedure was developed with the support of web service technology. The performance of the POK-based approach was tested using eight selected areas with differing landscapes from five different continents. An overall classification accuracy of over 80% was achieved. This indicates that the developed POK-based approach is effective and feasible for operational GLC mapping at 30 m resolution.

© 2014 The Authors. Published by Elsevier B.V. on behalf of International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/3.0/>).

1. Introduction

Information regarding land cover and its change over time is essential for a variety of societal needs, ranging from natural resources management, environmental studies, urban planning to sustainable development (Foley et al., 2005; Roger and Pielke, 2005; Running, 2008; Grimm et al., 2008; Zell et al., 2012; Sterling et al., 2013). Remote sensing has long been recognized as an effective tool for broad-scale land cover mapping (e.g., Carneggie and Lauer, 1966; Kushwaha, 1990; Townshend et al., 1991; Cihlar,

2000; Rogana and Chen, 2004; Hansen et al., 2013). As a result, a number of land cover datasets at a global scale have been developed with resolution ranging from 300 m to 1 km, using coarse resolution satellite imagery such as AVHRR, MODIS and MERIS (Hansen et al., 2000; Loveland et al., 2000; Friedl et al., 2002, 2010; Bartholomé and Belward, 2005; Bontemps et al., 2010). Although these GLC data products have been widely used, their quality is far from satisfactory for many applications (Coppin et al., 2004; Croke et al., 2004; He et al., 2006; Herold et al., 2008; Gong, 2009; Goward et al., 2011; Verburg et al., 2011). Various researchers (e.g., Iwao et al., 2006; Gong, 2009; Fritz et al., 2010) have highlighted the shortfalls of these datasets, e.g. the considerably low accuracies and low-level

* Corresponding author.

agreement amongst themselves. Consequently, the demand for new GLC products with improved spatial resolution and accuracy has been increasingly recognized by the remote sensing community e.g. the Group on Earth Observations (GEO) and the International Society for Photogrammetry and Remote Sensing (ISPRS) (Zell et al., 2012; Giri et al., 2013).

With the long-term archive and free availability of Landsat and similar image data, the development of GLC data products at 30 m resolution has become feasible. Such product have been considered a superior option for the next generation of GLC maps, since most significant human activities on the land system can be captured at this scale (Wulder et al., 2008; Giri et al., 2013). During the past two decades, the extraction of land cover information from Landsat-like imagery has been intensively studied, and a variety of automated and semi-automated methods/algorithms have been developed (e.g., Gong and Howarth, 1992; Ban, 2003; Coppin et al., 2004; Lu et al., 2004; Lu and Weng, 2007; Aitkenhead and Aalders, 2011; Chen et al., 2012; Huang and Jia, 2012; Ban and Jacob, 2013; Chen et al., 2013a,b). These have been applied to a number of national and regional land cover mapping projects using Landsat imagery (Liu et al., 1999; Xian et al., 2009; Johnson and Mueller, 2010; Hansen and Loveland, 2012). For instance, a set of 30 m land cover data with 13 different classes was produced by MacDonald Dettwiler and Associates Information Systems LLC (MDA, 2014), which covers the USA and a large proportion of Africa and Asia.

Land cover mapping with 30 m resolution at a global scale is much more complex than national or regional scale due to a number of factors, including the availability of good-quality imagery covering the land surface of the entire Earth (about 150 million km²) and the complex spectral and textual characterization of global landscapes. This makes the development of reliable 30 m GLC data products a very difficult task, as it requires a substantive level of technical innovation, as well as human and financial resources. This could be the reason why so far only global datasets with limited classes at this resolution have been reported (e.g. a global forest dataset at 30 m resolution was produced by Townshend et al. (2012) and Hansen et al. (2013)). Of course, forest data is very important for various applications, however a 30 m GLC product with a more comprehensive set of land cover types is desirable for wider studies.

In 2010 China launched a GLC mapping project, the aim of which was to produce a 30 m GLC data product (GlobeLand30) with 10 classes for years 2000 and 2010 within a four year period (Chen et al., 2011a). It was defined as an operational mapping project, with the production of reliable datasets as its clear objective. To achieve such an objective, it was necessary to utilize automated classification routines as much as possible. Therefore, the first attempt was to conduct an experimental evaluation on the usability of existing automated classification techniques (Gong et al., 2013). Four classifiers, i.e., Maximum Likelihood Classifier (MLC), J4.8 Decision Tree Classifier (DT), Random Forest Classifier (RF), and Support Vector Machine (SVM) were tested with more than 8000 images captured during the year 2000. It was found that the highest overall classification accuracy (OCA) was produced by SVM, at 64.9% (Gong et al., 2013). Such low accuracy is possibly attributable to significant spectral confusion among different land cover types. This essentially meant that it was not feasible to make use of fully automated classification techniques for such an operational project. A variety of mapping strategies and classification approaches were therefore investigated for deriving reliable 30 m GLC datasets (Chen et al., 2012, 2014a; Hu et al., 2014; Liao et al., 2014; Tang et al., 2014).

This paper presents a pixel-object-knowledge-based (POK-based) classification approach, the primary methodology used to produce China's 30 m GLC data product (GlobeLand30). Section 2

examines the two most critical issues for operational 30 m GLC mapping. Section 3 introduces the key concepts of the POK approach proposed for an operational mapping strategy. Section 4 describes how pixel-classifiers and object-based identification were integrated for per-class classification. Section 5 reports the use of knowledge for verifying and interactively improving mapping results. Section 6 presents the experimental results for selected areas and the accuracy assessment of the final GlobeLand30 product, as well as a comparison with regard to other similar 30 m data products. Section 7 provides discussion and conclusions.

2. Critical issues for operational 30 m GLC mapping

Experimental and operational classifications are two different approaches for large area land cover mapping and monitoring (Defries and Townshend, 1999; Hansen and Loveland, 2012). The former concentrates on the development and performance testing of novel algorithms and models, the latter focuses on the development and delivery of reliable data products within a pre-defined time schedule. From the point of view of operational GLC mapping at 30 m resolution, the selection and/or development of appropriate classifiers suitable for spectral and textual characterization of complex landscapes, as well as the quality of the resulting data products are two critical issues that need to be addressed.

2.1. Appropriate classifiers suitable for characterization of complex landscapes

Taking into consideration the massive task of operational GLC mapping and the existing land cover classification systems, China's GLC mapping project adopted a classification scheme consisting of 10 first-level classes, namely water bodies, wetland, artificial surfaces, cultivated land, permanent snow/ice, forest, shrubland, grassland, bareland and tundra. Table 1 presents these first-level classes and examples of their typical appearance on Landsat imagery collected from different sites around the world. On one hand, 30 m remote sensing imagery allows the observation of human impact on the Earth surface through scrutiny of shapes, sizes and patterns. Braided rivers and lakes, small villages and airports, irrigated round cultivated land, large-scale machinery agricultural patches can all be observed among others. On the other hand, it can easily be observed that there is high spectral heterogeneity within a single land cover class and significant spectral confusion among different classes. For instance, clear water from a reservoir, a river with a high sediment content, and a eutrophic lake may have very diverse spectral reflectance. Conversely, some subclasses of artificial surface, bareland and cropland have very similar spectral responses. The cultivated land class contains irrigated farmlands, paddy fields, green houses cultivated land, artificial tame pastures, economic cultivated land (such as grape, coffee, and palm), and abandoned arable lands. They may not have a unique spectral signature but have similarities with other land cover classes. This high spectral variation makes the per-pixel classification much more difficult than might otherwise be anticipated (Lu and Weng, 2007). For reliable operational mapping, it is impossible to select an appropriate single per-pixel classifier and a suitable set of variables for the entire globe. Therefore, new methods and approaches need to be adopted or developed to deal with the complex classification issues.

Image segmentation and object-based classification techniques have been developed to derive these structural elements by grouping pixels that have a relatively uniform spectral response into objects and identify the real world land cover features (Ban et al., 2010; Blaschke, 2010; Malinverni et al., 2011). Errors associated

Table 1
The 10 land cover classes and examples of their appearance on 30 m imagery.

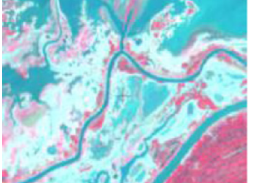




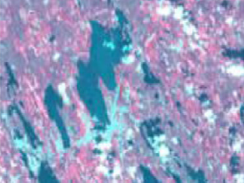


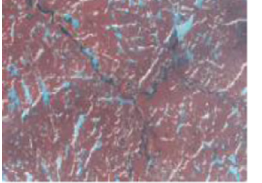
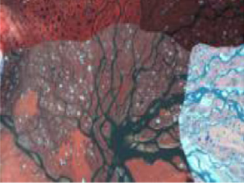
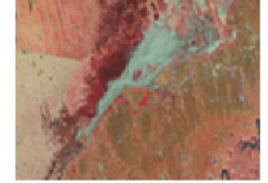




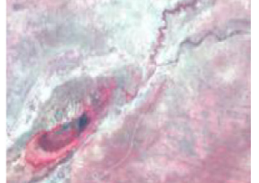



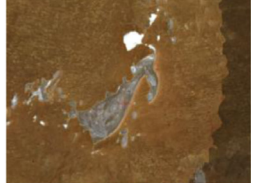

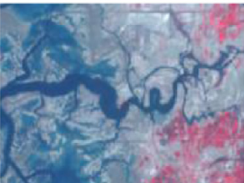


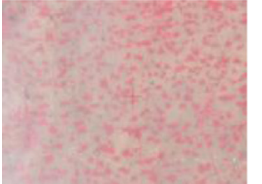



Land cover type	Contents and typical images on the 30m imagery				
Water bodies					
	(1)Amazon River, Brazil	(2)Lake Liangzi, China	(3)Reservoir Dalangdian, China	(4)Reservoir RiZhao, China	
					
	(5)Guangzhou, China	(6)Eastern Finland, Finland	(7)Nedalen, Sweden	(8)Nile River, Uganda	
					
	(9)Canada	(10)Lena River, Russia	(11)South Australia, Australia	(12)Sankey fishpond, China	
	Wetland				
		(1)Queensland, Australia	(2)Lake Tonle Sap, Cambodia	(3)Segou, Mali	(4)Okavango Delta, Botswana
					
		(5)Buenos Aires, Argentina	(6)Kherson Oblast, Ukraine	(7)Congo River, Congo (DRC)	(8)Coondambo, Australia
					
		(9)Sundarbans, Bangladesh	(10)Indus Delta, Pakistan	(11)Xinjiang, China	(12)Oruro, Bolivia
Artificial surfaces					
		(1)Anhui, China	(2)Cairo, Egypt	(3)Mo Bodhi Region, Mali	(4)Fremantle, Australia
		<i>(continued on next page)</i>			

Table 1 (continued)

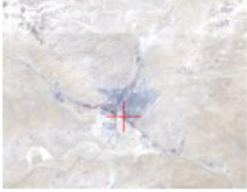

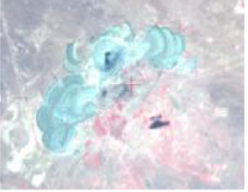


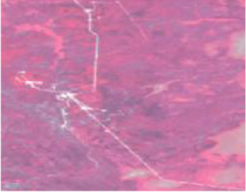
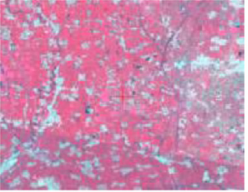

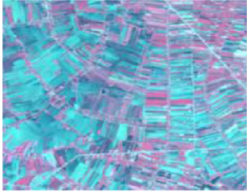
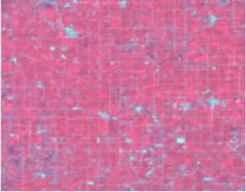
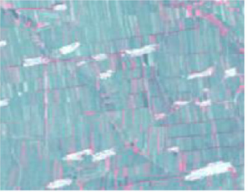


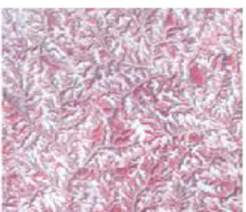




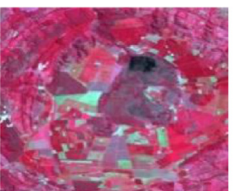
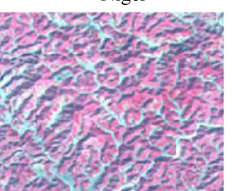


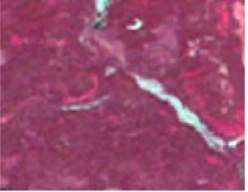
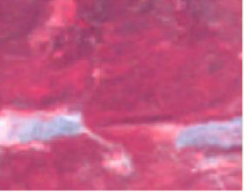
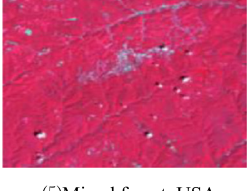
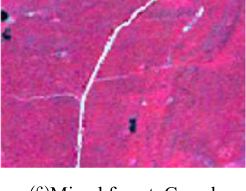
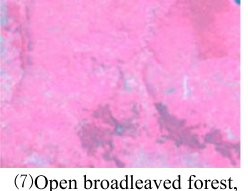

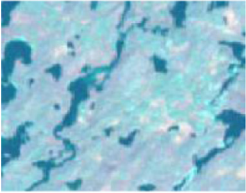


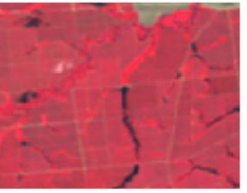

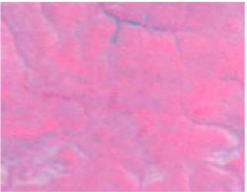
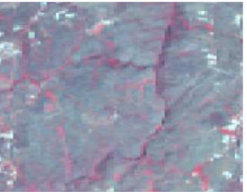
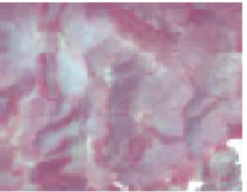

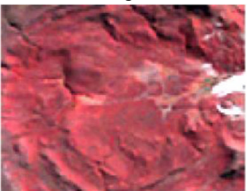
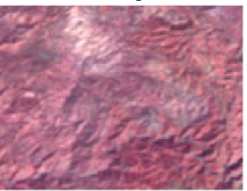
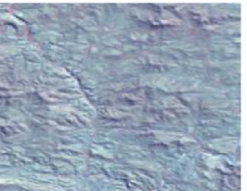
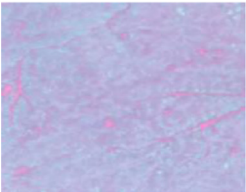




