Global land cover mapping using Earth observation satellite data: Recent progresses and challenges

1. Introduction

Land cover is an important variable for many studies involving the Earth surface, such as climate, food security, hydrology, soil erosion, atmospheric quality, conservation biology, and plant functioning. Land cover not only changes with human caused land use changes, but also changes with nature. Therefore, the state of land cover is highly dynamic. In winter snowy conditions under various other land cover types in higher latitudes. Flooding may persist for a long period in a year over low land areas in the tropical and subtropical regions. Forests may burn or clear cut in a few days and change to bare land. Within several months, the coverage of crops may vary from bare land to nearly 100% crops and then back to bare land following harvest. The highly dynamic nature of land cover creates a challenge in monitoring which remains to be adequately addressed. As economic globalization continues to intensify, there is an increasing trend of land cover/land use change, environmental pollution, land degradation, biodiversity loss at the global scale, timely and reliable information on land cover and its changes is urgently needed to mitigate the negative impact of global environment change.

2. Recent progresses in global land cover mapping and change monitoring

Remote sensing has been widely recognized as the most economic and feasible approach to derive land-cover information over large areas (Cihlar, 2000; Gong, 2012). Land cover is enlisted as one of the essential climate variable in the global climate observation system (GCOS) (Bojinski et al., 2014) that can be technically and economically feasible for systematic observation. Despite the fact that a large amount of research on land cover mapping at various spatial scales has been carried out (Lu and Weng, 2007; Yu et al., 2014), there are only a handful of land cover maps produced at global scale. To meet the need for global climate change studies, several global land cover (GLC) map products have been developed at spatial resolutions ranging from 100 km to 300 m (DeFries and Townshend, 1994; Loveland et al., 2000; Hansen et al., 2000; Bartholomé and Belward, 2005; Friedl et al., 2010; Arino et al., 2008; Tateishi et al., 2011). However, it has soon been realized that these data products are difficult to harmonize in terms of classification system and the large errors in areas with rapidly changing ecotones (Giri et al., 2005; Herold et al., 2008; Verburg et al., 2011). Furthermore, previous global land cover products derived using time series optical satellite data at coarse spatial resolution (300 m–1 km) did not provide sufficient thematic detail or change information for global change studies and for resource management (Giri et al., 2013). Following a survey on the land cover data requirements of climate system modelers, Bontemps et al. (2012) found that higher spatial resolution and more temporal frequent data products were needed. In the meantime, requirement for finer resolution land cover mapping at the global scale has emerged in the field of biodiversity, food security and forest carbon studies (Dobson, 2005; Buchanan et al., 2009; Fritz et al., 2013; Giri et al., 2013; Pereira et al., 2013). Higher resolution (~30 m) land cover characterization and monitoring permits detection of land change at the scale of most human activity and offers the increased flexibility of environmental model parameterization needed for global change studies. However, there are a number of challenges to overcome before producing such data sets including unavailability of consistent global coverage of satellite data, sheer volume of data, unavailability of timely and accurate training and validation data, difficulties in preparing image mosaics, and high performance computing requirements (Giri et al., 2013).

