A Review on Mapping Land Cover Change and the Various Techniques Used

Keerti Kulkarni[1], Apoorva Raghunandan[2]

[1]Assisant Professor, Department of ECE, BNM Institute of Technology, Bangalore

[2]Department of ECE, BNM Institute of Technology, Bangalore

Abstract: ***Global land cover types in 2001 and 2010 were mapped at 250 m resolution with the data from multiyear time series Moderate Resolution Imaging Spectrometer (MODIS) data. This was done with the help of data acquired in the preceding and subsequent years. Slope data and geographical coordinates of pixels were also used. The classification was done with the finer resolution observation and monitoring of global land cover (FROM-GLC) project and the results were further improved through post processing. A spatial-temporal consistency model, Maximum a Posteriori Markov Random Fields (MAP-MRF), was first applied to improve land cover classification for 3 consecutive years. The MRF outputs for 2001 and 2010 were then processed with a rule-based label adjustment method with MOD44B, slope and composited EVI series as auxiliary data. With maximum probabilities, the label adjustment process re-labelled the over-classified forests, water bodies and barren lands to alter-native classes. The longest record of global-scale medium spatial resolution earth observation data goes to Landsat data. Therefore, the current methods for large area monitoring of land cover change using medium spatial resolution imagery (10–50 m) use Landsat data. Forest cover change is quantified by large area products. Forests are an easy cover type to map and are the current focus of great environmental concern. Among existing change products, supervised or knowledge-based characterization methods predominate. Radiometric correction methods vary largely as a function of geographic/algorithmic scale. Temporal updating of cover change varies as a function of regional acquisition frequency, cloud cover and seasonality. With the Landsat archive opened for free access to terrain-corrected data, in the future it is very likely that product generation will be more data intensive. Per scene, interactive analyses will no longer be viable. With both free and open access to large data volumes with improved processing power there will be automated image pre-processing and land cover characterization methods. Such methods will need to grasp high-performance computing capabilities in advancing the land cover monitoring discipline. Robust validation efforts will be an essentiality to quantify product accuracies in determining the optimal change characterization methodologies.***

Keywords: MODIS Data, MAP-MRF, Landsat Data, Radiometric correction

1. Introduction

Documenting global land change is more pressing than ever, due to the changing state of global climate, biodiversity, food and fiber demand and other critical environmental/ecosystem services. Accurate global land cover data are not only required for improving performances of ecosystem, hydrological, and climate models at the global level but are also essential to understanding the global spread of diseases from natural causes. [2] The improved data availability, advanced processing methods and the strong need for information on environmental change will lead to an increase in our quantification of global land change in the coming months and years. While each of these three aspects is necessary for realizing this improved monitoring capability, data availability is the most critical of them all. The expansion of land monitoring methods and systems to other regions of the world and to other themes of interest will be accelerated best with a free and easy access data policy for global monitoring systems. In this paper, we review many large area medium spatial resolution land cover monitoring products, some of which employ Landsat data as inputs [1]. Till date, forest change products predominate due to the growing concern of forest change regarding carbon accounting, biodiversity monitoring, and other issues concerning forested landscapes. Another reason forests are the pivot of our study using remotely sensed data is due to the fact, forests are one of the most easily distinguished vegetation cover types when compared to other monitoring targets, such as croplands or urbanized landscapes. The forest monitoring methods included in this review employ digital image processing algorithms that quantify forest extent and change. The review is divided into a series of methodological inter-comparisons which include pre-processing, geographic/algorithmic and temporal scales and the change detection algorithms themselves. Validation is recognized as the key to future determination of not only the quality of individual products, but as a way to assess the aforementioned methodological considerations in generating large area change products. Validation of static large area land cover products is difficult, but change products are even more of a challenge. With the advent of large area land cover change products, the development of rigorous validation protocols is essential. The following review illustrates the wide variety of methods and the lack of consensus in large area monitoring approaches. Validation will be required to quantify the various strengths and weaknesses as the land cover and land cover change mapping community moves forward. [1]

Studies reviewed in this paper include land cover change products that employ digital image processing-based algorithms to quantify cover conversion at national scales or larger, typically greater than 1 Mkm^2 in area. A wide range of methods exist and are outlined in Table 1 for reference. Among the many methods, Landsat is the primary data source and is used in all studies; Indian Remote Sensing Advanced Wide Field Sensor (AWiFS) data are used as one of the inputs along with Landsat data in the NASS Cropland Data Layer (CDL) crop type maps. Forest cover and change is the primary thematic variable, excepting the NASS CDL crop type map (National Agricultural Statistics Service, 2011). Methods are often employed interactively or iteratively in deriving map outputs. Supervised learning algorithms are very beneficial. Most of the land cover products were produced for a single year with different classification algorithms or classification systems. The MODIS 500 m land cover products were generated on a yearly basis by using a supervised boosted decision tree classifier. The 300 m GLOB cover land cover data were produced for 2005 and 2009, respectively, using spectral, temporal and phonological information with an unsupervised classification method.

FROM-GLC data were only based on a single date Landsat image for a particular pixel. There-fore the land cover information produced with the 30m product may not reflect the dominant land cover type of any year. To overcome this problem, coarser resolution time-series have been combined with the 30m data through an approach consisting of image segmentation and random forest classification [2]

2. Product Inter-comparison

2.1 Operational Versus Experimental

The most important distinction between existing large area medium spatial resolution land cover change products is whether they are made in an operational or experimental setting. Standards for product deliverables are more demanding for operational products than experimental ones. Operational products exist as decision support systems for broader policy objectives and require close attention to product consistency and accuracy, along with a practical approach to mapping in meeting product delivery dates. Research based products have a higher tolerance for uncertainty. Operational ones need to meet quality and timeliness standards dictated by the administrative setting in which they operate. [1]