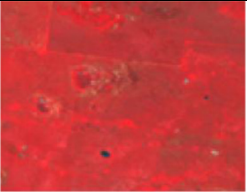
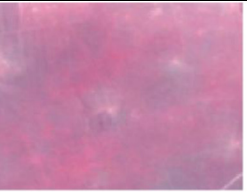





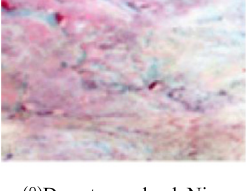
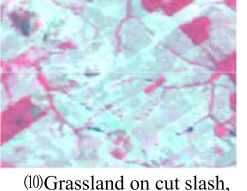






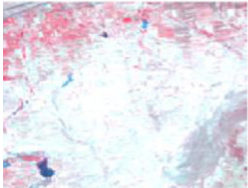

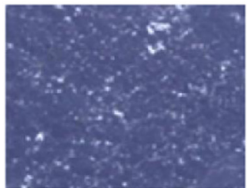



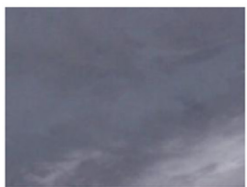

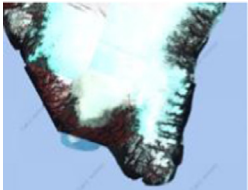
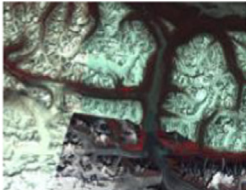
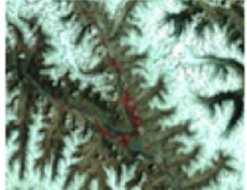

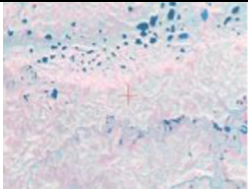



					
	(5)Northern of Saudi Arabia	(6)Sendai, Japan	(7)Cathode copper mine, Kazakhstan	(8)Boren ,America	
					
	(9)Port, Singapore	(10)Road, Russia	(11)Jixian County, China	(12)Cairo International Airport, Egypt	
Cultivated land					
	(1)Paddy, Thailand	(2)Large-scale machinery farmland, USA	(3)Harvested cropland, China	(4)Sloping cropland, Peru	
					
	(5)Irrigated round cropland, Egypt	(6)Cropland in Loess Plateau, China	(7)Terraces, Pakistan	(8)Cropland around Sahara, Niger	
					
	(9)Artificial tame, Russia	(10)Greenhouse,China	(11)Coffee Plantation, Brazil	(12)Paddy in mountain area, China	
	Forest				
		(1)Broadleaved forest, Brazil	(2)Broadleaved forest, Canada	(3)Coniferous forest, Canada	(4)Coniferous forest, Algeria
					
(5)Mixed forest, USA		(6)Mixed forest, Canada	(7)Open broadleaved forest, Congo(DRC)	(8)Open Coniferous forest, USA	

Table 1 (continued)

					
	(9) Sparse coniferous forest, Canada	(10) Young broadleaved forest, USA	(11) Olive, Tunisia	(12) Artificial forest, Indonesia	
Shrubland					
	(1) Closed shrubland, Argentina	(2) Tropical shrubland, Madagascar	(3) Sclerophyllous shrubland, Portugal	(4) Hilly shrubland, Turkey	
					
	(5) Alpine shrubland, China	(6) Mountainous shrubland, Argentina	(7) Mountainous low-shrubland, Mexico	(8) Sparse mountainous shrubland, South Africa	
					
	(9) Secondary shrubland, Ghana	(10) Low shrubland, USA	(11) Deserts and xeric shrubland, Kazakhstan	(12) Xeric shrubland, USA	
	Grassland				
		(1) Alpine meadow, China	(2) Pasture tame, Australia	(3) Meadow steppe, China	(4) Typical grassland, Kazakhstan
					
		(5) Mountainous grassland, USA	(6) Pampas steppe, Uruguay	(7) Savanna, Botswana	(8) Shrub grassland, Argentina
					
		(9) Desert grassland, Niger	(10) Grassland on cut slash, Brazil	(11) Grassland on fire slash, Burkina Faso	(12) Green of golf course, Zimbabwe

(continued on next page)

Table 1 (continued)

Bareland						
	(1)Dune desert, Saudi Arabia	(2)Stony desert, Sudan	(3)Gobi desert, Mongolia	(4)Sandy land, Mexico		
						
	(5)Saline-alkaline land, China	(6)Bare rock, China	(7)Gravel desert, Saudi Arabia	(8)Piedmont alluvial fan, Chile		
						
	(9)Biological crusts, Algeria	(10)Bare soil, Tunisia	(11)Dry salt lake bed, Tunisia	(12)Volcanic ash land, Reunion		
	Permanent snow and ice					
		(1)Greenland, Denmark	(2)Alaska Range, America	(3)Masherbrum, Kashmir	(4)Bayinguoleng, China	
		Tundra				
			(1) Wet tundra, Russia	(2) Grass tundra, USA	(3) Bare tundra, Canada	(4)Shrub tundra, Norway

with individual pixel misidentification may be avoided through this object-based classification approach (Aplin and Smith, 2011; Myint et al., 2011). It is natural to combine the spectral and structural information for solving the problems related to the characterization of complex landscapes (Ryherd and Woodcock, 1996; Trias-Sanz et al., 2008; Hussain et al., 2013). However, the determination of suitable segmentation parameters and incorporation of domain knowledge into the object identification process is extremely difficult, especially when the mapping area has a complex landscape with multi-scale patterns (Smith, 2013).

2.2. Assurance of data product quality

The reality consistency and logical consistency are used to evaluate the quality of geo-spatial data products (Heipke et al., 2008; Brisaboa et al., 2014; Chen et al., 2014b). For a land cover data product, the former refers to the thematic accuracy and up-to-dateness in comparison to the real world. The latter is related to the congruency of land cover features (or objects) within the same dataset, and between sequential datasets (Verburg et al., 2011; Robertson and King, 2011; Hansen and Loveland, 2012). Due to the errors introduced by image data manipulation and classification, it is common

to find inconsistency or imperfections in land cover datasets. Therefore the verification of preliminary classification results is an important task to ensure a high level of data quality.

A number of factors impact on the reality and logical consistency, such as the semantic interpretation of the adopted classification scheme, handling of the minimum mapping units, as well as time differences between the imagery used and the mapping time (Lu and Weng, 2007; Verburg et al., 2011; Costa et al., 2014). For instance, different understandings of the definition for wetland may lead to dissimilar results of interpretation from the same 30 m imagery and with the same mapping methodology (Comber et al., 2004), because the wetland class consists of several sub-types (lake swamp, river flooding wetlands, sea marsh, shrub/forest wetlands, mangrove forest, tidal flats/salt marshes). In addition, different individual interpretations of multi-epoch imagery might cause false changes. The reduction of omission and commission errors in land cover data products requires the formulation of data quality knowledge rules and the development of standardized verification procedures (Lu and Weng, 2007; Costa et al., 2014). However, the logical and temporal consistency of land cover category data is not as evident or explicit as the topographic features (Gerke et al., 2004; Brisaboa et al., 2014; Chen et al., 2007). One

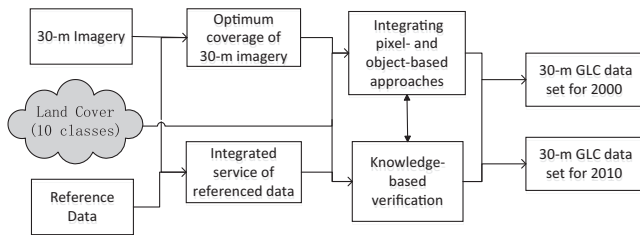


Fig. 1. Operational mapping strategy.

possible solution is to develop advanced verification tools and interactive mapping procedures by integrating land cover knowledge and all available ancillary data.

3. A POK approach for operational mapping

To address the critical issues outlined in Section 2, a POK-based approach was proposed. Its key elements are the integration of pixel- and object-based classification approaches, optimal coverage of 30 m imagery for two baseline years, web-service oriented integration of reference data, and knowledge-based verification for assuring the consistency and accuracy of 30 m GLC data products. An overview of the operational mapping strategy adopted is presented in Fig. 1.

3.1. Integrating pixel- and object-based classification techniques

As it is difficult to select or invent a classifier that is suitable for complex land cover mapping over large areas, the combination of different classification techniques has been investigated (e.g., Aitkenhead and Aalders, 2011; Hansen and Loveland, 2012). Specifically, the integration of pixel-based and the object-based classifiers for large area land cover mapping has been explored by several authors (e.g. Myint et al., 2011; Malinverni et al., 2011; Robertson and King, 2011; Smith, 2013; Costa et al., 2014). In this project, the object-based technique was used to determine the spatial extent of land features with their structural/contextual information to form land objects. For any given land object, pixel-based classifiers were used to derive variables (such as vegetation indices) and to identify the attribute value with the help of available reference data and expert knowledge.

Hierarchical classification strategies have also been tested by several researchers with a series of per-class classifiers to minimize the effect of spectral confusion among different land cover classes (Frazier and Page, 2000; Sulla-Menashe et al., 2011; Smith, 2013). Here it was proposed to decompose the complex GLC mapping into a series of simpler per-class classifications. Different pixel-based variables and segmentation approaches were selected and utilized for characterizing different land cover classes. The water bodies was extracted first and the corresponding pixels were masked out before the next processing stage. Wetland was characterized next, followed by permanent snow/ice, artificial surfaces, cultivated land, forest, shrubland, grassland, bareland and tundra.

3.2. Optimal coverage of 30 m imagery for two baseline years

Landsat TM/ETM+ images were selected as the primary data source for the two baseline years of 2000 and 2010. Imagery from the Chinese Environmental and Disaster satellite (HJ-1) were also used as supplementary data for the year 2010, as they have similar characteristics to the Landsat TM/ETM+ sensor in terms of spectral band set and spatial resolution (Hu et al., 2014).

The 30 m remotely sensed images were carefully selected to ensure optimal global coverage and minimal cloud contamination and such that they were captured within the local vegetation growing season. In addition, various image data useful to classification and validation were collected, including MODIS-NDVI time series data for extraction of seasonal information of vegetation (Verbesselt et al., 2010). Landsat TM/ETM+ provided by USGS were satisfactorily geo-referenced, meaning no further geometric registration was necessary. However, the HJ-1 imagery suffers from nonlinear distortion because of its wide-scene imaging capability (Zhang et al., 2012; Tang et al., 2014; Hu et al., 2014). Therefore, a specific geo-registration model that integrates the collinearity equations and Lagrange interpolation was developed for geometric processing of HJ-1 imagery. In total, 2640 HJ-1 scenes were orthorectified with a registration accuracy of less than two pixels, covering approximately 60% of the global land surface.

Radiometric correction was performed for both Landsat and HJ-1 imagery, including both atmospheric and topographic corrections. The atmospheric correction model of MODerate resolution atmospheric TRANsmission (MODTRAN) was used for Landsat imagery (Berk et al., 1999). A relative radiometric correction was adopted for HJ-1 imagery by selecting corresponding atmospherically corrected Landsat imagery that covered the same area but was acquired in different years as reference. For imagery covering mountainous areas with rough terrain, a new approach was developed to restore the radiometric information in areas of shadow cast by mountains by using a “continuum removal” (CR) spectral processing technique without the aid of a DEM (Zhou et al., 2014). The CR-based approach makes full use of the spectral information derived from both the shaded pixels and their neighboring non-shaded pixels of the same land cover type.

Thick cloud cover and Landsat ETM+ Scan Line Corrector (SLC)-off failure occasionally resulted in poor quality or missing data on Landsat TM/ETM+ images. The neighborhood similar pixel interpolator (NSPI) approach was employed to remove thick clouds (Zhu et al., 2012) and filled gaps in SLC-off images (Chen et al., 2011b). Based on the assumption that the same-class neighboring pixels around the missing pixels caused by thick clouds or SLC-off failure have similar spectral characteristics, and that these neighboring and missing pixels exhibit similar patterns of spectral differences at different dates, Neighborhood Similar Pixel based model (NSPI) interpolates the missing pixels by using information from an ancillary imagery with good quality but acquired at a different date. The results of case studies indicate that NSPI can restore the value of missing pixels very accurately, and that it works well in heterogeneous regions. In addition, it can work well even if there is a relatively long time interval or significant spectral changes between the input and target images. The filled images appear spatially continuous without any obvious striping patterns (Chen et al., 2011b; Zhu et al., 2012; Huang, 2012).

3.3. Web-based reference data integration

Ancillary data have proven effective in improving classification accuracy, such as the identification of land cover with higher spectral variation and distinctions between different land cover classes with similar spectral responses (Chen, 1984; Ehlers et al., 1989; Lu and Weng, 2007). At the global scale, there are a variety of reference data sets that can be used to support 30 m GLC mapping, such as existing GLC maps at coarser resolution, 30 m or higher resolution regional land cover data, global DEM data (SRTM and ASTER DEM), global 1:1 million topographic data (Hayakawa et al., 2008), and ecological zones (Olson et al., 2001). Online-distributed geospatial data assets and services (such as Google map, Map World, and OpenStreetMap), as well as land cover related services (such as Geo-Wiki) also provide valuable external and interoperable

ancillary sources of information (Fritz et al., 2012; Yu and Gong, 2012).

Ancillary data are less uniform than remotely sensed data, varying in format, accuracy and spatial resolution. In order to facilitate the use of such data and their incorporation into classification and verification processes, a bespoke web-based information service system was designed, developed and implemented by the Chinese GLC team. The system was used to integrate all the primary image data, reference data, ancillary data, and preliminary and final classification results (Han et al., 2015). Ancillary data were processed and published as web services according to OGC standards. The web-based system was connected to the image analysis systems, with data exchange facilitated by a shared user interface. Online tools were developed to support data browsing, information retrieval, comparison, geo-tagging, and verification. A special tool for change markup and reporting was developed following a publish/subscribe model for multi-tier applications.

3.4. Knowledge-based interactive verification

The verification of geo-spatial data is an important process to identify inconsistency and to remove, or minimize, errors in the datasets (Comber et al., 2004; Brisaboa et al., 2014). Knowledge-based verification has been explored for topographic mapping (Mills and Newton, 1996; Gerke et al., 2004; Chen et al., 2007) and land cover mapping (Comber et al., 2004; Verburg et al., 2011). A knowledge framework was proposed by Gao et al. (2004) for the generalization of land cover maps with nature-based, culture-based and application-specific knowledge. However, a priori knowledge of land cover and its change at a global scale is not available (Lu and Weng, 2007; Smith, 2013). In order to facilitate the logical consistency checking, knowledge about the geographical distribution of land cover was summarized and used to guide the data product verification. With the support of a web-based system, classification errors (commissions and omissions) on land cover and its change were checked with a variety of ancillary data and prior-knowledge. An allowable number of errors was defined for each land cover class on the basis of the minimum mapping units for a standardized quality controlling. This was followed by an interactive improvement to remove identified errors. A more detailed explanation of the knowledge-based verification approach adopted is provided in Section 5.