Thanks to the free availability of Landsat imagery, global land cover mapping at 30 m resolution has become possible. Based on the Landsat Thematic Mapper (TM) images acquired around 2010, China produced the world’s first 30 m global land cover maps in two versions, namely Finer Resolution Observation and Monitoring of Global Land Cover (FROM-GLC) and GlobeLand30. FROM-GLC contains two levels of land cover classes (10 Level-1 classes and 29 Level-2 classes) and was produced using four per-pixel image classification approaches (can be freely downloaded at: data.tsinghua.edu.cn). While efforts had been made in selecting images from the “greenest” and “wettest” seasons, the end results exhibited considerable “mosaic” effect due to different acquisition dates among neighboring images. The overall accuracies of Level-1 FROM-GLC range from 52.76% (Maximum Likelihood Classifier) to 53.88% (Maximum Likelihood Classifier) to 64.89% (Support Vector Machine) while the best overall accuracy of Level-2 FROM-GLC is 52.76%. Several classes including impervious areas, croplands, grasslands, and shrublands were poorly classified (Gong et al., 2013). Such low accuracy is possibly attributable to significant spectral confusion among different land cover types (Chen et al., 2015) due to the lack of temporal features in the single-date Landsat images used (Yu et al., 2013a). Recognizing the difficulties of fully automated classification techniques to attain GLC maps at sufficiently high accuracy for operational use, a pixel-object-knowledge based classification approach (POK) was developed to produce GlobeLand30, an operational 30 m global land cover dataset. GlobeLand30 consists of 10 major classes and two base-line years, 2000 and 2010 with an overall classification accuracy at 80.33% (Chen et al., 2015). It
should be noted that (1) the accuracies of FROM-GLC and GlobeLand30 were derived from different validation datasets thus are not directly comparable; and (2) the knowledge-based verification and refinement process in the POK method is rather labor intensive and time consuming. The GlobeLand30 dataset was donated to the United Nations on September 22, 2014 and available for open access (www.globeland30.org, Chen et al., 2014). This is a milestone achievement in the Earth Observation and open geo-information access.

Attempts were also made to use coarser resolution time-series Moderate Resolution Imaging Spectrometer (MODIS) data and auxiliary data to improve the FROM-GLC classification results and an overall classification accuracy of 67.08% was obtained (Yu et al., 2013a), a moderate improvement from the original 64.89%.

Although a number of single-category global data products such as cropland, urban and water bodies have been developed based on coarser resolution satellite data (Ramanakutty and Foley, 1998; Schneider et al., 2009; Carroll et al., 2009), single-category land-cover data product derived from Landsat data are still rare. One of the early attempts to develop single-category land cover product at the global scale with Landsat TM data is the mapping of global mangrove distribution (Giri et al., 2011). A 30 m global cropland map has been developed following an integrated analysis of the FROM-GLC cropland results with time-series MODIS data (Yu et al., 2013b). Recently, a number of fractional products on forest cover at the global scale were developed using Landsat TM data (Hansen et al., 2013; Sexton et al., 2013) and synthetic aperture radar data (ALOS PALSAR) at 25 m level for 2007–2010 (Shimada et al., 2014). For global urban mapping, automated urban extraction still remains a challenging task. Several efforts were made to develop methods for global human settlements extraction using high and very high resolution (0.5–10 m) optical data (Pesaresi et al., 2013), TanDEM SAR data at 18 m resolution (Esch et al., 2013), and ENVISAT ASAR wide swath mode at 75 m resolution (Gamba and Lisini, 2013). In response to a strong need of robust and operational methods for global urban extraction with a small number of SAR images at medium resolution, Ban et al. (2015) evaluated ENVISAT ASAR data for global urban mapping at 30 m resolution using a robust processing chain, the KTH-Pavia Urban Extractor. The overall average accuracy for 10 global cities was achieved at 85.4% using a single SAR image.

Change detection at global scale is more difficult than land cover classification for a single time instant. Due to random errors and the lack of sufficiently high classification accuracy for most land-cover products, it is generally not advised to compare land-cover maps produced at different times to derive land-cover change information (Friedl et al., 2010). Moreover, these data sets were produced using different data sources, classification system, and classification methodologies. However, in recent years a number of change detection and trajectory analysis algorithms have been proposed for single-category land-cover types such as forest change and urbanization (e.g., Kennedy et al., 2010; Huang et al., 2010; Xian et al., 2011; Wang et al., 2012; Yousif and Ban, 2013; Hu and Ban, 2014). New algorithms are being developed based on trajectory analysis to derive changes for multiple land-cover types (Zhu and Woodcock, 2014). Nevertheless, most of these algorithms have not yet been applied to the global level due to their requirement of large volume of data and/or intensive computation. At the global scale, Hansen et al. (2013) derived annual forest loss and gain maps from 2000 to 2012 with Landsat data while Song et al. (2014) determined annual forest loss with MODIS data for 2000–2012.