2.2 Pre- Processing

a) Geometrical Corrections

To perform time-series analyses of land cover change, a consistent geometric image becomes important. Historically, geometric rectification has been an hindrance in mass-processing of Landsat data and costly in-house rectification procedures have been employed to ready imagery for analysis. For example, the operational PRODES and NCAS data products use a set of base images to match new image sequences to. The NASS CDL project has purchased rectified imagery, such as the commercial provision of geometrically corrected AWiFS data. The realization that large area monitoring is dependent on consistent geometric corrections has led to the development of standard orthorectified products, like the Global Land Survey (Tucker et al., 2004) data set. The GLS data sets of orthorectified Landsat imagery were produced with global coverage for the 1990, 2000 and 2005 years. The North American LEDAPS disturbance mapping project of Masek et al. (2008) employed the GLS data sets for 1990 and 2000. Consistent image geometry of Landsat scenes across epochs is achieved through the use of an improved digital elevation model and ground control network. The resulting orthorectified, terrain corrected image base is now used for all USGS Landsat processing. Standard orthorectified imagery produced by data providers will be the wave of the future. For example, all data from the forthcoming Landsat Data Continuity Mission (LDCM) will be generated as standard orthorectified imagery processed with the same inputs currently used to generate all other USGS Landsat products, reducing the amount of effort needed by separate projects to create rectified time-series datasets and ensuring geometric consistency. The automatic rectification of imagery is relatively new and will enhance the implementation of more data-intensive multi-temporal studies. [1]

b) Radiometric Processing

A host of radiometric corrections are possible for pre-proc by methods reviewed in this study. While other standard products such as albedo or net primary productivity require surface reflectance inputs. Radiometric corrections for land cover are only required when a given land cover characterization model is to be extrapolated beyond a single image in space or time. This leads to an interesting mix of approaches. The radiometric corrections applied in these studies

c) Geographic/algorithmic scale

Algorithm implementation can be performed per scene, per stratified sub-unit, or over the entire study area. Applying the same rule-base over the entire study area has advantages i.e. it ensures consistency in the characterization across space. However, signature extrapolation limitations can worsen the impact of characterization (Cihlar, 2000). Additionally, spectral confusion between dissimilar cover types affects the quality of the characterization. We solve this issue by stratifying the study area and applying per stratum algorithm runs. The drawback to this is potential inconsistency in the individual models and the creation of seams/ inconsistencies at strata boundaries. Operational methods typically employ a stratified approach, whether per scene as with PRODES or per eco-zone as with NCAS. The NLCD and CI change products were derived per scene. The CDL product uses a state-based stratification that allows for a sub-stratification based on the richness of the calibration data. A per scene characterization of non-forest before applying the disturbance index algorithm within forests was done by The LEDAPS product. The more experimental products (LEDAPS and SDSU) employ a single model/index over the entire study are. the fundamental difference between the operational and research-based products can be observed from variance. However, many other applications, such as pasture-land or natural hazard monitoring require sub-annual monitoring. They may likely be tested given the open Landsat archive such as MODIS which can capture daily images globally, Landsat has a repeat cycle of 16 days. However, only the United States is acquired automatically each overpass. Due to the lack of In addition to limited acquisitions, other limitations such as seasonality and cloud cover also reduce the viability of annual land cover updates.

Many other countries and regions require more data intensive methods for removing clouds, and that may face fundamental data limitations to synoptic annual monitoring. Radiometric processing is often associated with include top of atmosphere reflectance, surface reflectance, bi-directional reflectance distribution/view angle normalization, and terrain normalization. Top of atmosphere reflectance calculations are typically bulk corrections applied to all the pixels in a given image to adjust primarily for sun angle and earth-sun distance. The most mature approach is that of LEDAPS (Masek et al., 2006), which estimates per pixel surface geographic/algorithmic scale Radiometric normalization is required to extrapolate signatures across multiple images either through space or time. In geographic/algorithmic scale, models either operate at the study area scale, or per a population of sub-units. Moreover, a reliance on standard radiometric response is required. These signals are characterized using a fixed model applied to all pixels over space and through time. Next, it relies on training data to model local to state-level radiometric responses in mapping large area land cover. The model attunes itself to local conditions and there is no standard rule relating spectral response to land cover that exists; there are a host of models. [1]

d) Temporal scale

Till now, medium spatial resolution imagery has been used to perform annual or multi-year epochal studies. Sub-annual change quantification has typically not been a requirement for land cover change products. INPE's DETER product is an exception, employing MODIS data to provide a near-real time alarm product in monitoring Legal Amazon deforestation. No Landsat products to date operate in such a manner. For most monitoring, objectives, annual change quantification is sufficient. It can also be stated as the goal of several global and national monitoring objectives (GCOS; U.S. Climate Change Science Program, 2003).

e) Algorithms

Physically-based modelling of land cover has not been demonstrated, and empirical methods are the rule. Traditionally, land cover characterization algorithms have been divided into supervised and unsupervised methods. In large area land cover mapping, which use medium spatial resolution inputs, supervised methods, or slight variations of supervised approaches are favored. All example products reviewed here, except for the LEDAPS and NLCD change products, directly employ training data for model development. In the case of PRODES, an analyst defines the spectral endmembers principally for creating a shade fraction layer that is subsequently input to segmentation and unsupervised clustering algorithms. For NCAS, training data are used with a spectral index, which is derived by reducing the feature space using a canonical variate analysis. To discriminate forest from non-forest, a thresholding approach is applied. Training data and a decision tree algorithm, is the method used for the SDSU and CI products . The NASS CDL uses one of the richest training datasets, labelled agricultural field polygons, also with a decision tree algorithm. The North America Disturbance Index method employed a knowledge-based disturbance index empirically tuned to be a relative measure between mature forest and bare soil. An index for creating a non-forest mask was also developed reflectance. LEDAPS causes scene spectral variations due to atmospheric effects including ozone concentration, column water vapor, elevation and surface pressure and aerosol optical thickness. The LEDAPS North American forest disturbance map employs a full surface correction application based on the MODIS 6S radiative transfer approach (Masek et al., 2006; Vermote et al., 1997).The benefits of removing atmospheric effects from input imagery are obvious, as consistent surface reflectance enables the application of standard models to all scenes, as in the LEDAPS product.[1] Surface-related spectral variations that may require additional correction include bi-directional reflectance distribution function (BRDF) effects (Schaaf et al., 2002). For wide-view angle sensors, these effects can be pronounced and can lead to confusion in analyses. But, even narrower swath medium spatial resolution sensors such as Landsat evidence BRDF effects, largely a function of solar zenith and sensor view angles (Danaher et al., 2001; Toivonen et al., 2006). While BRDF effects are target-dependent, bulk empirical adjustments have been shown to improve radiometric response and cover characterizations (Hansen et al., 2008). Fig. 1 shows an example of this for a scene. The SDSU and NCAS methods employ a per scene BRDF adjustment. As methods increase their use of multi-temporal imagery, and extend studies to multi-decadal time scales, radiometric consistency will only become more important regarding land cover monitoring (Wulder et al., 2008)using Landsat data.