3.5. The proposed POK approach

With the above considerations, a pixel-object-knowledge (POK) based classification approach was developed, as shown in Fig. 2. With the POK approach, the classification process of all 10 land cover types is handled with a split-and-merge strategy, i.e. firstly each class is determined in a prioritized sequence and then the results are merged together. The first step is called “per-class classification”.

In per-class classification, pixel-based classifiers and object-based identification approaches are integrated to determine the spatial extent and category of land cover features in each class. Both the spectral and multi-temporal signatures of the 30 m imagery are used for deriving variables. A knowledge-based interactive verification is then carried out to check and improve the reality and logical consistency of the data products with the support of web service technology.

For each class, the classification results of individual scenes were integrated into a map sheet of dimensions 6° in longitude and 5° in latitude. In total, 847 map sheets were produced. Inconsistency and conflicts within the individual classification results and at strata boundaries were checked and corrected by operators. All these per-class classifications were aggregated afterwards to produce comprehensive datasets for the years 2000 and 2010.

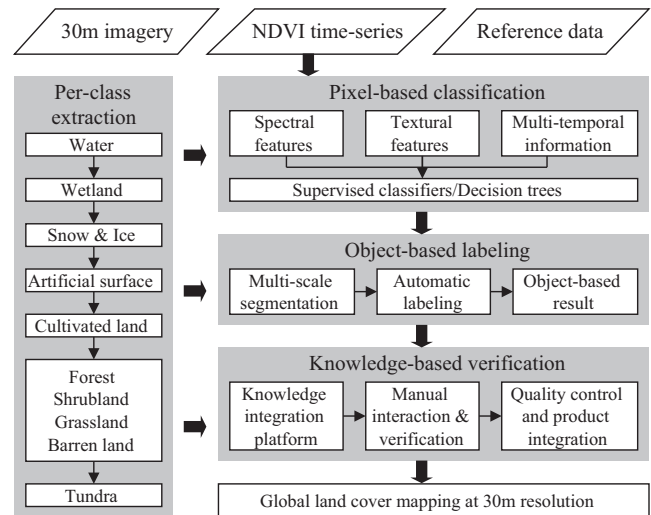


Fig. 2. The developed POK classification approach.

4. Hybrid pixel- and object-based classification

Pixel-based classification may generate a large number of misclassified pixels (the “salt-and-pepper effect”) due to the spectral confusion between land cover types and spectral diversity within the same land cover type. Therefore, object-based identification was adopted to integrate the pixel-based classification results with segmented objects generated using eCognition (v8) software. In this process, scale parameters ranging from 10 to 50 (with interval of 5) were used to obtain multiple scale segmentation results because landscapes of different land cover types vary significantly in multiple scenes. Based on pixel-based per-class extraction results and multiple segmentation layers, the majority criterion (Lu and Weng, 2007; Ok and Akyurek, 2012; Costa et al., 2014) and auxiliary information (such as DEM, slope, and reference GLC products) were applied as a decision rule to automatically or manually label the objects. Table 2 provides an overview of the pixel-based classifiers and object-based identification methods for each land cover type.

4.1. Water bodies

In general, water bodies exhibit a distinct spectral feature in Landsat TM/ETM+ imagery compared to other land cover types. A supervised Maximum Likelihood Classification (MLC) was employed based on the Normalized Differential Water Index (NDWI) and Wetness component of Tassel Cap transformation (TC-W). For the majority of mapped areas, the method worked well to ensure the accuracy and efficiency of water mapping (Liao et al., 2014). In some areas, water bodies, which included clear water, green water and turbid water, displayed spectral diversity. Clear water displayed lower reflectance in all bands, while turbid water had higher reflectance due to the blend of mud and sand, and green water caused by eutrophication displayed spectral features similar to green vegetation to some extent. A prior knowledge based decision trees classifier was adopted to deal with the complexity of land surface waters (Sun et al., 2012).

With the results of image segmentation, first the percentage of water pixels from the pixel-based classification results within each image object was calculated, and then a percentage threshold was defined, such as 20%, 15% and 10%, which was then used to compare with the percentage of water pixels within each image object. In this project, 10% was selected as the percentage threshold to

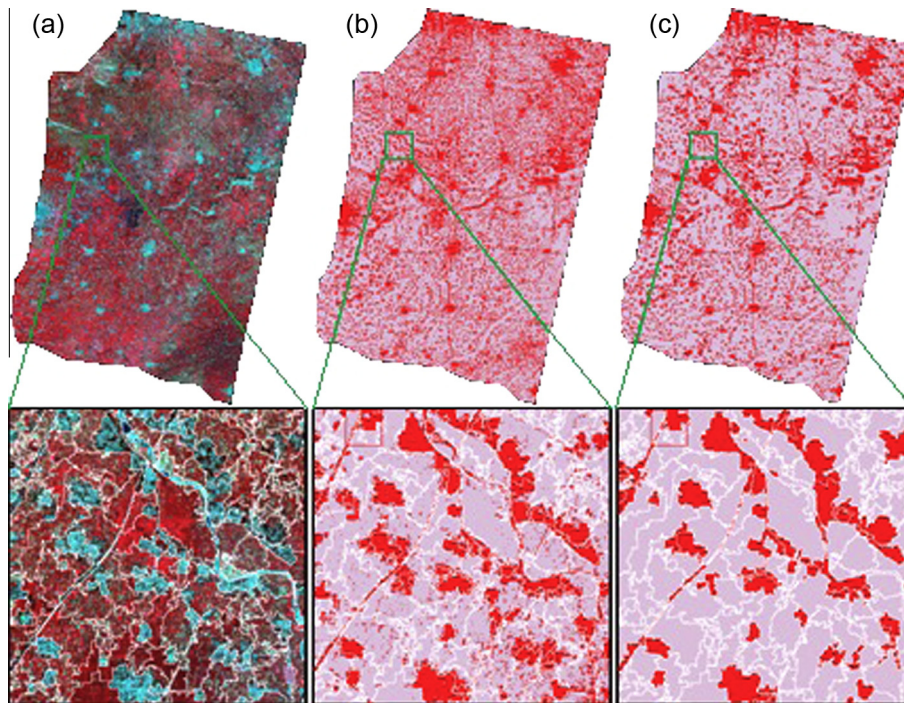


Fig. 3. Integration of pixel- and object-based classification of artificial surfaces: (a) Landsat TM image, (b) pixel-based classification result, and (c) object-based identification result. Red color denotes artificial surfaces, pink color is the other land cover types (background), and white polygons in the enlarged figures are the objects segmented based on the scale parameter of 20. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Per-class extraction with POK approach.

Land cover type	Pixel-based method	Object-based method	
	Features	Classifiers	
Water bodies	NDWI, TC-W, spectral bands, DEM and slope	MLC/SVM/DT	Percentage thresholding
Wetland	/	/	Expert interpretation
Permanent snow and ice	NDSI, spectral bands	Automated thresholding/MLC	/
Artificial surfaces	Spectral bands, texture, NDBI	SVM	Percentage thresholding
Cultivated land	Spectra bands, NDVI time-series and texture	MLC	Percentage thresholding
Forest	Spectral bands, NDVI time-series and priori probability	MLC	Expert interpretation
Shrubland			
Grassland			
Bareland			

label the object as “water” class after a comprehensive comparison (Sun et al., 2012; Liao et al., 2014).

4.2. Wetland

Wetland take multiple forms such as bogs, fens, meadows, marshes, peat land, swamp, floodplains, and mangroves. The spectral diversity within wetland is therefore significant. Different supervised classifiers were tested but the results were far from satisfactory (i.e. did not achieve acceptable accuracy). Since the successful extraction of wetlands depends on the contextual information and experts experiences, visual interpretation and inspection were used to accurately delineate wetland based on prior knowledge of wetland types, distribution, and temporal change patterns as well as the spectral characteristics.

4.3. Permanent snow and ice

Permanent snow and ice have unique spectral characteristics distinguishing it from other land cover types. An automated thresholding method was employed to identify snow and ice based on

the Normalized Difference Snow Index (NDSI) calculated from Landsat TM/ETM+ data (Yin et al., 2013). After careful comparison of several automated thresholding methods, Otsu’s method (Otsu, 1979) was selected as the operational approach as it selects an optimal threshold by the discriminant criterion to maximize the separability of the resultant classes in gray levels. With this method, a better accuracy was achieved in situations of various backgrounds and snow/ice fractions (Yin et al., 2013), allowing accurate and efficient snow and ice extraction.

4.4. Artificial surfaces

Artificial surfaces mainly consists of urban areas, roads, rural cottages and mines, which are primarily based on asphalts, concrete, sand and stone, bricks, glasses, and other materials. They can be divided into three sub-classes including high reflectance, low reflectance and vegetated type. Different from other land cover types, artificial surfaces usually holds specific textual patterns, which are helpful for further classification. Accordingly, a supervised pixel-based classifier (SVM) was first used to extract artificial surface using both spectral band information and a variance texture recommended by Lu and Weng (2007).

Table 3
Examples of Land cover and change knowledge used for data verification.

Type of knowledge	Examples	Implication for data verification
Nature-based knowledge	Land surface water and wetland reside in a relatively flat or low relief areas Arid/semi-arid zones have less rivers and lakes than an area with dense water networks Vegetation classes (forest, grass, tundra) might have different geographical distribution in distinct eco regions	<ul style="list-style-type: none"> • Calculate slope from DEM for filtering those area greater than a pre-defined threshold; • Identify areas covered by mountain shadows • Define different minimal allowable threshold about water omission or commission errors for these areas • Subdivide the whole globe using eco regions defined World Wildlife Fund (WWF) for minimizing confusion of discrete vegetation classes with similar spectral response
Culture-based knowledge	Woods are cut in forest areas and replanted later for commercial purpose Artificial surfaces areas may have a congruency with place names Land cover related socio-economic census data or inventory statistics serve as prior knowledge for the verification	<ul style="list-style-type: none"> • Identify those traces of cut wood which might be interpreted as farm lands (especially in Amazon area) • Use geographic names to locate those places in arid and semi-arid area where smaller towns are not easily identified from 30 m imagery • Compare the statistics resulted from 30-m GLC data with the available census or inventory statistics data for identifying gross and systematic error
Temporal constraints knowledge	Artificial surfaces will continue to exist once built up in most cases Land surface water has a great fluctuation Cultivated land with crops or after harvest have different spectra and may cause pseudo change	<ul style="list-style-type: none"> • Compute 2000–2010 artificial surfaces difference and check those areas with significant decrease • Correct the water area of different years with MODIS time-series data • Using MODIS time-series data and 30 m texture information to inspect fallow and harvested cultivated land

After the pixel-based classification, segmentation was performed at multiple scales, and areas of artificial surfaces were identified if the proportion of artificial surfaces pixels within an object was larger than a predefined threshold (i.e. 50%). Finally, the classification results of artificial surfaces at different scales were combined and verified by manual editing based on visual comparison with high resolution images (e.g. Google Earth) and other reference data. Fig. 3 gives an example of extraction of artificial surfaces by integrating pixel- and object-based classification methods.

4.5. Cultivated land

Cultivated land displays complicated spectral characteristics at the global scale. The spectra of cultivated land covered by plants are identical with those of natural vegetation, while the spectra of cultivated land without plants (fallow) are almost the same as those of soil. The most common spectra of cultivated land in growing season are the mixture of vegetation and soil. Moreover, cultivated land generally has regular distribution patterns such as circles or rectangles as well as texture which is different from natural vegetation due to human management. However, the utilization of texture and spectral characteristics is very difficult because they are determined not only by the type and size of

cultivated land, but also by the phenology of the crops, which contains planting, growing, maturity and harvest. In this study, a new cultivated land extraction method was developed based on both crop phenology and regular distribution pattern of the cultivated land. In order to obtain the phenological information of the crops, a new method, known as NDVI Linear Mixing Growth Model (NDVI-LMGM), was proposed for generating high spatial-resolution NDVI time series. By using multi-temporal 250 m MODIS NDVI data and Landsat imagery (Rao et al., 2014), this method was able to cope with the limitation of coarse spatial resolution MODIS-NDVI products for 30 m resolution land cover mapping. After that, the 30 m NDVI time series data were used by a pixel-level supervised classifier (MLC) to identify potential cultivated land pixels.

The objects obtained by segmentation were overlaid onto potential cultivated land images. Only objects displaying regular man-made patterns such as circles or rectangles and for which the proportion of potential cultivated land pixels within the object was larger than a predefined threshold (i.e. 70%) were identified as cultivated land via visual interpretation. Fig. 4 gives an example of cultivated land extraction in Amazon area. Some cultivated land were mis-classified as traces of cutting woods. With the object segmentation results and the help of higher resolution images, they were identified and modified as cultivated land patches.

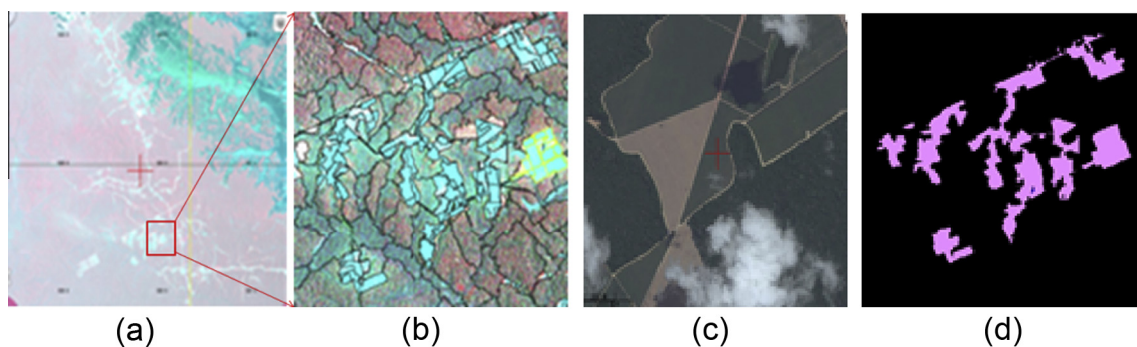


Fig. 4. Object-based cultivated land extraction: (a) TM image in Amazon Forest area where some cultivated land were miss-classified as traces of cutting woods; (b) segmentation results for cultivated land identification; (c) higher resolution image used for identifying the missed cultivated land patches; and (d) modified cultivated land patches.