### 3. Remaining challenges and opportunities

Mapping land cover over large areas is considerably more difficult than mapping small areas. For small areas, it is easier to collect data because a single-date image may cover the entire area of interest thus reducing inter-scene variability. In addition, it is possible to use manual interpretation which is usually more accurate when the image interpreter is familiar with the area. It is easy to conduct field visit on those areas that are hard to interpret. It is also easier to conduct validation and refinement of preliminary mapping results. As the study area increases to the continental and global scale, we face with a number of challenges at both the cognition and technology levels. At the cognition level, some of the challenges include:

- Can the design of classification system meet the need of a specific application and take advantage of the available remotely sensed data? Different application requires different classification scheme. For example, to model the carbon sequestration by the terrestrial ecosystem or the energy and water exchange between land and atmosphere, the research community requires that the land cover classes include classes as specific as plant functioning types. However, to help plan for meeting the Aichi Biodiversity Targets, conservation specialists and policy makers require information on land cover categories as broad as settlement, cropland, forest, wetland, water, etc.

- Can the image analyst(s) correctly associate each class with remotely sensed data? In supervised image classification, this capability determines the quality of training sample and in turn the quality of the outcome of the classification algorithm used. In unsupervised image classification, this capability also determines quality of cluster labeling. As the spatial scale increases, the level of requirement on the geometrical knowledge and interpretation skill of an image analyst increases considerably because it is not practical to verify every uncertain land cover class on site.

- Can we collect sufficient representative samples for training and validation? More image analysts are needed at large spatial scale as the size of training or validation sample increases. Differences always exist in image interpretation by several analysts. Large inconsistencies among different image analysts would increase confusions in algorithm calibration and in accuracy assessment. How to reduce the inconsistancies among different image analysts is both an educational problem and a project management problem. Accumulation of successful global knowledge on land cover classification and building and maintaining a high quality team of image analysts are key elements in global land cover mapping and change monitoring.