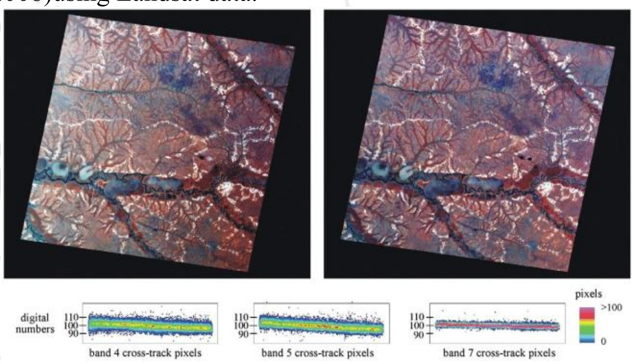

Figure 1

Another approach is the use of spectral vectors or change indices from a reference state, a method employed in the NLCD land cover change layer.. Here a set of indices, including NDVI, were interactively threshold in identifying and mapping change (Xian et al., 2009). An approach worth noting is that of the CDL layers, is in fact not a change detection approach. For this product, annual crop type extent is mapped independently and used in a regression estimator procedure to refine annual crop acreage estimates. In this manner, the CDL derives annual acreage estimates that are directly comparable from year to year.[1]

3. Improving Data Processing

Landsat data dominate the field of national-scale land cover change monitoring to date. This is due to the value of Landsat observation continuity (Wulder et al., 2011) as well as the new free Landsat data access policy (Woodcock et al., 2008). In addition to a free and accessible data policy, global acquisition strategies are needed to ensure consistent national scale and larger mapping efforts. Distributed and costly mapping models that require hosts of analysts will no longer be required. [1] depictions of regional, continental, biome and global scale land cover extent and change. To date, Landsat ETM+ is the only medium spatial resolution sensor to have a global acquisition strategy (Arvidson et al., 2001). It allows for the sharing of methods and experiences from using data from a common sensor. This is important for programs such as the UNFCCC's Reducing Emissions from Deforestation and degradation in Developing countries initiative (REDD+). However, it is incumbent that data policies do not backslide given the increased interest in, for example, carbon accounting (GOFC-GOLD, 2010), and the resulting possible policy initiatives to commoditize imagery.

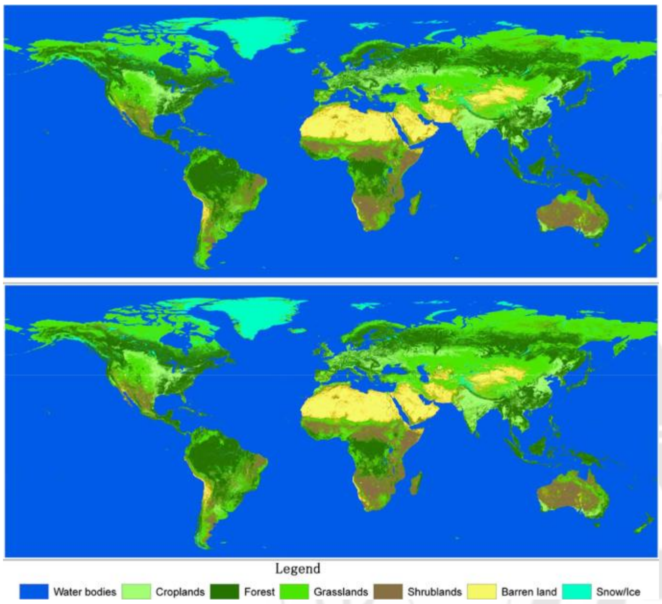

Figure 2: Global mapping results for 2001 (top) and 2010 (bottom)

4. The End of Per Scene Analysis

Three of the more mature methods, all within operational settings, the PRODES, NCAS and NLCD forest or land change products, require significant per scene analyst interaction. This is made possible due to the viability of single-image updates per year, such as largely cloud-free imagery captured during the respective dry seasons within Brazil and Australia. For many other areas, cloud cover is a confounding factor and precludes the use of singe image updates of land cover per year. Per pixel processing methods using multi-date imagery are required for such areas, and are exhibited here in the Congo/Indonesia/ European Russia and CDL examples.

Regardless of current approaches, the future will lie in mass-processing of Landsat or other like data. Single-date comparison methods are a legacy of prohibitive data costs and a lack of automation in geometric and radiometric pre-processing. For many, the capacity to process thousands of images from raw radiometric state to finished land cover extent and change product simply does not exist. The derivation of standard products, is needed. One example of a pre-processed set of user-friendly Landsat data, based on the MODIS processing model (Justice et al., 2002), is the Web-Enabled Landsat Dataset (WELD), which exists for the continental United States and Alaska The advent of Landsat-resolution pre-processed time-series data sets will greatly reduce the costs and logistical complexity of mapping

5. Training Samples

Our training samples are selected from the FROM-GLC project (Gong et al., 2013; Zhao et al., submitted for publication). The training samples were interpreted and cross-checked based on Landsat TM and Google Earth data, and then rechecked twice by high quality image interpreters using EVI time series for 2010 as auxiliary data. The validation samples were preset in a systematic unaligned manner, and then were interpreted and checked by high quality image interpreters. As a final step, one of the best interpreters finalized the validation samples to further improve accuracy and consistency. The original samples are based on Landsat TM at 30 m resolution. The large samples in Gong et al. (2013) and Zhao et al. (submitted for publication) are those with homogeneous areas greater than 500 m 500 m. We used only the large samples in the 250 m MODIS classification. For classes with few large validation samples, large training samples were also used. The spatial distribution of the samples is shown in Fig. 1. Fig. 2 shows the number of samples for each class. The sample numbers of different classes differ significantly. Despite this undesirable proportional distribution it is the best sample currently available. [2]

6. Data and Sample Preparation

6.1 Data Preparation

We first set the cloudy pixels of EVI, RED, NIR, BLUE and MIR time series from MOD13Q1 to missing values, and smoothed the data with a method based on the Savitzky–Golay filter (Chen et al., 2004). Then, time series in each band were aggregated to 32-day averages. The averaging reduces data volume and helps., to improve classification accuracy (Friedl et al., 2010; Hüttich et al., 2011). The DEM data were re-projected, sampled and gridded to the MODIS tiles, so that the resultant data were registered pixel-to-pixel to the tiles of MOD13Q1. Abnormally high values appear at the land–ocean boundary when computing the slope data from the DEM data, these values were replaced with zeroes. For each MOD13Q1 tile, we computed latitude for all pixels and saved them in an input layer. In total, there are 62 input features for a single year, including monthly EVI (12), monthly spectral bands (12 4), slope (1) and latitude (1)