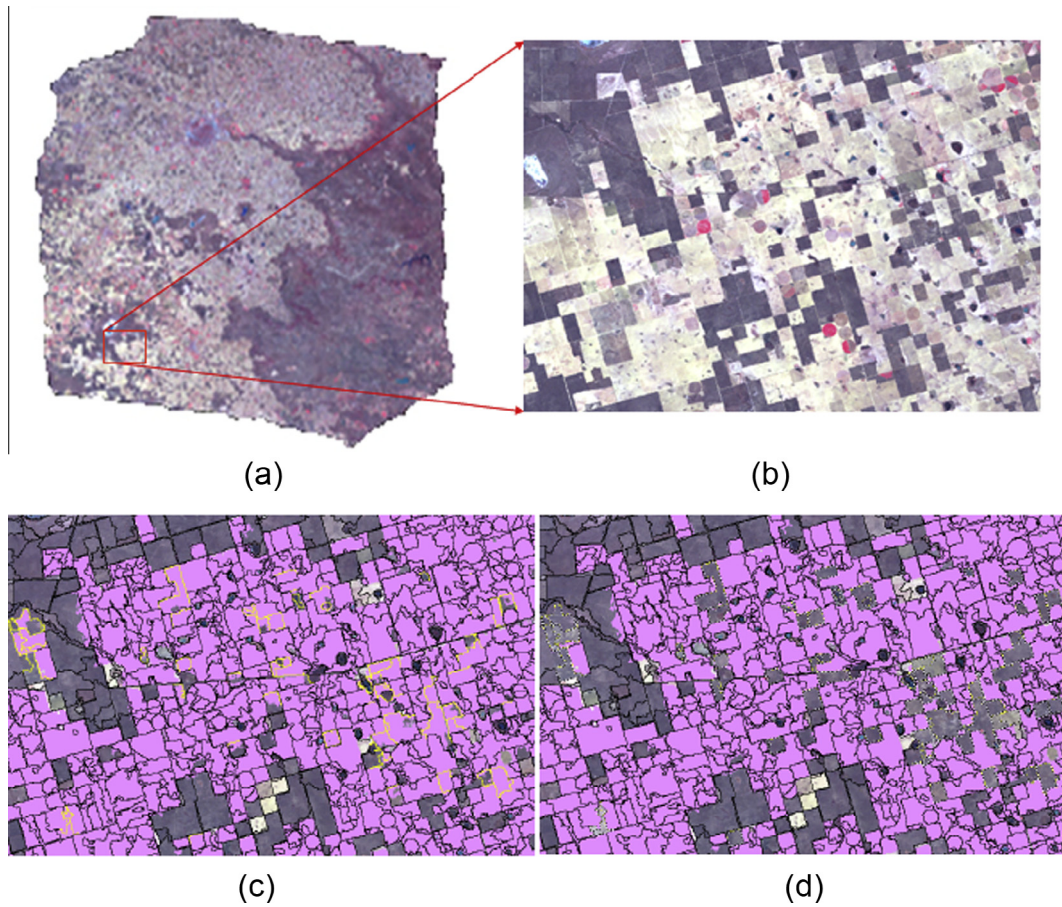


Fig. 5. Object-based grassland extraction (a) original image; (b) enlarged image; (c) example grassland (yellow) that was misclassified as cultivated land; and (d) correct classification results of grassland. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

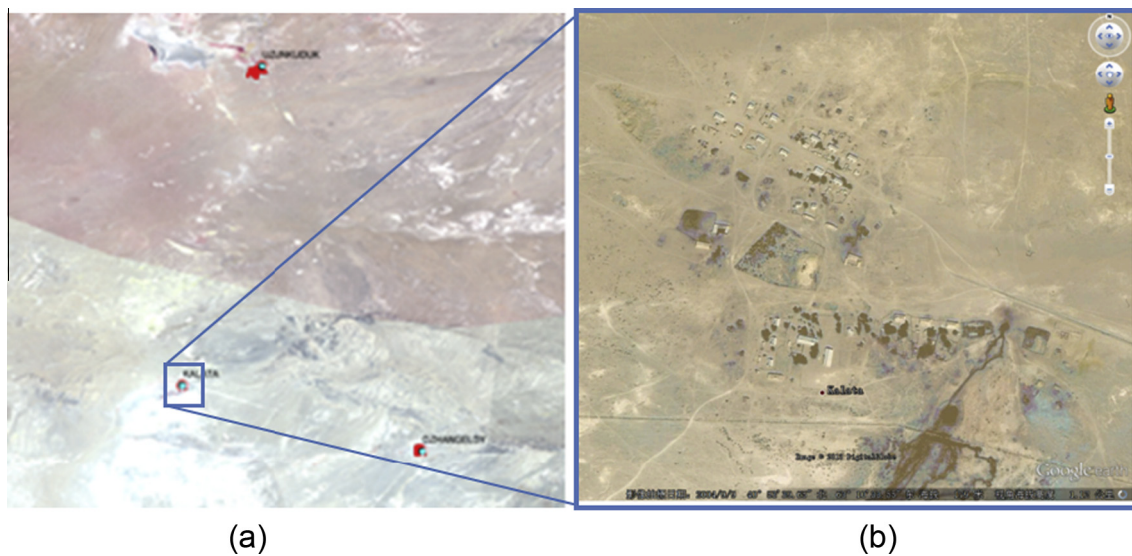


Fig. 6. Verifying missing artificial surface with geographic names and Google imagery (a) locating the missing artificial surface in an arid area by matching geographic names and (b) adding artificial surface with the help of Google imagery.

4.6. Forest, shrubland, grassland and bareland

After completing the classification of water, wetlands, snow and ice, artificial surfaces and cultivated land, it become relatively straightforward to classify forest, shrubland, grassland and

bareland because a large number of pixels with spectral similarity to these land cover types were masked out by the previous procedures. Fig. 5 illustrates the object-based extraction of grassland (yellow). Considering that the natural vegetation types and bareland are characterized by their own specific phenological patterns,

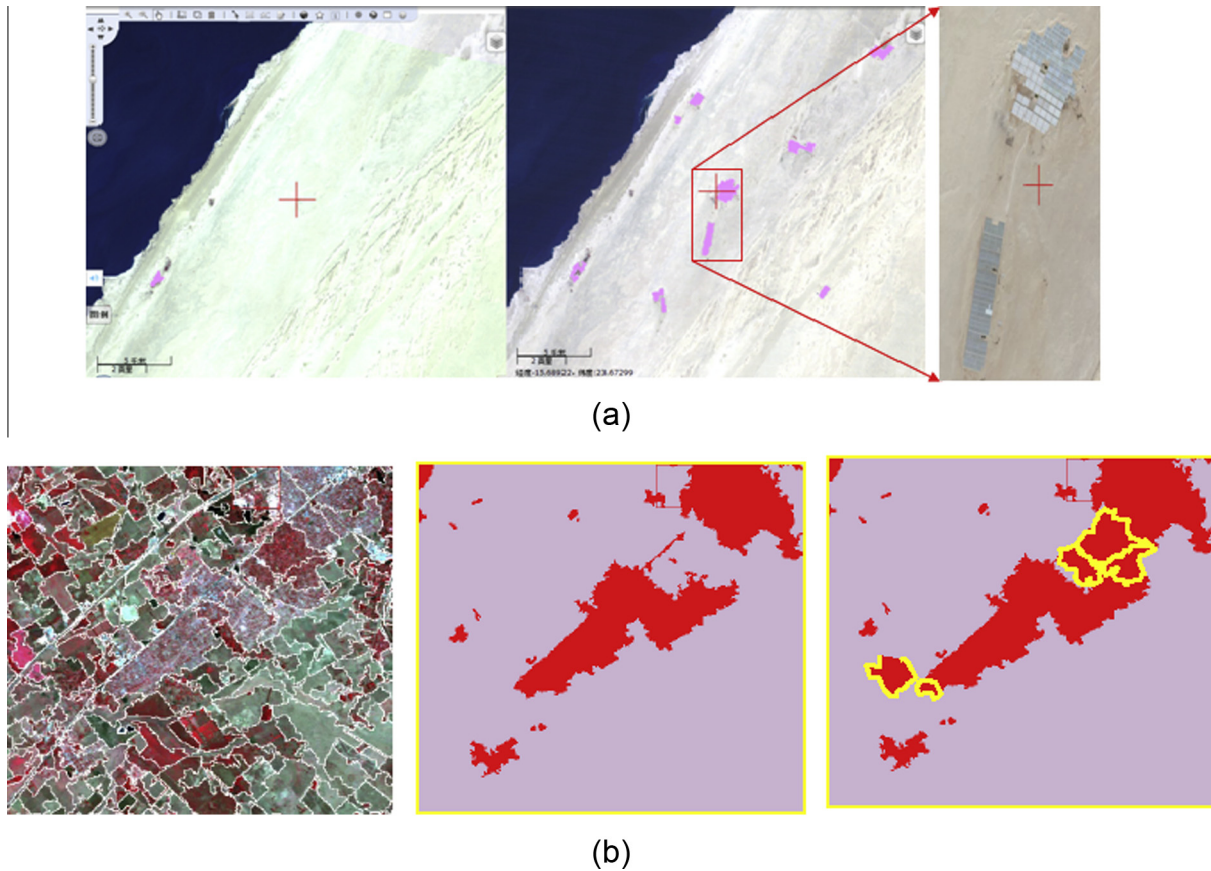


Fig. 7. Interactive verification with the support of the web-based system (a). Verifying the increase of farm land in the Western Sahara using higher resolution imagery (b). Artificial surfaces with high vegetation density was omitted and re-collected after verification.

six spectral bands of Landsat TM/ETM+ imagery, together with 23 NDVI bands per year (generated by the NDVI-LMGM method using multi-temporal 250 m MODIS NDVI data and Landsat imagery), were used as input features in the MLC classifier to classify forest, shrubland, grassland and bareland synchronously. Here, the MLC classifier was selected because it made better use of the existing land cover information through a priori probability. Consequently, before the classification, the a priori probability of each pixel in a Landsat TM/ETM+ image belonging to forest, shrubland, grassland and bareland was calculated according to the existing GLC datasets such as 300 m GlobCover 2009 (Bontemps et al., 2010) and 500 m BU-MODIS land-cover map (Friedl et al., 2010). Both land cover datasets were recorded according to the land cover types used in this research, and resampled to a corresponding 30 m pixel size. The priori probability of a pixel for each class was calculated by dividing the number of pixels of the class by the total number of pixels within a predefined window centered by the target pixel, and the averaged value of the a priori probability from the two global datasets was used as the input a priori probability for MLC classification. The a priori probability data made use of the latest land cover mapping experiences, and also helped to verify the training and validation samples.

5. Knowledge-based interactive verification

5.1. Land cover and change knowledge

Nature-based and culture-based knowledge of the geographical distribution of land cover are indispensable for land cover data verification. As two successive data products for years 2000 and 2010

were subject to verification, temporal consistency was added as another important knowledge.

5.1.1. Nature-based land cover knowledge

Nature-based knowledge refers to the interaction and relations among different elements of the natural environment (Gao et al., 2004). For instance, water and wetland reside in relatively flat or low relief areas. In areas of high latitude, surface water might be confused with mountain shadows due to solar altitude. With such knowledge, slope could be calculated from a global DEM and used to identify errors of omission or commission in surface water (Liao et al., 2014). Another typical example is that an arid or semi-arid zone has generally fewer rivers and lakes than a humid or semi-humid region with dense water networks. Different criteria on data quality control can be adopted for these two different types of area. The minimal allowable threshold for omission or commission errors of water objects (features) could be larger than that in a dense water network region. A few examples of nature-based knowledge and implications on the verification of land cover characterization are given in Table 3.

5.1.2. Culture-based land cover knowledge

Culture traditions, political policies and socio-economic events all have an impact on land cover distribution and changes. Such knowledge can also help in the verification of land cover datasets. For instance, people cut down trees and replant them in forest areas for commercial purposes. Without this knowledge, some clearings in Amazonian areas might be interpreted as farmlands. Artificial surfaces are considered as a proxy measure of human impact on the environment and is associated with centered

Table 4
Minimum mapping unit (MMU) and allowable minimum errors for each land cover type per scene.

Land cover type	MMU (pixels)	Allowable minimum error of omission or commission per-image scene (pixels or percentage)				
Water bodies	Lake, reservoir	3 × 3	Area	Dense water region	Arid and dense farmland region	General case
			$A \geq 4 \times 4$	0	0	0
			$3 \times 3 \leq A < 4 \times 4$	4	1	2
	River	3 × 1	Width ≥ 3	Must be continuous		
			$2 \leq A < 3$	Discontinuous is allowed		
Wetland	9 × 9	Sub-type	Grass			Wood
			Lake	2		3
			River	1		3
			Sea/estuary	1		2
Permanent snow and ice	3 × 3	Area	Mid-low latitude mountain regions with large terrain variation		General case	
			$A \geq 4 \times 4$	0	0	
			$3 \times 3 \leq A < 4 \times 4$	4	2	
Artificial surfaces	4 × 4	Area	General case	Difficult areas		
			$A \leq 8 \times 8$	2	4	
			$A > 8 \times 8$	0	0	
Cultivated land	6 × 6	Area	Large cultivated land in plain region	Arid/semi-arid, hilly mountain regions	General case	
			$A \geq 10 \times 10$	3%	3%	3%
			$6 \times 6 \leq A < 10 \times 10$	1	4	2
			$A < 6 \times 6$	No restrict		
Forest	8 × 8	Area	Continuous forest region	Arid/semi-arid, hilly or mountain/pastoral/farmland/urban regions	General case	
			$A \geq 12 \times 12$	3%	3%	3%
			$8 \times 8 \leq A < 12 \times 12$	1	4	2
			$A < 8 \times 8$	No restrict		
Shrubland	10 × 10	Area	Hilly valley/forest-shrub transition region	Arid/semi-arid region/mountain regions above forest line	General case	
			$A \geq 15 \times 15$	5%	5%	5%
			$10 \times 10 \leq A < 15 \times 15$	6	2	4
			$A < 10 \times 10$	No restrict		
Grassland	10 × 10	Area	Hilly/pastoral transition region/arid/semi-arid grassland/urban broken surface regions	Typical grassland/tropical savannah/mountain meadow regions	General case	
			$A \geq 15 \times 15$	5%	5%	5%
			$10 \times 10 \leq A < 15 \times 15$	6	2	4
			$A < 10 \times 10$	No restrict		
Bareland	6 × 6	Area	Hilly/plain/broken barren land region	Arid/semi-arid, hilly mountain regions	General case	
			$A \geq 10 \times 10$	3%	3%	3%
			$6 \times 6 \leq A < 10 \times 10$	1	4	2
			$A < 6 \times 6$	No restrict		

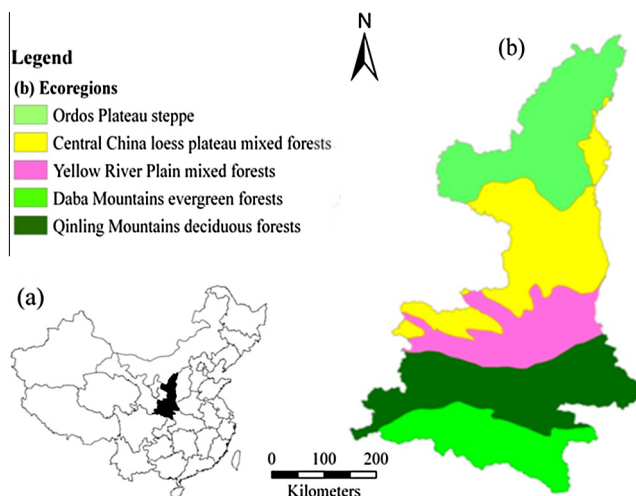


Fig. 8. Location, ecological regions of Shaanxi Province, China. (a) Location of Shaanxi Province (black) in mainland of China and (b) ecological regions of Shaanxi Province.

population (Elvidge et al., 2007; Sutton et al., 2009). For some smaller towns in arid and semi-arid areas, the spectral response is much closer to that of the background landscape and may not be easily identified from 30 m imagery. They could be located using geographic names from digital topographic maps at 1:1 million or large scale and then mapped with the help of online higher resolution imagery or other maps. Fig. 6 gives one example.