Since minimizing the involvement of image analyst may help increase the objectivity of final image classification results and has potential for increased efficiency and reproducibility, there are tendencies of relying on technological advancement to overcome the shortcoming of land cover mapping methods involving large amount of human intervention. At regional and continental scale, a number of maps have been produced with unsupervised classification approaches (Cihlar, 2000; Bartholomé and Belward, 2005; Hu et al., 2014). However, so far supervised classification or image interpretation still has its superiority in terms of final classification accuracies (Zhang et al., 2014). Thus, the amount of efforts made by expert analysts determines to a large extent the quality of final map product. To reduce the involvement of image analysts in the classification process, it is important to accumulate samples that can be used for signature extension and validating multiple land-cover data products (Demir et al., 2013; Zhao et al., 2014).
At the technology level, a number of reviews have emphasized on the need for collecting new and finer resolution data at hard-to-map areas, fractional land-cover (continuous field) data in addition to distribution of land-cover types, and monitoring of more rapidly changing areas as primary challenges (Herold et al., 2008; Verboug et al., 2011; Giri et al., 2013). In 2000, Cihlar (2000) suggested some future needs in large area mapping including (1) better geometric and atmospheric correction, (2) creating better image composites, (3) increasing robustness of algorithms, (4) incorporating more non-spectral features such as spatial and biophysical parameters and, (5) minimizing human involvements. Since 2000, enormous progress has been achieved in data processing at the global scale, particularly for data acquired on board of the U.S. satellites such as Terra, Aqua and the Landsat series. Reflectance data corrected to the top-of-the-atmosphere level are available in image composites for every 16 days at 30 m, and daily for 250 m and 500 m resolution with average geometric accuracies better than 1 pixel (e.g., Roy et al., 2008, 2010). More consistent data in combination with recent new machine learning algorithms have improved the robustness of image classification procedures. Longer and denser time-series of satellite data allow us to derive better phenological features that can be used to separate crop species that would otherwise impossible to differentiate (Zhong et al., 2014). It has been recently found that although newer classification algorithms are being continuously proposed, the contribution of classifiers to the improvement of land cover mapping is not as big as that of optical features and increasing representativeness of training sample (Li et al., 2014; Yu et al., 2014). However, in global land cover mapping, use of spatial features is rare (Durieux et al., 2007). Some super-resolution algorithms that combine data acquired from slightly different angles and data fusion algorithms to increase spatial resolution may create more opportunities for use of spatial features (Shen et al., 2009; Gao et al., 2006). Synergistic use of data from thermal and microwave sensors with optical data is also rare in global land cover mapping. At small spatial scale, the fusion of SAR and optical data or combination of thermal and optical data has proven to be helpful in improving land cover and land use classification (e.g. Lambin and Ehrlich, 1996; Ban et al., 2010; Julien et al., 2011; Warner and Nerry, 2009; Ban and Jacob, 2013). Combining LiDAR data with multispectral data has also proven to be able to substantially improve land cover classification accuracy (Chiri et al., 2013). Processing SAR data from the European Space Agency satellites in the same manner as what has been done for the Japanese ALOS PALSAR data (http://www.eorc.jaxa.jp/ALOS/en/palsar_fni/fni_index.html) could provide an important historical archive that complement the use of optical data for land-cover mapping over cloudy regions of the world. Obtaining representative training sample over large spatial area is still a challenge. Image analysts tend to use what they are familiar with in collecting training sample units, thus leaving some biases in the training sample. The short of representative training sample with in collecting training sample units, thus leaving some biases in the training sample. The short of representative training sample and standardized data processing can be done, the situation of lacking for data could be substantially improved. On the data pre-processing aspect, although geometric and atmospheric correction algorithms have gained a lot of progress, the correction for topographic effects has lagged behind (Huang et al., 2008; VanOeneckelen et al., 2013). Reducing the involvement of the image analysts in the classification process – accumulate samples that can be used for multiple tasks will have considerable benefits to future global land-cover mapping.

4. Contributions in this special theme issue

The theme issue intended to bring together a volume of high quality research papers on the current situation in terms of progress, challenges, and opportunities on the mapping and monitoring of global land cover conditions at various scales, particularly at 30 m resolution. This theme issue contains ten papers covering a variety of topics and shows that tremendous efforts have been made to tackle many of the challenges discussed in the previous section and significant progresses have been made in five major areas as summarized below:

- **Improved global land cover mapping using fine-resolution data:** Two papers are devoted to this topic. In the first paper, Chen et al. developed a pixel-object-knowledge (POK) based approach for operational global land cover mapping at 30 m resolution (GlobeLand30) for 2000 and 2010 using Landsat TM images. For the identification of each class, automated pixel- and object-based classifications were initially used. 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 evaluated using eight selected areas with differing landscapes from five different continents. An overall classification accuracy of over 80% was achieved. These results indicate that the POK-based classification approach is feasible and reliable for global land cover mapping. In the second paper, Ban et al. evaluate ENVISAT ASAR data for improved global urban mapping using a robust processing chain, the KTH-Pavia Urban Extractor. The processing chain includes urban extraction based on spatial indices and Grey Level Co-occurrence Matrix (GLCM) textures, an existing method and several improvements i.e., SAR data preprocess- ing, enhancement, and post-processing. Using single-date, single polarization ASAR data, average urban extraction accuracy at 85.4% was achieved for 10 global cities, much higher than that of MODIS 500 and GlobeCover urban class. The results indicate that efficient and reliable global urban mapping is possible at low cost using spaceborne SAR data.