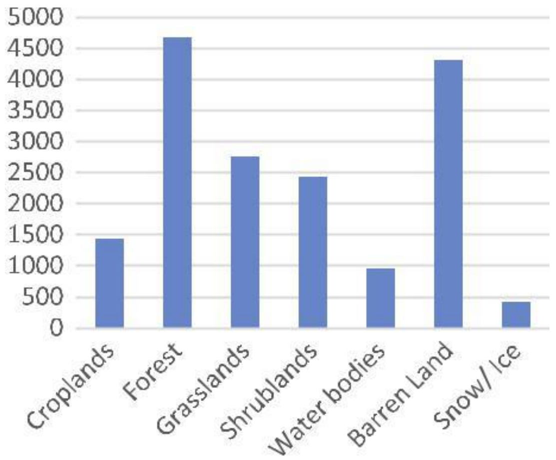

Figure 3 Sample distribution by class

6.2) Sample Change In Time

The original samples were collected for the year 2010. Based on the geographical location, we compared samples in other years. For each of the six years, a random forest (Breiman 2001) classifier of 200 trees was trained, with the slope, the latitude, the monthly EVI and the monthly spectral bands as input features .transition model (Liu et al., 2008), and by counting pixel pairs with-in each tile. [2]

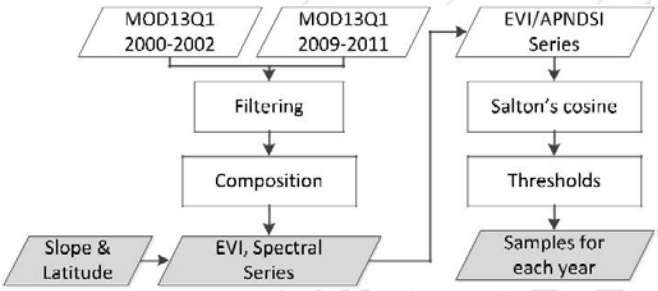

Figure 4: Shows a flowchart of the methodology for data and sample preparation.

7. Land Cover Mapping Method

The trained classifier was then applied to all the pixels to get the class labels and the posterior probabilities. The posterior probability estimated by the random forest classifier indicates a pixel's probability of belonging to a class given the input features. A pixel's class label is determined per the maximum probability value.

The MAP-MRF model was applied to the probabilities of 3 consecutive years to improve the accuracy and consistency of the land cover classification. The model is defined and computed in the same way as in Liu and Cai (2012). Using the Iterated conditional modes (ICM) algorithm (Besag, 1986), the solution can be obtained by iteratively maximizing the conditional posterior probability of each pixel. For a pixel of spatial-temporal position (s, t), the resultant label is computed:

$$L_{(s,t)} = \arg\max_{L_{(s,t)}} \{P(L_{(s,t)} | L_{N(s,t)}) f(X_{(s,t)} | L_{(s,t)})\}$$

where X is the data; L is the class label; $N(s, t)$ is the set of spatial-temporal neighbors for pixel (s, t); f is the marginal likelihood. We mapped the global land cover data into Level 2 classes, and the results were merged into Level 1 classes when needed. With data from adjacent years, we used a spatial-temporal consistency model, Maximum a Posteriori Markov Random Fields (MAP-MRF) (Liu et al., 2006, 2008; Liu and Cai, 2012), to improve the land cover classification. Finally, we performed a fine label adjustment on the output data for 2001 and 2010. Fig. 4 shows the land cover mapping workflow. [2]

Rule Based Label Adjustment

Although input samples are critical to final classification accuracy, it is impossible to include all variations with training samples at the global scale. The MAP-MRF process can improve the accuracy and consistency of land cover maps, but there is still a need for label adjustment based on prior knowledge and/or auxiliary data. For each pixel, the corresponding probabilities estimated by the random forest classifier were used to find the most probable new label when the pixel's label was to be changed. The alternative class with the maximum probability value was selected as the new label. We first filtered the output data with the percent tree cover of the VCF data. A pixel of the VCF data was regarded as forest if the percent tree cover value was in [t, 100], where t is the given thresh-old. The forest/non-forest threshold was derived from comparing the VCF data for 2010 and the FAO global forest cover statistics 2010 (FRA, 2010). The forest area of the VCF data is closest to the FRA 2010 results, 4 billion ha, when the threshold value is set to 21. We selected the value 18, so that the forest area of the filtered land cover mapping in 2010 was close to the FRA 2010. [2]

For 2001 and 2010, if a pixel was classified as forest and the percent tree cover value was less than 18 or greater than 100, then its label would be changed to the second most likely class after forests. Finally, we filtered barren lands (Level 2 classes) with EVI data. If the average of the two highest EVI values was greater than a given threshold, we would compare the probability values of grasslands (Level 2 classes) and shrub lands (Level 2 classes). If the maximum probability value was also greater than a given threshold, the label would be changed to the corresponding class. The thresholds were set manually tile-by-tile with visual interpretation.

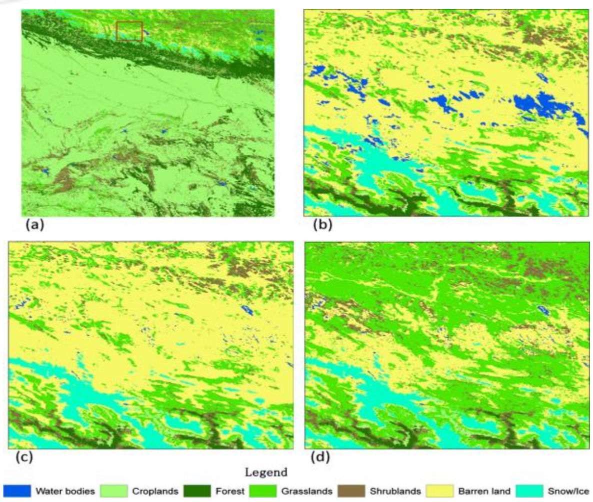

Figure 5 Regional results for the year 2010 in North China Plain. (a) Land cover in MODIS tile h27v05, the red box

indicates the area shown in b–d, (b) output from random forest, (c) Output from MAP-MRF, and (d) output from label adjustment.

Pixels with a large proportion of shadows located in mountain areas were often classified as water bodies. We filtered all water bodies with the slope data. If a pixel was labeled as water body and the corresponding slope was greater than a given threshold, it would be relabeled to the second most likely class after water bodies. If the corresponding VCF value was not between 18 and 100, forest would not be chosen as a substitute to water bodies. The threshold of slope was set manually tile-by-tile with visual interpretation. All threshold values are around 10 degrees, and the fine-tuning was performed to reduce errors between misclassified water bodies and real water bodies.

The MAP-MRF algorithm is a process for optimizing land cover classification with information from spatial-temporal neighbourhoods. [2]

These were not corrected by the MAP-MRF model, as shown in Fig. 8b. After water filtering, most incorrectly classified water bodies were eliminated. However, a small fraction of incorrectly classified water pixels still exists. These pixels contain large proportions of shadows, but their corresponding slope values are low.