In addition, land cover related socio-economic census data or inventory statistics may also serve as prior knowledge for verification. For example, macro-inconsistency was conducted by comparing statistics from 30 m GLC mapping with thematic studies, such as global agricultural lands statistics (Ramankutty et al., 2008), inland water surface areas (Raymond et al., 2013), urban land cover areas in larger and smaller cities (Potere et al., 2009; Angel et al., 2011), and wetland statistics from the Wetland Handbook (Whigham, 2009).

5.1.3. Temporal-constraint knowledge

Differences between land cover datasets at regular time intervals can be spatially-explicit GLC change and are critical assets for global environmental change studies (Goldewijk and Ramankutty, 2004; Verburg et al., 2011; Sterling et al., 2013).

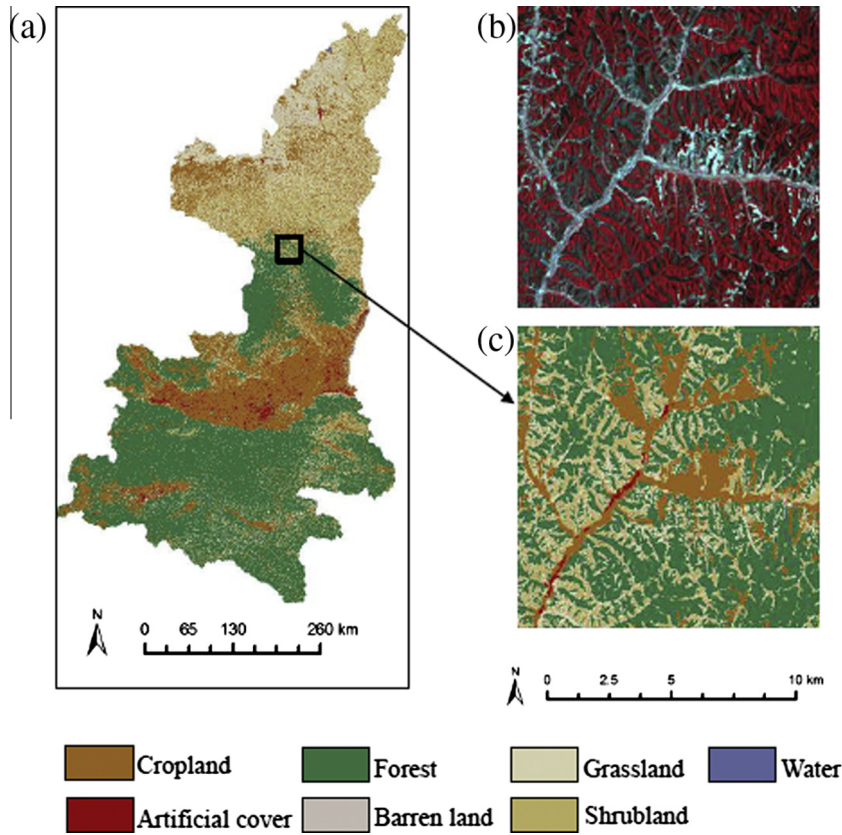


Fig. 9. (a) Land cover map of Shaanxi province in 2000 produced by POK, (b) detail comparison between Landsat imagery and (c) classification result.

Table 5
Error matrix of POK classification results of Shaanxi Province in base-year 2000.

Number of pixels	Reference data							Total	UA (%)
	Water	Artificial surface	Cultivated-land	Forest	Shrub-land	Grass-land	Barren land		
<i>Classification</i>									
Water bodies	3224	16	16	0	0	0	0	3256	99.0
Artificial surfaces	38	6103	419	0	2	6	0	6568	92.9
Cultivated land	285	1400	11530	98	0	184	51	13548	85.1
Forest	127	1	183	2181	2	98	0	2592	84.1
Shrubland	223	46	563	48	416	339	98	1733	24.0
Grassland	121	68	1409	33	149	1100	1113	3993	27.5
Bareland	30	18	141	0	13	31	2539	2772	91.6
Total	4048	7652	14261	2360	582	1758	3801	34462	
PA (%)	79.6	79.8	80.8	92.4	71.5	62.6	66.8		
	Overall accuracy = 78.6%, Kappa coefficient = 0.720								

Table 6
Comparison of the classification accuracies with various classification methods in Shaanxi province, China.

Method	Overall accuracy (%)	Kappa coefficient
Pixel-object-knowledge-classifier	78.6	0.720
Support vector machine	73.4	0.608
Decision tree (J4.8)	70.4	0.562
Random forest	71.6	0.584
Maximum likelihood classification	63.1	0.462

However, some differences might not come from an actual change in the real world, but from random, gross or systematic errors in land cover characterization. There are a number of temporal constraints between successive land cover datasets which can be used for improving their consistency. For example, in most cases,

artificial surfaces will continue to exist once created. A significant disappearance in the dataset of 2010 with respect to 2000 would be considered as suspicious and be subject to an interactive verification with higher resolution imagery or other source of data. Some exceptions can be found, however, such as the conversion of some buildings into green parks in Beijing in preparation for the 2008 Summer Olympic Games. National statistics from different countries were also used to aid discovery of unreasonable changes in artificial surface.

As land surface water has great fluctuation, differences derived from imagery of 2000 and 2010 may reflect real change (such as the dramatic areal decrease of the Aral Sea), or be subject to seasonal difference and annual rainfall. Such knowledge of temporal constraints was utilized for better understanding and verification of water changes between two 30 m datasets (Liao et al., 2014). The temporal inconsistency caused by the original remotely sensed data was also corrected using MODIS time-series data for the

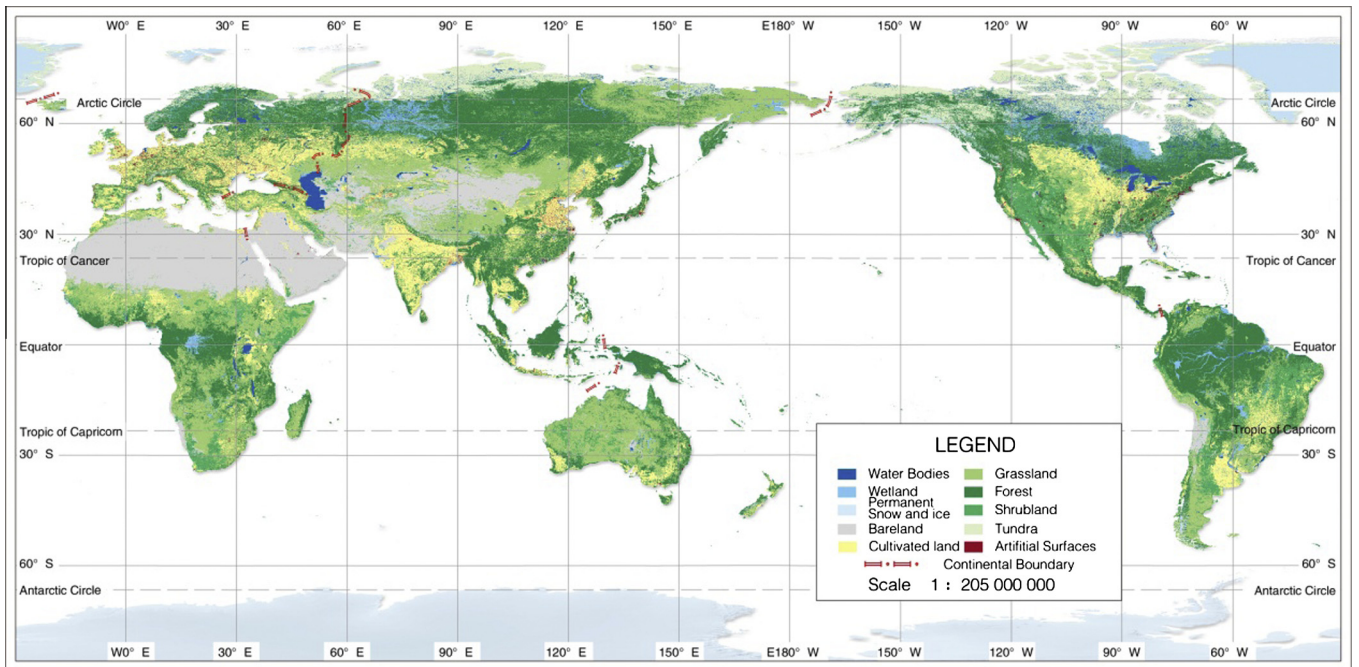


Fig. 10. GlobeLand30 map for the year 2010.

Table 7
Distribution of map sheet samples in continents.

Continent	Asia	Europe	Africa	America	Oceania
Land area ratio	33.01%	7.32%	22.17%	31.19%	6.31%
Size of map sheet samples	26	6	18	25	5
Sum of map sheet samples	80				

calculation and comparison of water areas, water ration and coefficient of spatial variation (Cao et al., 2014).

5.2. Interactive verification and improvement with the support of the web system

During the preparation and implementation of the Chinese GLC project, as much knowledge as possible on land cover and change was analyzed and collected. However, not all the knowledge about

land cover and change could be formalized and used for automated data verification. The success of the project depended greatly on the interactive verification and modification by skillful operators with rich experience. With the support of the web-based system mentioned in Section 3.3, interactive verification was accomplished by a group of expert operators with different land cover knowledge (such as water resources, wetlands, and urban planning) to check possible gross errors or misclassifications by comparison with reference data. The verified results were then sent back to the production stage for removal and reduction of uncertainties from the object-based classification results. Fig. 7 gives two examples of interactive verification conducted with the support of the web-based system.

Table 4 lists the Minimum Mapping Unit (MMU) and allowable minimum error of omission or commission per scene for each land cover type. Levels were defined according to the spatial characteristics of each land cover class and used for guiding the verification process.

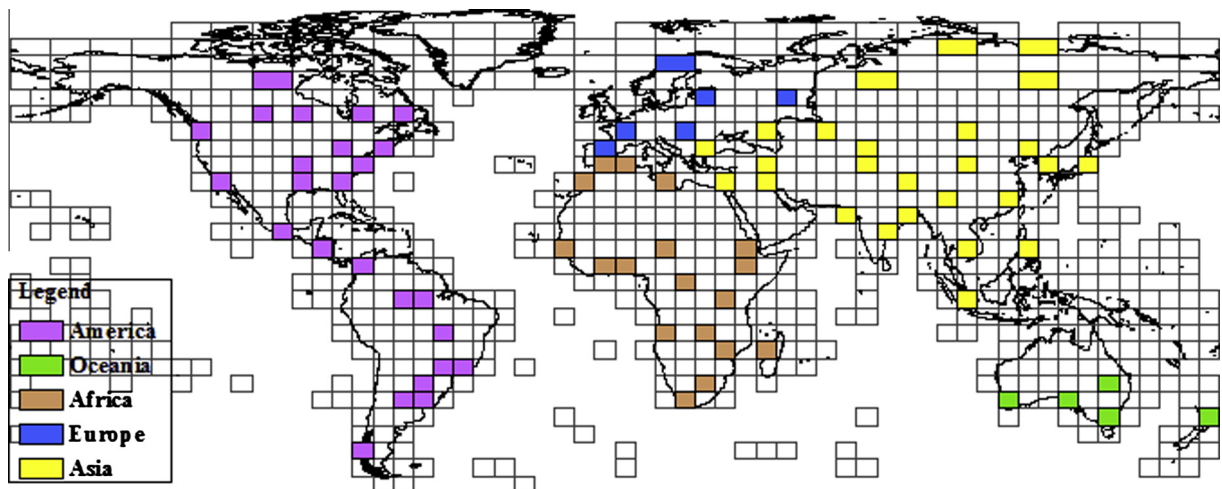


Fig. 11. Spatial distribution of map sheet samples of GlobeLand30.