- **Cutting-edge technological issues related to global land cover mapping, such as the utilization of new sensor data, longer and denser time-series data, fusion of data from multiple sources, advanced classification algorithms:** Four papers are dedicated to this topic. In the first paper, Wang et al. (2015) mapped global land cover types in 2001 and 2010 at 250 m resolution with multiple year time series MODIS data. A random forest classifier was used to obtain original class labels and to estimate class probability accuracies for 2000–2002, and 2009–2011. The overall accuracies estimated from cross validation of samples were 74.93% for 2001 and 75.17% for 2010. The classification results were further improved through post processing using a spatial–temporal consistency model, Maximum a Posteriori Markov Random Fields (MAP-MRF), and a rule-based label adjustment method with MOD44B, slope and composited EVI series as auxiliary data. The label adjustment process labeled the over-classified forests, water bodies and barren...
lands to alternative classes with maximum probabilities. The research also identifies the importance of regional information and knowledge, the number of samples and distributions of samples in the feature space, as well as more training and validation points for improving GLC classification accuracies. In the second paper, See et al. (2015) demonstrate how existing medium resolution global land cover maps can be integrated into a hybrid product using GWR and a training dataset obtained through crowd sourcing. Two different methods were used: (a) the GWR was used to determine the best land cover product at each location; (b) the GWR was only used to determine the best land cover at those locations where all three land cover maps disagree, using the agreement of the land cover maps to determine land cover at the other cells. The results show that the hybrid land cover map developed using the first method resulted in a lower overall disagreement than the individual global land cover maps. The hybrid map produced by the second method was also better when compared to the GLC2000 and GlobCover but worse or similar in performance to the MODIS land cover product depending upon the metrics considered. The reason for this may be due to the use of the GLC2000 in the development of GlobCover, which may have resulted in areas where both maps agree with one another but not with MODIS, and where MODIS may in fact better represent land cover in those situations. These results serve to demonstrate that spatial analysis methods can be used to improve medium resolution global land cover information with existing products. In the third paper, Clinton et al. (2015) investigated decision-level fusion of several land cover classification results to improve global land cover map accuracy. In this paper, they described how stacking of geographically allocated classifications could create a map composite of higher accuracy than any of the individual classifiers. This method has the advantage of not requiring posterior probabilities or level of support for predicted class labels. Both voting algorithms and trainable classifiers with a set of validation data were used to combine individual land cover maps. To demonstrate the technique, they used Landsat based 30-meter land cover maps with globally distributed validation samples to compose the maps and compute accuracy. The results showed that geographic stacking could improve individual map accuracy by up to 6.6%. The voting methods could also achieve higher accuracy than the best of the input classifications. The results were illustrated on a Landsat scene in California, USA. The compositing technique described here has broad applicability in remote sensing based map production and geographic classification. In the fourth paper, Han et al. (2015) presents the design and development of a web-based system for supporting 30-m resolution GLC data production by combining geo-spatial web-service and Computer Support Collaborative Work (CSCW) technology. Based on the analysis of the functional and non-functional requirements from GLC mapping, a three tiers system model is proposed with four major parts, i.e., multisource data resources, data and function services, interactive mapping and production management. This web-based system not only facilitates the integration of heterogeneous data and services required by GLC data production, but also provides online access, visualization and analysis of the images, ancillary data and interim 30 m global land-cover maps. The system further supports online collaborative quality check and verification workflows. It has been successfully applied to China’s 30-m resolution GLC mapping project, and has significantly improved the efficiency of GLC data production and verification. The concepts developed in this study could also benefit other GLC or regional land-cover data production efforts.

- **Global land cover change detection and identification of rates and patterns:** One paper in the theme issue is focused on change detection. Song et al. (2015) demonstrate methods for extending the span of forest-cover change analysis from the Landsat-5 and -7 era (1984 to present) to the previous era covered by the Corona archive in two study areas: one area covered predominantly by urban and sub-urban land uses in the eastern US and another area by tropical forest in central Brazil. Forest/non-forest classification of Corona images was based on texture measures and a SVM classifier while forest/non-forest classification Landsat images were conducted using SVM and Training Data Automation-SVM. Post-classification change detections were then performed based on the resulting thematic dataset. Classification accuracies were approximately 95% and 96% for forest/non-forest discrimination for the temperate urban and tropical forest study areas, respectively. The trajectory of change from the 1960s to 2000s for the two study areas were presented and analyzed. These results demonstrate the potential of Corona data in land cover and change studies. However, using Corona imagery for land cover change studies over larger areas will require more automated georegistration and classification methods. Georeferencing Corona images must be automated to avoid overwhelming human involvement, especially in the identification of ground control points. Automatic training data selection methods, similar to others used for Landsat-based classification, also need to be further developed for regional to global retrievals.