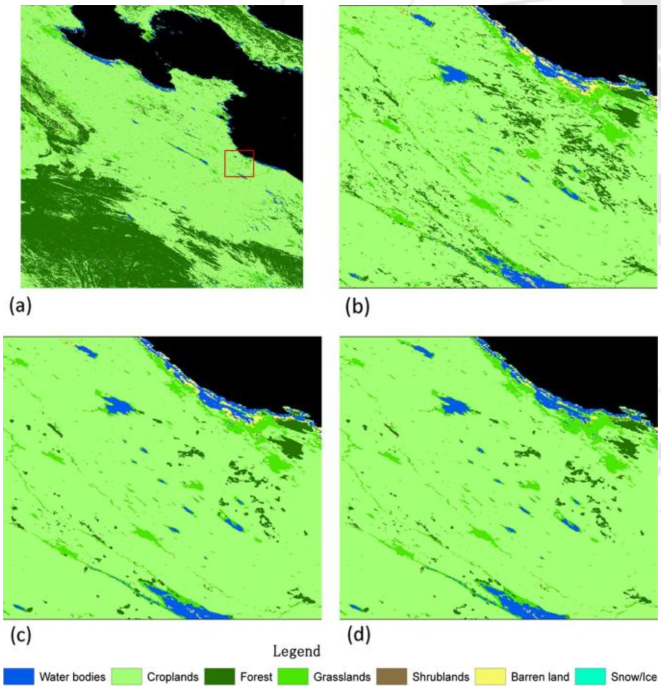

Figure 6. Regional results for the year 2010 in Tibet. (a) and cover of a MODIS tile with identifier h25v06, red box indicates the range of the sample region, (b) output from MAPMRF,

The final results Fig. 8d were obtained from barren lands filtering. Grasslands and shrub lands with weak vegetation signals were separated from the barren lands classification with information from spatial-temporal neighbor-hoods. Although the spatial consistency is less significant at 250 m resolution than those at finer resolutions, the temporal neighbors can still provide important information. Thus the classification accuracy can be improved with the MAP-MRF algorithm. Fig. 7 shows regional results for the year 2010. The region is located in the North China Plain, as indicated in Fig. 7a (MODIS tile h27v05) with a red box. Croplands are the main land cover type in this region. Some over-classified forests were correctly relabeled as croplands by represent the percentage of cover types in the relabeled barren lands pixels. The pixels with a large proportion of shadows located in large mountain areas are mostly barren lands. There are also some grasslands and shrub lands with weak vegetation signals, and some snow/ice in the high-altitude areas. Training samples are always expensive to obtain for global land cover mapping and an independent validation sample set requires a tremendous amount of time to acquire. Friedl et al. (2010) measured the map accuracy by performing a 10-fold cross validation with an effort to avoid spatial autocorrelation. Their samples were from thousands of polygons. Our samples were independent points. With the random forest classifier, we estimated the map accuracy with a stratified 10-fold cross validation directly. Table 3 is the confusion matrix for 2001 and Table 4 for 2010. Croplands, grasslands and validation for 2000, 2002, 2009, and 2011. The overall accuracy values are 74.95%, 75.13%, 75.25%, and 74.62%, respectively. The out-of-bag (OOB) error can be estimated as an accuracy measurement when training the random forest with all samples (Breiman, 2001). The OOB error estimated for Level 2 classes is 0.3769 for 2001, and 0.3781 for 2010.

The overall accuracy for 2010 estimated for Level 2 classes is 62.48%. Mixed forests tend to be confused with other Level 2 classes of forests. Seasonal croplands exhibit low producer's accuracy, and they were mainly misclassified to other croplands. Dry lake/ river bottoms were mainly misclassified to dry salt flats and ex-posed bare rock. With 2010 data, we also assessed the contribution of each step of post-processing to the improvement of image classification. We used 80% of the total samples for training and the remaining 20% for testing. The overall accuracy was 74.30% for outputs directly from random forest and 82.07% for outputs from MAP-MRF. The inclusion of spatial–temporal contextual information has clearly increased the overall classification accuracy. However, shrub lands tend to be confused with each other, and barren lands and water bodies have high accuracies. [2]

8. Improving Data Processing and Characterization

As illustrated in this review, there are a variety of methods currently in use in support of either operational or research-based large area monitoring of land cover change at medium spatial resolutions. The advent of both the open Landsat archive and improved high-performance computing capabilities will lead to dramatic changes in current methods. Past approaches that rely on bi-temporal image comparisons, i.e. single-date time 1and time 2 analyses, as per the PRODES, NCAS, NLCS, LEDAPS and CI methods, will be replaced by more exhaustive and data intensive methods that employ time-series data sets (Huang et al., 2009; Kennedy et al., 2009). [1] Such methods will use all appropriate available data in characterizing change. The SDSU approach uses all available data in the Landsat EROS archive based on a cloud cover metadata threshold. The CDL product combines AWiFS and Landsat ETM+ and TM data in characterizing annual crop type extent. Once these methods are sufficiently mature and validated, their adoption for national monitoring purposes will follow. In effect, countries that have been limited by past data policies and processing limitations will be able to 'leapfrog' directly to more data intensive methods. The development of user-friendly versions of such methods is required, however. Efforts in support of initiatives such as REDD+ should include the creation of pre-processed inputs, such as the WELD data, as well as algorithmic interfaces that enable the appropriate country agencies to exploit image archives in more automated manners.

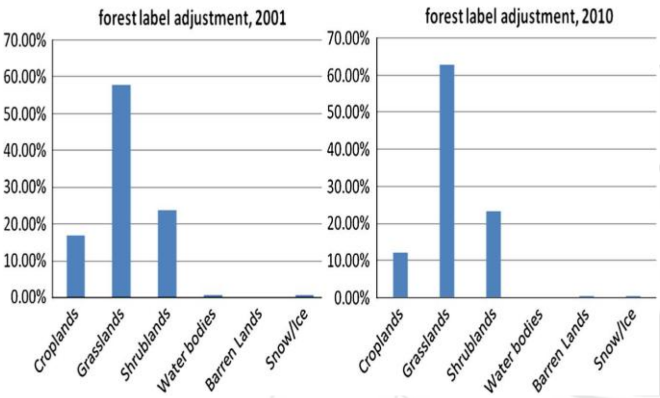

Figure 7: Classes relabelled from forests after forest filtering for 2001 (left) and 2010 (right).