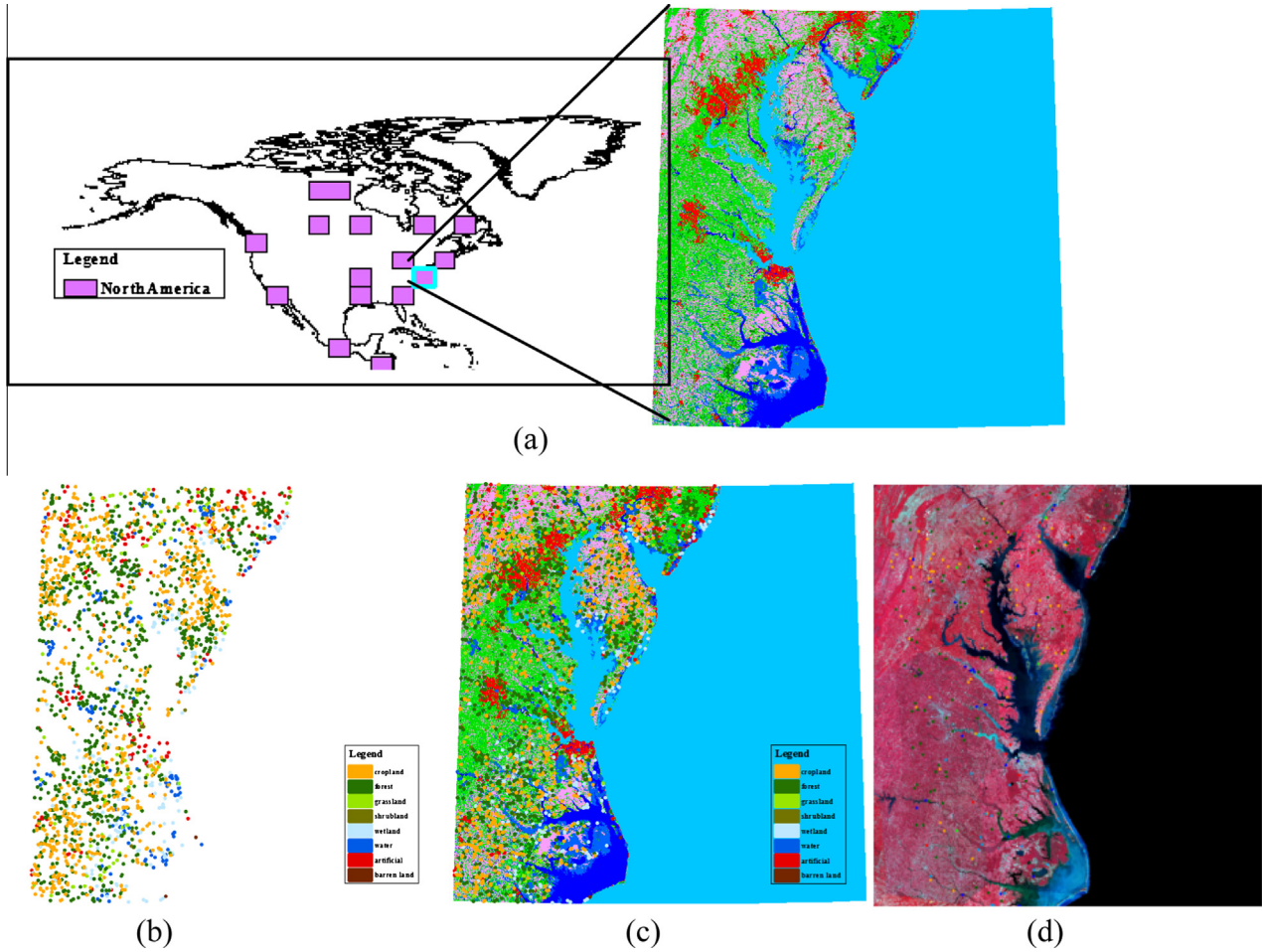


Fig. 12. An example of feature sampling in one map sheet sample in North America. (a) The selected map sheet sample in North America. (b) Pixel samples of all land cover types. (c) Pixel samples in the map sheet. (d) Pixel samples in the corresponding TM image.

Table 8
Distribution of map sheet samples and pixel samples in different regions.

Region	Map sheet samples	Pixel samples (2010 year)
Northeast Asia and Central Asia	13	28,924
Southeast Asia and West Asia	13	30,938
Europe	6	12,765
Africa	18	26,229
America	25	45,817
Oceania	5	9913
Total	80	154,586

Table 9
Preliminary result of accuracy validation of GlobeLand30-2010.

Class	2010			
	User's accuracy (%)	Area ratio	Overall accuracy	Area-weighted overall accuracy
Cultivated land	82.76	0.150478	80.33% ± 0.2%	79.26% ± 0.2%
Forest	83.58	0.299418		
Grassland	72.16	0.244509		
Shrubland	72.64	0.073392		
Wetland	74.87	0.026222		
Water bodies	84.70	0.024507		
Artificial surfaces	86.70	0.009458		
Bareland	81.76	0.153051		
Permanent snow and ice	75.79	0.018964		
Kappa			0.75	

6. Experimental evaluation

6.1. Comparison of POK with other classifiers

To evaluate the performance of the POK-based classification approach, Landsat TM images of Shaanxi province, China, were used for experiments. The area is located in the centre of China (Fig. 8(a)) with various land cover types, latitudinal ecological regions (Fig. 8(b)) and complex terrain.

Besides POK, four other widely used classifiers (SVM, decision tree, random forest and MLC) were selected for comparison. Intensive field investigations were conducted in 2010 by the National Geomatics Center of China (NGCC), and 792 acres (34462 pixels) of ground truth data were collected for accuracy assessment.

In the classification procedure, the training samples required for the supervised classifiers were selected by visual interpretation of TM imagery, which was independent of the collection of the ground truth data.

The land cover map of Shaanxi province produced by POK in 2000 is shown in Fig. 9. By using the collected ground truth data, an error matrix of the POK classification result was obtained (Table 5). There are seven classes in total, as snow/ice, wetland, tundra and bareland do not exist in this area. In general, classification accuracies of water, artificial surface, cultivated land and

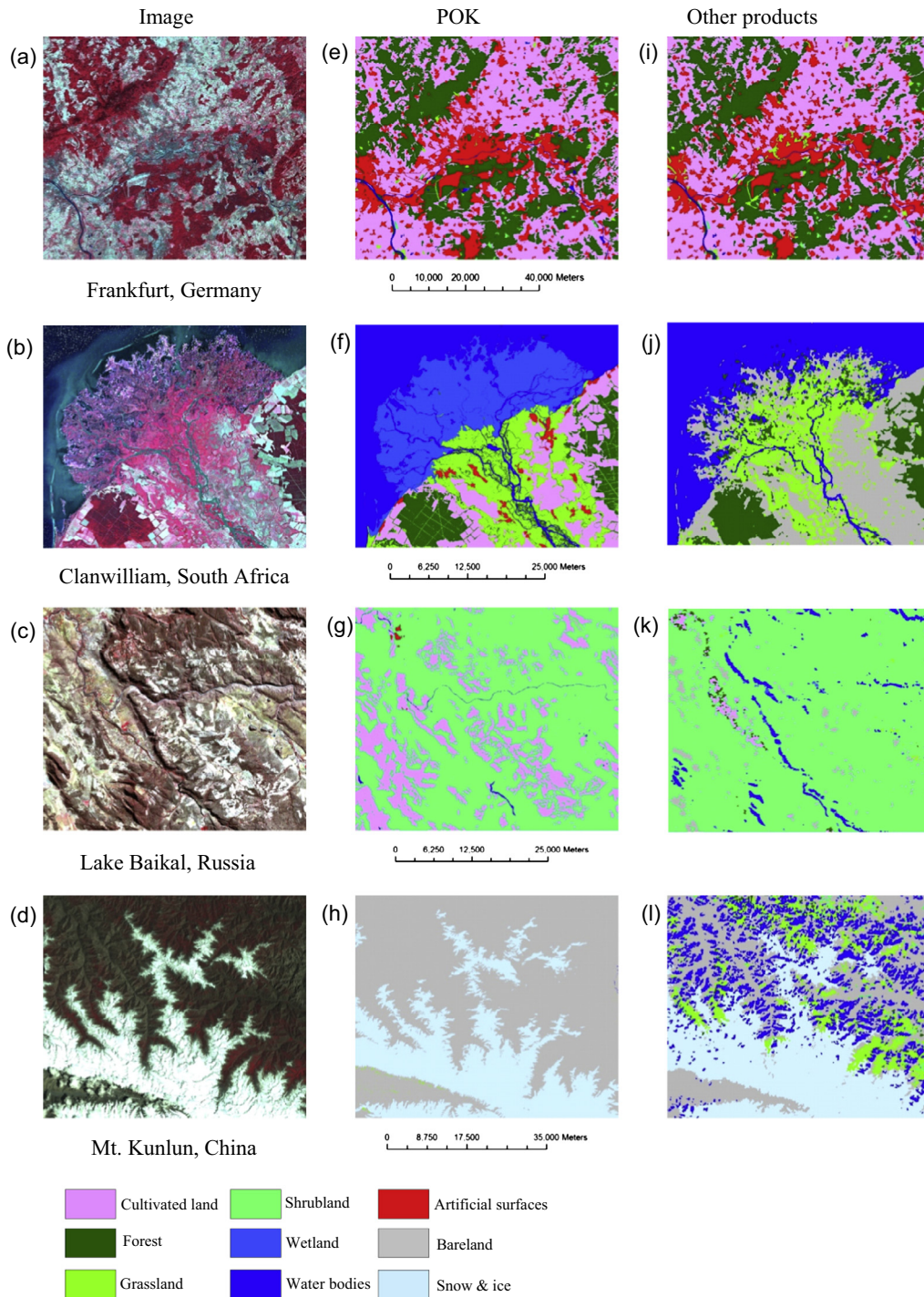


Fig. 13. Comparison between the results from POK and other 30 m land cover products. (a)–(d): Landsat TM images; (e)–(h): classification results by POK; (i) and (j): CORINE 1: 100,000 product; (k) and (l): FROM 30 m product.

forest were found to be relatively higher. However, there was much confusion among grassland, shrub land and bareland because these land cover types are usually mixed. The overall accuracies of POK and other four methods are shown in Table 6. It is noticeable that the POK achieves the highest overall accuracy and Kappa coefficient (0.7017), followed by SVM.

6.2. Preliminary accuracy assessment of GlobeLand30

The proposed POK approach was used as the primary approach by China’s GLC mapping team to produce GlobeLand30, comprising

two 30 m resolution GLC maps for 2000 and 2010. Fig. 10 illustrates the GlobeLand30 map for the year 2010.

A preliminary accuracy assessment was conducted by third-party experts with a two-rank sampling strategy. The first-rank sampling (also called map sheet sampling) involved selecting map sheet samples from global map sheets, and the second-rank sampling (also called feature sampling) selected feature samples of each land cover type within each elected map sheet (Tong et al., 2011; Tong and Wang, 2012). 80 map sheet samples were selected from a total of 847 map sheets in the first-rank sampling, and these samples were spatially distributed into five continents

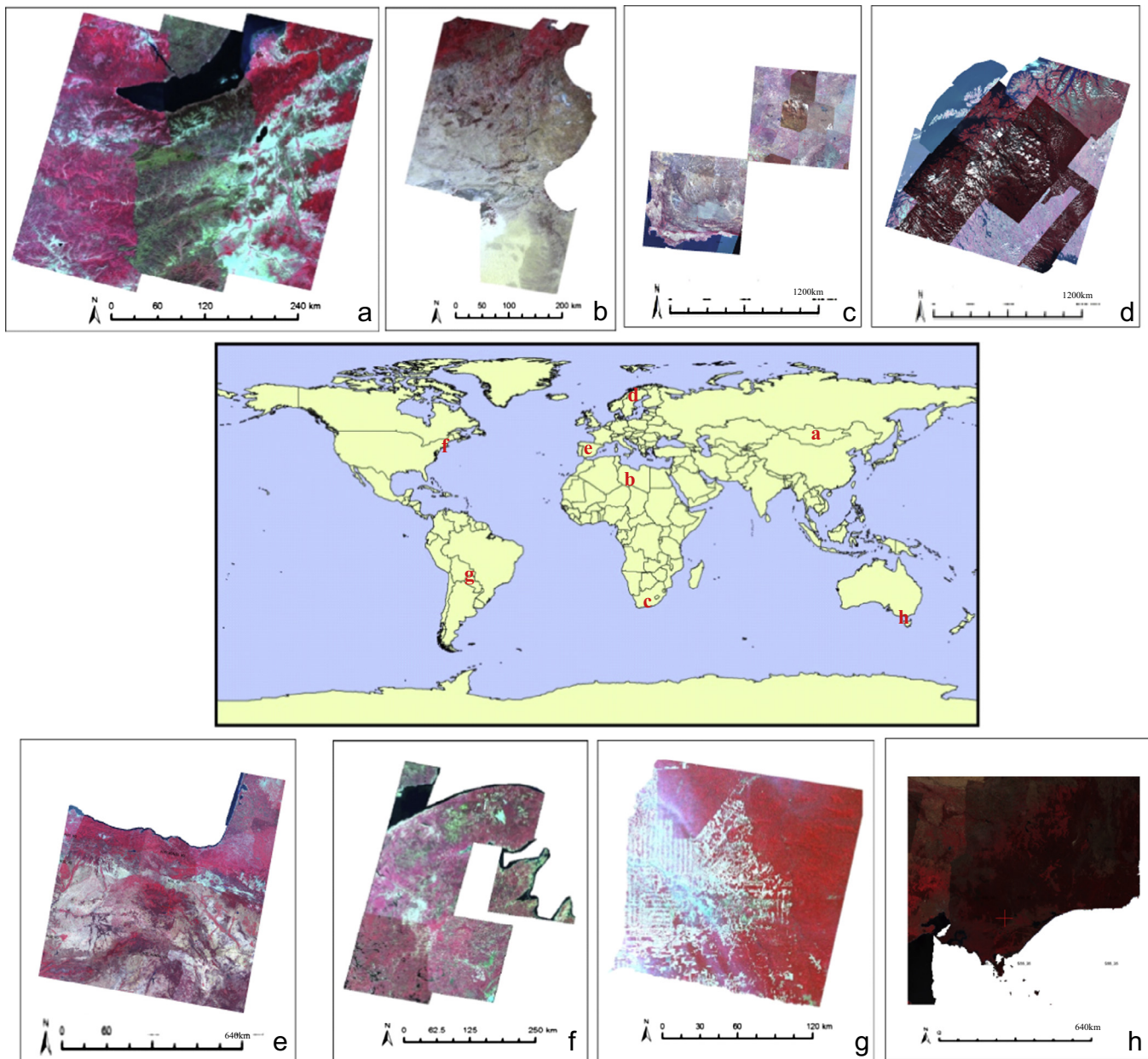


Fig. 14. Eight test areas across the five continents.

based on the land area ratio. Table 7 shows the distribution of map sheet samples, and Fig. 11 shows the spatial location of the 80 map sheet samples.

A total of 159,874 pixel samples were selected for the assessment of GlobeLand30 in the baseline year of 2010, in which 154,586 pixel samples can be definitely validated as correct or incorrect ones. Feature samples were selected using spatial stratification sampling within each selected map sheet, in which the features were stratified by land cover types. The total sample size of feature samples in each map sheet was first calculated based on the optimization sampling model, and the number of feature samples (pixels) in each layer was allocated according to landscape index and layer area ratio to the total sample size. The locations of feature samples were then spatially determined layer by layer based on spatial correlation analysis. Fig. 12 shows an example of feature sampling in one map sheet sample in North America, whilst Table 8 shows the distribution of map sheet samples and pixel samples in different regions.