- **Validation of global land cover datasets:** One paper in the theme issue is related to this topic. Tsendbazar et al. (2015) analyzed metadata information from 12 existing and forthcoming GLC reference datasets and assessed their characteristics and potential uses in the context of 4 GLC user groups, i.e., climate modellers requiring data on Essential Climate Variables (ECV), global forest change analysts, the GEO Community of Practice for Global Agricultural Monitoring and GLC map producers. The authors found that suitability of the datasets is highly dependent upon specific applications by the user communities considered. The LC-CCI, GOF-GOLD, FAO-FRA and Geo-Wiki datasets had the broadest applicability for multiple uses. The re-usability of the GLC reference datasets would be greatly enhanced by making them publicly available in an expert framework that guides users on how to use them for specific applications.

- **Global land-cover observation capacity from earth observation satellites and data quality:** Two papers are focused on this topic. In the first paper, Belward and Skøien present a compendium of civilian earth observation satellites with the potential to gather global land-cover observations. They find that, since the launch of Landsat-1 in 1972, government and/or private entities in 33 other sovereign states and geopolitical groups have chosen to finance such missions and 197 individual satellites with a global land-cover observing capacity have been successfully launched. Of these 98 were still operating at the end of 2013. The average number of satellites launched per-year/per-decade has increased from 2 to 12 and spatial resolution increased from around 80 m to less than 1 m multispectral and less than half a meter for panchromatic; synthetic aperture radar resolution has also fallen, from 25 m in the 1970s to 1 m post 2007. More people in more countries have access to data from global land-cover observing spaceborne missions at a greater range of spatial resolutions than ever before. In the second paper, Nagol et al. (2015) examined the seasonal magnitude and latitudinal variation of Bidirectional Reflectance Distribution Function (BRDF) effects on Landsat reflectance and vegetation indices using reflectance and BRDF parameters derived from daily MODIS.
reflec
tive estimates and the sun–target–sensor geometry of the Landsat-5 orbit. They found that bidirectional effects are prevalent in Landsat measurements, but their effect varies by latitude. Seasonal variations in illumination geometry affect reflectance as well as vegetation indices, adding spurious seasonality in otherwise stable evergreen phenology at a lower-latitude tropical forest site. The ±7.5° variation in View Zenith Angle (VZA), usually considered negligible, resulted in as much as 20% along-scan variation in reflectance at the lower latitude site. Although variation in Solar Zenith Angle (SZA) is less pronounced in the tropics than in the upper latitudes, its impact is most significant in the tropics due to the relative proximity of Landsat scans to the principle plane. For this reason, even small variations in viewing and illumination geometry have large impacts on reflectance and vegetation indices near the equator. In contrast, the amplitude of SZA and phenological variations have the greatest impact on reflectance estimates in mid- and upper latitudes, where interaction between phenological variation and Landsat’s 16-day orbital cycle can be greater than the combined effect of BRDF and orbital drift. The analysis suggests that accounting or correction for BRDF effects will increase precision of land cover mapping and monitoring based on Landsat data alone or in combination with other sensors.

We are pleased to offer this theme issue to the scientific community, and hope that it has accomplished our goal of highlighting recent progress and challenges in global land cover mapping. We would like to take this opportunity to thank all authors for submitting their papers to the theme issue, the reviewers for their constructive comments, and the Editor-in-Chief, Prof. Derek Lichti, for his support to this theme issue.

References