The generation of standard products as inputs will continue to evolve. The creation of input data sets that include geometric, atmospheric, BRDF and results derived directly from supervised classification. We also performed stratified cross using the MAP-MRF model and the label adjustment further refined some details. Globally, over-classified forests were changed to grasslands, shrub lands, croplands and other classes by label adjustment. By forest filtering, about 8,926,510 km^2 of forests for 2010 and 10,129,466 km^2 for 2001 were relabeled to other classes. Fig. 9 shows the percentage of each class in relabeled pixels from forests after forest filtering. Grasslands, shrub lands and croplands are easily confused with forests, and the bar plots for 2001 and 2010 are similar in shape.

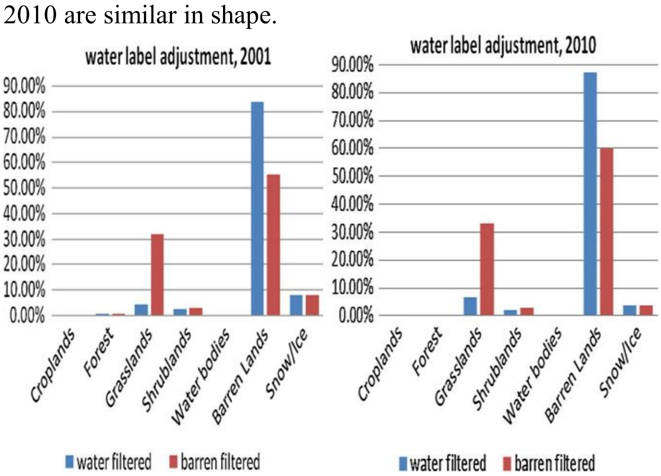

Figure 8: Classes re-labelled from water in water filtering for 2001 (left) and 2010 (right). The blue bars are percentage of cover types relabelled after water filtering, and the red bars are the results after some barren lands are relabelled.

About 84,980 km^2 of water bodies for 2010, and 120,230 km^2 for 2001 were identified as other cover types with a large proportion of shadows. Fig. 10 shows the percentage of each class relabeled from water. The blue bars represent the percentage of cover types in the relabeled water pixels, and the red bars represent the percentage of cover types in the re-labeled barren lands pixels. [2] Topographic corrections is an important objective. However, each step of processing should be evaluated in quantifying the proposed improvements to input data sets. Ju et al. (this issue) perform such a comparison of a proposed WELD atmospheric correction methodology with that of LEDAPS for Landsat processing. Additionally, multi-sensor cross-calibration will become an important focus of research as more data sets from more sensors are used in ensemble approaches. [1]

Part of the evolution of processing will include cloud-computing, such as that currently being developed by NASA in the form of the NASA Earth Exchange (NEX) (Nemani et al., 2011) and by Google in the form of the Google Earth Engine (GEE) (http://earthengine. googlelabs.com). Both systems rely on parallel processing to reduce computing time in the derivation of remotely-sensed earth system science products. While none of the products reviewed here used cloud computing, such processing systems will be more accessible to research and operational users in the future, improving product iteration/analysis capabilities and reducing the latency of product delivery. By partnering with these or other such computing facilities, agencies or research institutes interested in large area land cover and change mapping will be able to focus on characterization methods while avoiding the overhead associated with developing and maintaining high-performance computing infrastructures. Alternatively, computing costs will continue to decline, enabling users to build their own high-performance computing infrastructures.

9. Advancing Thematic Outputs

The large area land cover change products reviewed here capture land cover conversion processes, largely due to the comparative ease with which such dynamics can be detected and mapped. Other more subtle disturbance dynamics, such as within-cover modifications, are considerably more challenging to quantify. Forest recovery and degradation are two such examples. While selective logging has been quantified for the major logging states of the Brazilian Amazon by Asner et al. (2005) and such methods are being improved (Souza & Roberts 2005), the level of maturity for operational implementation at national scales and larger has not yet been realized. The admitted difficulty in capturing such subtle cover changes consistently through time and across space does not lessen their importance. Many land use dynamics occur in the context of cover modification, not conversion, and are critical to quantifying many earth system dynamics (Ramankutty et al., 2006). [1]

10. Completing the Lands at Archive

Finally, for global studies, there is a significant need to improve the amount of Landsat coverage available for use in land change studies. For environment, technical, and programmatic reasons, the geographic and temporal coverage of global Landsat data is uneven. A significant amount of unique Landsat coverage dating from 1972 to the present is found in international archives and are either not accessible or are in formats inconsistent with those distributed by the USGS. The USGS is working with its network of International Cooperators to collect and consolidate unique historical Landsat coverage in the USGS Landsat archive. Preliminary analysis suggests that an additional 2.5–3.0 million unique Landsat scenes may be available. In the future, improved global coverage will come about through acquisition capacity improvements planned for Landsat Data Continuity Mission and other follow-on Landsat missions, and through international collaboration with other compatible medium resolution missions, such as the European Sentinel-2 mission. [1]

11. Conclusion

Currently, there is a wide variety of methods employed in large area land cover change characterization. While much of the methodological variation described here will persist, future methods will evolve and adapt to greater data volumes and processing capabilities. The land cover change monitoring community is poised for a dramatic increase in characterization capabilities, due to the new Landsat data policy and concurrent advances in high-performance computing. Heritage change mapping methods relying on analyst interaction with individual scenes should decline over time given the improved ability to process and characterize rich time-series of medium spatial resolution data. However, such methods will be tested against institutional requirements for thematic accuracy. Near-term research objectives will require robust validation data sets in establishing which data-intensive methods are most appropriate in quantifying large area land cover change. [1] A MAP-MRF model was used to get consistent results with the auxiliary data from the preceding and subsequent years of 2001 and 2010.

Using MOD13Q1 data for 2000, 2001, 2002, 2009, 2010 and 2011, slope data and latitude coordinates as inputs, we mapped the global land cover for 2001 and 2010 at 250 m resolution. Samples for each year were those unchanged samples and topography. This algorithm can be applied to any year when MODIS time series data of three consecutive years are available. [2]

13) References

[1] Hansen MC and Loveland TR. A review of large area monitoring of land cover change using Landsat data. Remote Sens Environ

[2] Mapping Land cover in 2001 and 2010 with spatial temporal consistency at 250m resolution Jie Wang a, Yuanyuan Zhao b, Congcong Li c, Le Yub, Desheng Liu d, Peng Gong.

[3] Arino, O., Bicheron, P., Achard, F., Latham, J., Witt, R., Weber, J.L., 2008. GLOBCOVER The most detailed portrait of Earth. ESA Bulletin–European Space Agency 136, 24–31.