Table 9 lists the result of the preliminary accuracy assessments for each land cover type in the baseline year 2010. In the table, the user's accuracy and its area ratio for each land cover type are presented, and the area-weighted overall accuracy is calculated by summing the multiplication of user's accuracy and its area ratio of each land cover type (without tundra type). It can be observed that the overall accuracy presented in GlobeLand30 in the year of 2010 is better than 80%, and the kappa coefficient is 0.75.

6.3. Comparison of GlobeLand30 with other 30 m land cover products

The POK classification results (Fig. 13(e–h)) were also compared with other land cover products at 30 m resolution, e.g. CORINE (Fig. 13(i and j)) and FROM (Fig. 13(k and l)). It was found that the quality of GlobeLand30 was at the similar level to CORINE data, which covers only Europe in the base-year of 2000. Although FROM is also at 30 m resolution, it was found to be of much lower quality compared to the POK classification results as it resulted from fully

Table 10

Comparison of the overall accuracy and Kappa of POK and other land cover products.

Global testing areas	Ground truth samples (pixels)	Overall accuracy (kappa coefficient)		
		POK 30 m	CORINE	FROM
Southeast of Baikal Lake, Russia	9012	78.9% (0.737)	/	72.9% (0.615)
Northeastern area, Tunisia	300	93.0% (0.890)	/	72.0% (0.603)
Westcentral area, South Africa	1445	79.72% (0.676)	/	28.46% (0.15)
Northern area, Sweden	2345	81.99% (0.766)	63.68% (0.529)	/
Northwestern area, Spanish	1742	84.73% (0.78)	76.63% (0.666)	/
Southeastern area, Canada	67,983	97.2% (0.950)	/	90.6% (0.802)
Amazon river basin, Brazil	144	79.2% (0.677)	/	65.7% (0.262)
Southeastern area, Australia	2194	96.67% (0.950)	/	42.4% (0.315)

automated classification (Gong et al., 2013). There was also much classification error identified between shadow and water in these examples (Fig. 13(k and l)).

Further quantitative accuracy assessments were conducted in eight test areas across all continents in the world for the year 2000 (Fig. 14). The corresponding ground truth data were collected by third-party researchers through interpretation of high-resolution images (including aerial photos and Google Earth imagery) and field observations. Detailed information of the experimental sites and the ground truth data are shown in Table 10. To achieve consistent accuracy assessment, land cover types were firstly transformed to the classification system used in this paper based on Herold et al. (2008), because the other GLC products used different classification systems. Table 10 lists the overall accuracy and kappa coefficient of the POK product and other two land cover products. From the assessment results, it is clear that the POK product achieves the highest accuracy among these three land cover products. It is noticed that the accuracy of FROM data varies greatly with overall accuracy (as low as 28.46%), implying that the automated classification technique is not robust enough for global mapping at 30 m resolution. These experimental results indicate that the POK-based classification approach is reliable for global land cover mapping. The other land cover maps at coarser resolution were not compared here because they were found to be too different in resolution to obtain meaningful comparison.

7. Conclusions

This paper has reported the primary methodology used in the operational GLC mapping project of China. With the realization that, at global scale with 30 m resolution, fully automated classification techniques are not currently able to meet the requirement of operational mapping, a compromise between effectiveness (accuracy) and efficiency (level of automation) was adopted. As such, in addition to automated classifiers, interactive processes were employed in cases of classification in complex areas and for quality control, resulting in the pixel-object-knowledge based (POK-based) operational approach.

In the developed POK-based approach, the determination of land feature classes takes two steps. The first is to decompose the classification of the 10 land cover classes (types) into simpler per-class classifications in a prioritized sequence. The second step is to merge per-class classification results together. In the determination of each class, an optimal integration of pixel-and object-based classification was developed. To improve the quality of the data products, a knowledge-based interactive verification procedure was developed with the support of web service technology.

The performance of the POK-based approach was tested with eight selected areas with various landscapes and from five different continents. An overall classification accuracy of over 80% was achieved. Based on this limited test, it may be concluded that this POK-based approach is a feasible method for operational mapping at 30 m resolution on a global scale.

With the POK approach as the primary methodology, two 30 m GLC data sets for the years 2000 and 2010 have been produced, entitled GlobeLand30 (<http://www.globallandcover.com>). GlobeLand30 provides more detailed land cover patterns and reflects land cover changes induced by human activities from 2000 and 2010. To assess the quality of GlobeLand30, researchers from China, Greece, Italy, Mexico and Sweden have carried out preliminary evaluations. Large samples have been taken from all over the world and an overall accuracy in excess of 80% has been obtained. From such results, it might be concluded that GlobeLand30 is a reliable product for a number of various applications.

GlobeLand30 has been released for open access and non-commercial utilization at the end of September, 2014. It will be applied to support a variety of international and national scientific programs and development activities. These include the United Nations' post-2015 Sustainable Development Agenda, GEO's GEOSS, ICSU (International Council of Science Union) Future Earth initiative, and the Biodiversity Indicators Partnership program from UNEP-WCMC. This will greatly promote data sharing in the field of geosciences and Earth observation.

Moreover, the sharing of GlobeLand30 will also stimulate GLC mapping and collaborative information services worldwide. More efforts will be mobilized and devoted to the development of increasingly automated approaches for continual updating and data refinement, especially to develop long-time series GLC datasets. An international validation of GlobeLand30 will be organized with the support of UNGGIM and GEO over the next two years.

Acknowledgements

This work has been funded by 863 Project of China (2009AA122001/003) and the National Science Foundation of China (Project #41231172). The authors wish to thank Profs. Guanhua Xu, Jiyan Liu, Chuang Liu, Zhilin Li, Ping Tang, John Townshend, Zengyuan Li, Marguerite Madden and all other team members for their advices and contribution. They also thank all the anonymous reviewers for their constructive comments.

References

- Aitkenhead, M.J., Aalders, I.H., 2011. Automating land cover mapping of Scotland using expert system and knowledge integration methods. *Remote Sens. Environ.* 115 (5), 1285–1295.
- Angel, S., Parent, J., Civco, D.L., Blei, A., Poter, D., 2011. The dimensions of global urban expansion: estimates and projections for all countries, 2000–2050. *Progr. Plann.* 75, 53–107.
- Aplin, P., Smith, G.M., 2011. Introduction to object-based landscape analysis. *Int. J. Geograph. Inf. Sci.* 25 (6), 869–875.
- Ban, Y., 2003. Multitemporal ERS-1 SAR and Landsat TM data for agricultural crop classification: comparison and synergy. *Can. J. Remote Sens.* 29 (4), 518–526.
- Ban, Y., Jacob, A., 2013. Object-based fusion of multi temporal multi angle ENVISAT ASAR and HJ-1 multispectral data for urban land-cover mapping. *IEEE Trans. Geo Sci. Remote Sens.* 51 (4), 1998–2006.
- Ban, Y., Hu, H., Rangel, I., 2010. Fusion of quick bird MS and RADARSAT-1 SAR data for land-cover mapping: object-based and knowledge-based approach. *Int. J. Remote Sens.* 31 (6), 1391–1410.

- Bartholomé, E., Belward, A.S., 2005. GLC2000: a new approach to GLC mapping from earth observation data. *Int. J. Remote Sens.* 26 (9), 1959–1977.
- Berk, A., Anderson, G., Bernstein, L., Acharya, P., Dothe, H., Matthew, M., Adler-Golden, S., Chetwynd Jr., J., Richtsmeier, S., Pukall, B., Allred, C., Jeong, L., Hoke, M., 1999. MODTRAN4 radiative transfer modeling for atmospheric correction. In: *Proceedings SPIE Annual Meeting 3756*, Denver, CO, pp. 348–353.
- Blaschke, T., 2010. Object based image analysis for remote sensing. *ISPRS J. Photogrammetry Remote Sens.* 65, 2–16.
- Bontemps, S., Defournay, P., Van Bogaert, E., Arino, O., Kalogirou, V., Perez, J.P., 2010. GLOBCOVER2009 Products Description and Validation Report, <http://due.esrin.esa.int/globcover/LandCover2009/GLOBCOVER2009_Validation_Report_2.2.pdf>.
- Brisaboa, N.R., Miguel, R., Luaces, M., Rodríguez, A., Seco, D., 2014. An inconsistency measure of spatial data sets with respect to topological constraints. *Int. J. Geograph. Inf. Sci.* 28 (1), 56–82.
- Cao, X., Chen, J., Chen, L., Liao, A., Sun, F., Li, Y., Li, L., Chen, J., He, C., Peng, S., Lin, Z., Pang, Z., 2014. Preliminary analysis of the spatio-temporal pattern of global land surface water. *Sci. China-Earth Sci.* 40 (8), 1661–1670.
- Carnegie, D.M., Lauer, D.T., 1966. Uses of multiband remote sensing in forest and range inventory. *Photogrammetria* 21 (4), 115–141.
- Chen, J., 1984. Utilization of DTM in improving remote sensing classification accuracy. *J. Wuhan Univ. Surveying Mapp.* 9 (1), 69–81.
- Chen, J., Liu, W., Li, Z., Zhao, R., Cheng, T., 2007. Detection of spatial conflicts between rivers and contours in digital map updating. *Int. J. Geograph. Inf. Sci.* 21 (10), 1093–1114.
- Chen, J., Chen, J., Gong, P., Liao, A., He, C., 2011a. Higher resolution GLC mapping. *Geomatics World* 4 (2), 12–14 (in Chinese).
- Chen, J., Zhu, X., Vogelmann, J.E., Gao, F., Jin, S., 2011b. A simple and effective method for filling gaps in Landsat ETM+ SLC-off images. *Remote Sens. Environ.* 115, 1053–1064.
- Chen, X., Chen, J., Shi, Y., Yamaguchi, Y., 2012. An automated approach for updating land cover maps based on integrated change detection and classification methods. *ISPRS J. Photogrammetry Remote Sens.* 71, 86–95.
- Chen, J., Lu, M., Chen, X., Chen, J., Chen, L., 2013a. A spectral gradient difference based approach for land cover change detection. *ISPRS J. Photogrammetry Remote Sens.* 85, 1–12.
- Chen, J., Wu, H., Li, S., Liao, A., He, C., Peng, S., 2013b. Temporal logic and operation relations based knowledge representation for land cover change web services. *ISPRS J. Photogrammetry Remote Sens.* 80 (8), 140–150.
- Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., Peng, S., Han, G., Zhang, H., He, C., Wu, H., Lu, M., 2014a. Some concepts and considerations for 30 m global land cover mapping. *Acta Geodaetica et Cartographica Sinica* 43 (6), 551–557.
- Chen, J., Wang, D., Zhao, R., Zhang, H., Liao, A., Liu, J., 2014b. Fast updating of national geo-spatial databases with high resolution imagery: China's methodology and experiences. *Proc. ISPRS Commission IV Symp.*, 1–10.
- Cihlar, J., 2000. Land cover mapping of large areas from satellites: status and research priorities. *Int. J. Remote Sens.* 21 (6–7), 1093–1114.
- Comber, A., Fisher, P., Wadsworth, R., 2004. Integrating land-cover data with different ontologies: identifying change from inconsistency. *Int. J. Geograph. Inf. Sci.* 18 (7), 691–708.
- Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B., 2004. Digital change detection methods in ecosystem monitoring: a review. *Int. J. Remote Sens.* 25 (9), 1565–1596.
- Costa, H., Carrão, H., Bação, F., Caetano, M., 2014. Combining per-pixel and object-based classifications for mapping land cover over large areas. *Int. J. Remote Sens.* 35 (2), 738–753.
- Croke, B.F.W., Merritt, W.S., Jakeman, A.J., 2004. A dynamic model for predicting hydrologic response to land cover changes in gauged and ungauged catchments. *J. Hydrol.* 291, 115–131.
- Defries, R.S., Townshend, J.R.G., 1999. GLC characterization from satellite data: from research to operational implementation? *Glob. Ecol. Biogeogr.* 8, 367–379.
- Ehlers, M., Edwards, G., Bedard, Y., 1989. Integration of remote sensing with geographic information systems: a necessary evolution. *Photogrammetric Eng. Remote Sens.* 55, 1619–1627.
- Elvidge, C.D., Tuttle, B.T., Sutton, P.C., Baugh, K.E., Howard, A.T., Milesi, C., Bhaduri, B.L., Nemani, R., 2007. Global distribution and density of constructed impervious surfaces. *Sensors* 7, 1962–1979.
- Foley, J., Ruth, D., Asner, G.P., Barford, C., Bonan, G., Carpenter, S.R., Chapin, F.S., Coe, M.T., Daily, G.C., Gibbs, H.K., Helkowski, J.H., Holloway, T., Howard, E.A., Kucharik, C.J., Monfreda, C., Patz, J.A., Prentice, I.C., Ramankutty, N., Snyder, P.K., 2005. Global consequences of land use. *Science* 309 (5734), 570–574.
- Frazier, P., Page, K., 2000. Water body detection and delineation with Landsat TM data. *Photogrammetric Eng. Remote Sens.* 66 (12), 1461–1467.
- 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., Schaaf, C., 2002. Global land cover mapping from MODIS: algorithms and early results. *Remote Sens. Environ.* 83 (1–2), 287–302.
- Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., Huang, X.M., 2010. MODIS collection 5 global land cover: algorithm refinements and characterization of new datasets. *Remote Sens. Environ.* 114 (1), 168–182.
- Fritz, S., See, L., Rembold, F., 2010. Comparison of global and regional land cover maps with statistical information for the agricultural domain in Africa. *Int. J. Remote Sens.* 31, 2237–2256.
- Fritz, S., McCallum, I., Schill, C., Perger, C., See, L., Schepaschenko, D., van der Velde, M., Kraxner, F., Obersteiner, M., 2012. Geo-Wiki: an online platform for improving global land cover. *Environ. Modell. Software* 31, 110–123.
- Gao, W., Gong, J., Li, Z., 2004. Thematic knowledge for the generalization of land use data. *Cartographic J.* 41 (3), 245–252.
- Gerke, M., Butenuth, M., Heipke, C., Willrich, F., 2004. Graph-supported verification of road databases. *ISPRS J. Photogrammetry Remote Sens.* 58, 152–165.
- Giri, C., Pengra, B., Long, J., Loveland, T.R., 2013. Next generation of global land cover characterization, mapping, and monitoring. *Int. J. Appl. Earth Observation Geo Inf.* 25, 30–37.
- Goldewijk, K.K., Ramankutty, N., 2004. Land cover change over the last three centuries due to human activities: the availability of new global data sets. *Geo J.* 61, 335–344.
- Gong, P., 2009. Assessment of GLC map accuracies using flux net location data. *Progr. Nat. Sci.* 19, 754–759.
- Gong, P., Howarth, P.J., 1992. Frequency-based contextual classification and gray level vector reduction for land-use identification. *Photogrammetric Eng. Remote Sens.* 58 (4), 423–437.
- Gong, P., Wang, J., Yu, L., Zhao, Y., Zhao, Y., Liang, L., Niu, Z., Huang, X., Fu, H., Liu, S., Li, C., Li, X., Fu, W., Liu, C., Xu, Y., Wang, X., Cheng, Q., Hu, L., Yao, W., Zhang, H., Zhu, P., Zhao, Z., Zhang, H., Zheng, Y., Ji, L., Zhang, Y., Chen, H., Yan, A., Guo, J., Yu, L., Wang, L., Liu, X., Shi, T., Zhu, M., Chen, Y., Yang, G., Tang, P., Xu, B., Giri, C., Clinton, N., Zhu, Z., Chen, J., Chen, J., 2013. Finer resolution observation and monitoring of GLC: first mapping results with Landsat TM and ETM+ data. *Int. J. Remote Sens.* 34 (7), 2607–2654.
- Goward, S., Williams, D., Arvidson, T., Irons, J., 2011. The future of Landsat-class remote sensing. In: Ramachandran, B., Justice, C.O., Abrams, M.J. (Eds.), *Land Remote Sensing and Environmental Change. Remote Sensing and Digital Image Processing 11*. Springer Science and Business Media, Berlin, Germany, pp. 807–834.
- Grimm, N.B., Faeth, S.H., Golubiewski, N.E., Redman, C.L., Wu, J., Bai, X., Briggs, J.M., 2008. Global change and the ecology of cities. *Science* 319, 756–760.
- Han, G., Chen, J., He, C.Y., Li, S.N., Wu, H., Liao, A.P., Peng, S., 2015. A web-based system for supporting GLC data production. *ISPRS J. Photogrammetry Remote Sens.* 103, 66–80.
- Hansen, M.C., Loveland, T.R., 2012. A review of large area monitoring of land cover change using Landsat data. *Remote Sens. Environ.* 122, 66–74.
- Hansen, M.C., Defries, R.S., Townshend, J.R.G., Sohlberg, R., 2000. GLC classification at 1km spatial resolution using a classification tree approach. *Int. J. Remote Sens.* 21 (6–7), 1331–1364.
- Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., Townshend, J.R.G., 2013. High-resolution global maps of 21st-century forest cover change. *Science* 342 (15), 851–853.
- Hayakawa, Y.S., Oguchi, T., Lin, Z., 2008. Comparison of new and existing global digital elevation models: ASTER G-DEM and SRTM-3. *Geophys. Res. Lett.* 35 (17), L17404.
- He, C., Okada, N., Zhang, Q., Shi, P., Zhang, J., 2006. Modeling urban expansion scenarios by coupling cellular automata model and system dynamic model in Beijing, China. *Appl. Geography* 26 (3–4), 323–345.
- Heipke, C., Woodsford, P.A., Gerke, M., 2008. Updating geospatial databases from images. In: Li, Zhilin, Chen, Jun, Baltasavias, E. (Eds.), *Advances in Photogrammetry, Remote Sensing and Spatial Information Sciences*. CRC Press/Balkema, pp. 355–362.
- Herold, M., Mayaux, P., Woodcock, C., Baccini, A., Schullius, C., 2008. Some challenges in GLC mapping: an assessment of agreement and accuracy in existing 1 km datasets. *Remote Sens. Environ.* 112, 2538–2556.
- Hu, C.W., Zhang, Feng, Z., Tang, P., 2014. Landsat TM/ETM+ and HJ-1A/B CCD data automatic relative radiometric normalisation and accuracy verification. *J. Remote Sens.* 18 (2), 267–276.
- Huang, B., 2012. Spatiotemporal reflectance fusion via sparse representation. *IEEE Trans. Geosci. Remote Sens.* 50, 3707–3716.
- Huang, Z., Jia, X., 2012. Integrating remotely sensed data, GIS and expert knowledge to update object-based land use/land cover information. *Int. J. Remote Sens.* 33 (4), 905–921.
- Hussain, M., Chen, D., Cheng, A., Wei, H., Stanley, D., 2013. Change detection from remotely sensed images: from pixel-based to object-based approaches. *ISPRS J. Photogrammetry Remote Sens.* 80, 91–106.
- Iwao, K., Nishida, K., Kinoshita, T., Yamagata, Y., 2006. Validating land cover maps with degree confluence project information. *Geophys. Res. Lett.* 33, L23404.
- Johnson, D.M., Mueller, R., 2010. The 2009 crop land data layer. *Photogrammetric Eng. Remote Sens.* 11, 1201–1205.
- Kushwaha, S.P.S., 1990. Forest-type mapping and change detection from satellite imagery. *ISPRS J. Photogrammetry Remote Sens.* 45, 175–181.
- Liao, A., Chen, L., Chen, J., He, C., Cao, X., Chen, J., Peng, S., Sun, F., Gong, P., 2014. High-resolution remote sensing mapping of global land water. *Sci. China Earth Sci.* 40 (8), 1634–1645.
- Liu, J., Liu, M., Deng, X., Zhuang, D., Zhang, Z., Luo, D., 1999. The land use and land cover database and its relative studies in China. *J. Geographic Sci.* 12 (3), 275–282.
- Loveland, T.R., Reed, B.C., Brown, J.F., Ohlen, D.O., Zhu, Z., Yang, L., Merchant, J.W., 2000. Development of a GLC characteristics database and IGBP discover from 1 km AVHRR data. *Int. J. Remote Sens.* 21 (6–7), 1303–1330.
- Lu, D., Weng, Q., 2007. A survey of image classification methods and techniques for improving classification performance. *Int. J. Remote Sens.* 28 (5), 823–870.
- Lu, D., Mausel, P., Brondizio, E., 2004. Change detection techniques. *Int. J. Remote Sens.* 28 (5), 823–870.
- Malinverni, E.S., Tassetti, A.N., Mancini, A., Zingaretti, P.Z., Frontoniand, E., Bernardini, A., 2011. Hybrid object-based approach for land use/land cover