[4] Bartholomé, E., Belward, A.S., 2005. GLC2000: a new approach to global land cover mapping from Earth observation data. Int. J. Remote Sensing 26, 1959–1977.

[5] Besag, J., 1986. On the statistical analysis of dirty pictures. J. R. Statist. Soc. Ser. B (Methodological), 259–302.

[6] Bontemps, S., Defourney, P., Van Bogaert, E., Arino, O., 2010. GLOBCOVER2009 products description and validation report.

[7] Breiman, L., 2001. Random forests. Mach. Learn. 45,5–32. Chen, J., Jönsson, P., Tamura, M., Gu, Z., Matsushita, B., Eklundh, L., 2004. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter. Remote Sensing Environ. 91 (3), 332–344.

[8] DiMiceli, C.M., Carroll, M.L., Sohlberg, R.A., Huang, C., Hansen, M.C., Townshend, J.R.G., 2011. Annual Global Automated MODIS Vegetation Continuous Fields (MOD44B) at 250 m Spatial Resolution for Data Years Beginning Day 65, 2000– 2010, Collection 5 Percent Tree Cover. University of Maryland, College Park, MD, USA.

[9] Egghe, L., Leydesdorff, L., 2009. The relation between Pearson's correlation coefficient r and Salton's cosine measure. J. Am. Soc. Inform. Sci. Technol. 60, 1027–1036.

[10] FRA 2010. Global Forest Resources Assessment 2010 (Food and agriculture organization of the United Nations, 2010). <http://www.fao.org/forestry/fra/fra2010/en/>.

[11] 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 Sensing Environ. 83, 287–302.

[12] Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., Huang, X., 2010. MODIS Collection 5 global land cover: algorithm refinements and characterization of new datasets. Remote Sensing Environ. 114, 168–182.

[13] Gong, P., Howarth, P.J., 1990. The use of structural information for improving land-cover classification accuracies at the rural-urban fringe. Photogrametric Eng. Remote Sensing 56 (1), 67–73.

[14] Gong, P., Howarth, P.J., 1992. Frequency-based contextual classification and grey-level vector reduction for land-use identification. Photogrammetric Eng. Remote Sensing 58 (4), 423–437.

[15] 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., Ciri, C., Clinton, N., Zhu, Z., Chen, J., Chen, J., 2013. Finer Resolution Observation and Monitoring of Global Land Cover: First Mapping Results with Landsat TM and ETM+ Data. Int. J. Remote Sensing 34, 2607–2654.

[16] Hansen, M.C., Defries, R.S., Townshend, J.R.G., Sohlberg, R., 2000. Global land cover classification at 1km spatial resolution using a classification tree approach. Int. J. Remote Sensing 21, 1331–1364.

[17] Homer, C., Dewitz, J., Fry, J., Coan, M., Hossain, N., Larson, C., Herold, N., McKerrow, A., VanDriel, J., Wickham, J., 2007. Completion of the 2001 national land cover database for the counterminous United States. Photogrammetric Eng. Remote Sensing 73, 337–341.

[18] Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., Ferreira, L.G., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sensing Environ. 83, 195–213.

[19] Hüttich, C., Herold, M., Wegmann, M., Cord, A., Strohbach, B., Schmullius, C., Dech, S., 2011. Assessing effects of temporal compositing and varying observation periods for large-area land-cover mapping in semi-arid ecosystems: Implications for global monitoring. Remote Sensing Environ. 115, 2445–2459 Jarvis, A., Reuter, H.I., Nels on, A., Guevara, E., 2008. Hole-filled SRTM for the globe Version 4, available from the CGIAR-CSI SRTM 90 m Database.

[20] Liang, L., Xu, B., Chen, Y.L., Liu, Y., Cao, W.C., Fang, L.Q., Feng, L.M., Goodchild, M.F., Gong, P., 2010. Combining spatial-temporal and phylogenetic analysis approaches for improved understanding onglobal H5N1 transmission. PLoS ONE 5 (10), e13575. http://dx.doi.org/10.1371/journal.pone.0013575.

[21] Liu, D., Kelly, M., Gong, P., 2006. A spatial–temporal approach to monitoring forest disease spread using multi-temporal high spatial resolution imagery. Remote Sensing Environ. 101, 167–180.

[22] Liu, D., Song, K., Townshend, J.R., Gong, P., 2008. Using local transition probability models in Markov random fields for forest change detection. Remote Sensing Environ. 112, 2222–2231.

[23] Liu, D., Cai, S., 2012. A spatial-temporal modeling approach to reconstructing land-cover change trajectories from multi-temporal satellite imagery. Ann. Assoc. Am. Geogr. 102, 1329–1347.

[24] Loveland, T.R., Reed, B.C., Brown, J.F., Ohlen, D.O., Zhu, Z., Yang, L., Merchant, J.W., 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. Int. J. Remote Sensing 21, 1303–1330

[25] Thenkabail, P.S., Biradar, C.M., Noojipady, P., Dheeravath, V., Li, Y.J., Velpuri, M., Gumma, M., Reddy, G.P.O., Turral, H., Cai, X.L., Vithanage, J., Schull, M., Dutta, R., 2009. Global irrigated area map (GIAM), derived from remote sensing, for the end of the last millenium. Int. J. Remote Sensing 30, 3679–3733

[26] Sexton, Feng, M., Narasimhan, R., Kim, D., Song, K., Song, D.X., Song, X.P., Noojipady, P., Tan, B., Hansen, M.C., Li, M.X., Wolfe, R.E., 2012. Global characterization and monitoring of forest cover using Landsat data: opportunities and challenges. Int. J. Digital Earth 5, 373–397.

[27] Coppin, P., Jonckheere, I., Nackaerts, K., & Muys, B. (2004). Digital change detection methods in ecosystem monitoring: A review. International Journal of Remote Sensing, 25, 1565‒1596.

[28] Dymond, J. R., & Shepherd, J. D. (1999). Correction of the Topographic Effect in Remote Sensing. IEEE Transaction on Geoscience and Remote Sensing, 37, 2618‒2620.

[29] European Space Agency (2009). http://www.esa.int/esaEO/SEMXK570A2G_index_0.html GCOS, Technical Document No. 1143. Geneva: World Meteorological Organization. Global Forest Survey of India

[30] (2008). State of the forest report 2005. Dehradun, India:GOFC-GOLD (2010). A sourcebook of methods and procedures for monitoring and reporting anthropogenic greenhouse gas emissions and removals caused by deforestation, gains and losses of carbon stocks in forests remaining forests, and forestation, GOFC-GOLD Report version COP16-1. Edmonton, Alberta, Canada: GOFC-GOLD Project Office, Natural Resources Canada 203 pp.

[31] Government of Indonesia/World Bank (2000). Deforestation in Indonesia: A review of the situation in 1999. Jakarta: Government of Indonesia/Work Bank.

[32] Hansen, M. C., DeFries, R. S., Townshend, J. R. G.,Carroll, M., Dimiceli, C., & Sohlberg, R. A.(2003). Global percent tree cover at a spatial resolution of 500 meters: First results of the MODIS vegetation, continuous fields algorithm. Earth Interactions, 7(10)

[33] Hansen, M. C., Egorov, A., Roy, D. P., Potopov, P., Ju, J., Turubanova, S., et al. (2011). Con- tinuous fields of land cover for the conterminous United States using Landsat data: First results from the Web-Enabled Landsat Data (WELD) project. Remote Sensing.

[34] Hansen, M., Loveland, T., Quirk, B., Stensas, G., & Christopherson, J. (2006). A constella-tion of mixed-orbit micro-satellites for monitoring global land change and ecosystem dynamics. SDSU EROS Decadal Study Report.

[35] Hansen, M. C., Stehman, S. V., & Potapov, P. V. (2010). Quantification of global gross forest cover loss. Proceedings of the National Academy of Sciences, 107, 8650‒8655.

[36] Harper, G. J., Steininger, M. K., Tucker, C. J., Juhn, D., & Hawkins, F. (2007). Fifty years of deforestation and forest fragmentation in Madagascar. Environmental Conservation, 34, 1‒9.

[37] Holben, B. N., & Justice, C. O. (1980). The topographic effects on spectral response from nadir-point sensors. Photogrammetric Engineering and Remote Sensing, 46, 1191‒1200.

[38] Huang, C., Goward, S. N., Masek, J. G., Thomas, N., Zhu, Z., & Vogelmann, J. E. (2009).An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. Remote Sensing of Environment, 114,183‒198.

[39] Justice, C. O., Townshend, J. R. G., Vermote, E. F., Masuoka, E., Wolfe, R. E., Saleous, N., et al. (2002). An overview of MODIS Land data processing and product status. Remote Sensing of Environment, 83, 3‒15.

[40] Kennedy, R. E., Yang, Z., & Cohen, W. B. (2009).

Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr Temporal segmentation algorithms. Remote Sensing of Environment, 114, 2897–2910.

[41] Khorram, S. (1999). Accuracy assessment of remote sensing-derived change detection.

[42] Bethesda, Md: American Society for Photogrammetry and Remote Sensing Mono-graph Series 65 pp..

[43] National Agricultural Statistics Service (2011). Cropland data layer. Washington, DC: U.S. Department of Agriculture. www.nass.usda.gov/research/Cropland/SARS1a.htm Nemani, R., Votava, P., Michaelis, A., Melton, F., & Milesi, C. (2011). Collaborative super- computing for global change science. EOS Transactions, 92, 109–110.

[44] Potapov, P. V., Turubanov, S. A., Hansen, M. C., Adusei, B., Broich, M., Alstatt, A., Mane, L., & Justice, C. O. (2012). Quantifying forest cover loss in Democratic Republic of the Congo, 2000–2010, with Landsat ETM + data. Remote Sensing of Environment, 122, 106–116 (this issue).

[45] Ramankutty, N., et al. (2006). In E. Lambin, & H. Geist (Eds.), Land use and land cover change: Local processes, global impacts (pp. 9–39). Berlin: Springer.

[46] Roy, D. P., Ju, J., Kline, K., Scaramuzza, P. L., Kovalskyy, V., Hansen, M. C., et al. (2010). Web-enabled Landsat Data (WELD): Landsat ETM+ composited mosaics of the conterminous United States. Remote Sensing of Environment, 114, 35–49.

[47] Sandmeier, S., & Itten, K. I. (1997). A physically-based model to correct atmospheric and illumination effects in optical satellite data of a rugged terrain. IEEE Transactions on Geoscience and Remote Sensing, 35, 708–717.

[48] Schaaf, C., Gao, F., Strahler, A., Lucht, W., Li, X., Tsang, T., et al. (2002). First operational BRDF, albedo and nadir reflectance products from MODIS. Remote Sensing of Envi-ronment, 83(1and2), 135–148.

[49] Shimabukuro, Y. E., Batista, G. T., Mello, E. M. K., Moreira, J. C., & Duarte, V. (1998). Using shade fraction image segmentation to evaluate deforestation in Landsat Thematic Mapper images of the Amazon region. International Journal of Remote Sensing, 19, 535–541.

[50] Souza, C., & Roberts, D. (2005). Mapping forest degradation in the Amazon region with Ikonos images. International Journal of Remote Sensing, 26, 425–429. Strahler, A. H., Boschetti, L., Foody, G. M., Friedl, M. A., Hansen, M. C., Herold, M., et al. (2006). Global land cover validation: Recommendations for evaluation and accuracy

[51] Teillet, P. M., Guindon, B., & Goodenough, D. G. (1982). On the slope-aspect correction of multispectral scanner data. Canadian Journal of Remote Sensing, 8, 84–106.

[52] Toivonen, T., Kalliola, R., Ruokolainen, K., & Malik, R. N. (2006). Across-path DN gradient in Landsat TM imagery of Amazonian forests: A challenge for image interpretation and mosaicking. Remote Sensing of Environment, 100, 550–562.

[53] Woodcock, C. E., Allen, A. A., Anderson, M., Belward, A. S., Bindschadler, R., Cohen, W. B., et al. (2008). Free access to Landsat imagery. Science, 320, 1011.

[54] Wulder, M. A., White, J. C., Gillis, M. D., Walsworth, N., Hansen, M. C., & Potapov, P (2009). Multi-scale satellite and spatial information and analysis framework in support of a large-area fores t monitoring and inventory update. Environmental

[55] Remote Sensing of Environment, 112, 955–969.Wulder, M. A., White, J. C., Masek, J. G., Dwyer, J., & Roy, D. P. (2011). Continuity of Landsat observations: Short term considerations. Remote Sensing of Environment, 115, 747–751.

[56] Potapov, P., Turubanova, S., & Hansen, M. C. (2011). Regional-scale boreal forest cover and change mapping using Landsat data composites for European Russia. Remote Sensing of Environment, 115, 548–561

Author Profile

Keerti Kulkarni received her B.E degree from the university of Bombay and M.Tech degree from VTU. She is currently pursuing her P.hD in VTU and is working as Assistant Professor at BNMIT.

Apoorva Raghunandan is a final year B.E student at BNMIT. Her research interests include Digital Image Processing, Embedded Systems and Environmental Sciences.