- mapping using high spatial resolution imagery. *Int. J. Geographical Inf. Sci.* 25 (6), 1025–1043.
- MDA, 2014. EarthSatGeoCover LC Overview, <<http://www.mdafederal.com/geocover/geocoverlc>>.
- Mills, J.P., Newton, I., 1996. A new approach to the verification and revision of large-scale mapping. *ISPRS J. Photogrammetry Remote Sens.* 51, 17–27.
- Myint, S.W., Gober, P., Brazel, A., Grossman-Clarke, S., Weng, Q., 2011. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sens. Environ.* 115 (5), 1145–1161.
- Ok, A.O., Akyurek, Z., 2012. A segment-based approach to classify agricultural lands by using multi-temporal optical and microwave data. *Int. J. Remote Sens.* 33 (22), 7184–7204.
- Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'Amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P., Kassem, K.R., 2001. Terrestrial ecoregions of the world: a new map of life on earth. *Bioscience* 51 (11), 933–938.
- Otsu, N., 1979. A threshold selection method from grey-level histograms, *IEEE Transactions on Systems, Man, and Cybernetics SMC-9*, pp. 62–66.
- Potere, D., Schneider, A., Angel, S., Civco, D.L., 2009. Mapping urban areas on a global scale: which of the eight maps now available is more accurate? *Int. J. Remote Sens.* 30 (24), 6531–6558.
- Ramakutty, N., Evan, A.T., Monfreda, C., Foley, J.A., 2008. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Global Biogeochem. Cycles* 22, GB1003.
- Rao, Y., Zhu, X., Chen, J., Wang, J., 2014. A direct method for producing high spatial-resolution NDVI time series datasets with multi-temporal MODIS NDVI data and a Landsat TM/ETM+ image. *IEEE Trans. Geosci. Remote Sens.* (submitted for publication).
- Raymond, P.A., Hartmann, J., Lauerwald, R., Sobek, S., McDonald, C., Hoover, M., Butman, D., Striegl, R., Mayorga, E., Humborg, C., Kortelainen, P., Meybeck, M., Ciais, P., Guth, P., 2013. Global carbon dioxide emissions from inland waters. *Nature* 503, 355.
- Robertson, L.D., King, D.J., 2011. Comparison of pixel- and object-based classification in land cover change mapping. *Int. J. Remote Sens.* 32 (6), 1505–1529.
- Rogana, J., Chen, D., 2004. Remote sensing technology for mapping and monitoring land-cover and land-use change. *Progr. Plann.* 61, 301–325.
- Roger, A., Pielke, Sr., 2005. Land use and climate change. *Science* 310, 1625–1626.
- Running, S.W., 2008. Ecosystem disturbance, carbon, and climate. *Science* 321, 652–653.
- Ryherd, S., Woodcock, C.E., 1996. Combining spectral and texture data in the segmentation of remotely sensed images. *Photogrammetric Eng. Remote Sens.* 62 (2), 181–194.
- Smith, G., 2013. Hybrid pixel- and object-based approach to habitat condition monitoring. In: *Proceedings GI Forum - Creating the GI Society* 3756. Wichmann-Verlag, Berlin, pp. 552–555.
- Sterling, S.M., Ducharme, A., Polcher, J., 2013. The impact of global land-cover change on the terrestrial water cycle. *Nat. - Climate Change* 3, 385–390.
- Sulla-Menashe, D., Friedl, M.A., Krankina, O.N., Baccini, A., Woodcock, C.E., Sibley, A., Sun, G., Kharuk, V., Elsakov, V., 2011. Hierarchical mapping of northern Eurasian land cover using MODIS data. *Remote Sens. Environ.* 115, 392–403.
- Sun, F., Sun, W., Chen, J., Gong, P., 2012. Comparison and improvement of methods for identifying water bodies in remotely sensed imagery. *Int. J. Remote Sens.* 33 (21), 6854–6875.
- Sutton, P.C., Anderson, S.J., Elvidge, C.D., Tuttle, B.T., Ghosh, T., 2009. Paving the planet: impervious surface as proxy measure of the human ecological footprint. *Prog. Phys. Geogr.* 33 (4), 510–527.
- Tang, P., Zhang, H.W., Zhao, Y., Niu, Z., Zhong, B., Huang, C., Shan, X., 2014. Practice and thoughts of the automatic processing of multispectral images with 30 m spatial resolution on the global scale. *J. Remote Sens.* 18 (2), 231–241.
- Tong, X.H., Wang, Z.H., 2012. Fuzzy acceptance sampling plans for inspection of geospatial data with ambiguity in quality characteristics. *Comput. Geosci.* 48 (11), 256–266.
- Tong, X.H., Wang, Z.H., Xie, H., Liang, D., Jiang, Z.Q., Li, J.C., Li, J., 2011. Designing a two-rank acceptance sampling plan for quality inspection of geospatial data products. *Comput. Geosci.* 37 (10), 1570–1583.
- Townshend, J., Justice, C., Li, W., Gurney, C., McManus, J., 1991. GLC classification by remote sensing: present capabilities and future possibilities. *Remote Sens. Environ.* 35, 243–255.
- Townshend, J.R., Masek, J.G., Huang, C., Vermote, E.F., Gao, F., Channan, S., Sexton, J.O., Feng, M., Narasimhan, R., Kim, D., Song, K., Song, D., Song, X., Noojipady, P., Tan, B., Hansen, M.C., Li, M., Wolfe, R.E., 2012. Global characterization and monitoring of forest cover using Landsat data: opportunities and challenges. *Int. J. Digital Earth* 5, 373–393.
- Trias-Sanz, R., Stamon, G., Louchet, J., 2008. Using colour, texture, and hierarchical segmentation for high-resolution remote sensing. *ISPRS J. Photogrammetry Remote Sens.* 63 (2), 156–168.
- Verbesselt, J., Hyndman, R., Newnham, G., Culvenor, D., 2010. Detecting trend and seasonal changes in satellite image time series. *Remote Sens. Environ.* 114 (1), 106–115.
- Verburg, P.H., Neumann, W., Linda, N.L., 2011. Challenges in using land use and land cover data for global change studies. *Glob. Change Biol.* 17, 974–989.
- Whigham, D.F., 2009. Global distribution, diversity and human alterations of wetland resources. In: Barker, Tom, Maltby, Edward (Eds.), *The Wetlands Handbook*. Blackwell Publishing Ltd.
- Wulder, M.A., White, J.C., Goward, S.N., Masek, J.G., Irons, J.R., Herold, M., Cohen, W.B., Loveland, T.R., Woodcock, C.E., 2008. Landsat continuity: issues and opportunities for land cover monitoring. *Remote Sens. Environ.* 112, 955–969.
- Xian, G., Collin, H., Fry, J., 2009. Updating the 2001 national land cover database land cover classification to 2006 by using Landsat imagery change detection methods. *Remote Sens. Environ.* 113 (6), 1133–1147.
- Yin, D., Cao, X., Chen, X., Shao, Y., Chen, J., 2013. Comparison of automatic thresholding methods for snow-cover mapping using Landsat TM imagery. *Int. J. Remote Sens.* 34 (19), 6679–6700.
- Yu, L., Gong, P., 2012. Google earth as a virtual globe tool for earth science applications at the global scale: progress and perspectives. *Int. J. Remote Sens.* 33 (12), 3966–3986.
- Zell, E., Huff, A.K., Carpenter, A.T., Friedl, L.A., 2012. A user-driven approach to determining critical earth observation priorities for societal benefit. *IEEE J. Selected Topics Appl. Earth Observations Remote Sens.* 5 (6), 1594–1602.
- Zhang, W., Wu, W., Cui, Y., Wen, Q., 2012. HJ-1 satellite image geometric correction system design, *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 22–27, Munich, pp. 4351–4354.
- Zhou, Y., Chen, J., Guo, Q., Rao, Y., Zhu, X., 2014. Restoration of information obscured by mountainous shadows through Landsat TM/ETM+ images without the use of DEM data: a new method. *IEEE Trans. Geosci. Remote Sens.* 52 (1), 313–328.
- Zhu, X., Gao, F., Liu, D., Chen, J., 2012. A modified neighborhood similar pixel interpolator approach for removing thick clouds in Landsat Images. *IEEE Geosci. Remote Sens. Lett.* 9 (3), 521–525.