# MONITORING FOREST COVER CHANGE IN BOREAL FORESTS: A METHODOLOGICAL APPROACH

*Ouns Kissiyar[1,2], Svyatoslav S. Bartalev[1], and Frédéric Achard[1]*

1. Institute for Environment and Sustainability, Global Environment Monitoring Unit, Joint Research Centre of the European Commission, TP 440, 21027 Ispra (VA), Italy; ouns.kissiyar(at)jrc.ec.europa.eu
2. AGIV, Koningin Maria Hendrikaplein 70 bus 110, 9000 Gent, Belgium

## ABSTRACT

The purpose of this study is to develop a monitoring tool for boreal forest cover change on continental level at high resolution. The system is based on Landsat satellite imagery and has been implemented for the period 1990-2000-2010. For the identification and classification of the forest cover within a large amount of satellite imagery, a robust methodological approach combining multi-date image segmentation and cluster based supervised automated classification was chosen. Thus, an object based, automatic classification method with a regional expert validation are combined to produce regional scale land cover statistics over Russia and Mongolia. High resolution satellite imagery is used to accurately estimate land cover and land cover change for the epochs 1990-2000-2010. The overall method consists of four distinct steps: (i) automatic image preprocessing and pre-interpretation, (ii) validation by regional expert, (iii) statistic computation and (iv) accuracy assessment. The automated procedures have as main objective to unequivocally identify the objects so as to maximally reduce the post-classification interventions of manual procedures and of visual interpretation. A total of 14 different land cover classes are defined in the legend. Given the focus on forests, special attention was devoted to the differentiation of 8 different forest cover types, going up to species level.

## INTRODUCTION

This study was set up as a compliment to the European Commission's Joint Research Centre efforts to monitor tropical forest cover change at continental level based on satellite imagery. The results will contribute to the reduction of uncertainties in estimating forest cover change and related carbon emissions. Such information is desired in the context of the ongoing international initiatives and the existing convention on climate change.

Boreal forests represent a critical component of the global carbon cycle, storing the largest amount of carbon of all global biomes (1). Additionally 75% of the boreal forests are situated in Russia (2). Within a broader context Russia accounts for 48.6% of total area and 51.1% of the total forest stock of the boreal and temperate zones (3).

On a global scale Russia accounts for 20% of the world forests. Forests occupy over half of the land surface of the country. The FAO expects the forest land to expand from 882 million ha in 2010 by 0.9 to 1.15% by 2030. This would mainly happen due to reforestation of abandoned and inconvenient agricultural land and as a result of forest expansion on non-forested land and Tundra. (4)

The objectives of this study were to design an appropriately automatic classification method for the high resolution mapping of forestry related land cover change in Russia and Mongolia during the reference years 1990, 2000 and 2010.

Given the geographical and temporal extent of the area under study the method would ideally rely heavily on automatic processing and classification. However, the expected variations in vegetation state, imagery acquisition conditions and sensor characteristics suggest that a purely automatic classification procedure will not yield the desired uniformity of accuracy levels across the selected test sites.

DOI: 10.12760/02-2014-1-15

## METHODS

### Data

We adapted a global sampling strategy that concentrates on sample sites of 20 km by 20 km systematically located at each integer confluence of the geographic grid (e.g. 50°N, 100°E). In the tropical regions this covers approximately 4% of the land surface and is a statistically representative sample set (5). Due to the convergence of the earth's meridians towards the poles a subset of this sampling rate was taken. The subset approximates the 4% coverage of the land surface at increasing latitudes towards the poles. A total of 1470 sample sites are thus identified. From this sampling a random selection of 502 sites (or approximately 34%) was chosen for this study.

For all sample units image subsets were extracted from Landsat Thematic Mapper (TM) for 1990 and 2010, or Enhanced Thematic Mapper (ETM+) for the year 2000. The images were acquired as close as possible to these reference years. For each sample unit intensive screening for good quality (haze and cloud free) imagery has been performed. Most data were obtained from the United States Geological Survey's (USGS) National Center for Earth Resources Observation and Science (http://glovis.usgs.gov) at full spatial resolution (30 m). Additional images were provided by the CEOS (Committee on Earth Observation Satellites) Land Surface Imaging Constellation program (LSI), mainly for filling remaining gaps due to missing or low quality data (6). The six reflective spectral bands (1-5, 7) of Landsat TM and ETM+ images were subjected to a pre-processing chain that included the conversion to top of atmosphere reflectance (ToAR), haze correction and relative normalization, and cloud masking (7).

An implementation scheme of the applied classification approach is given in Figure 1.

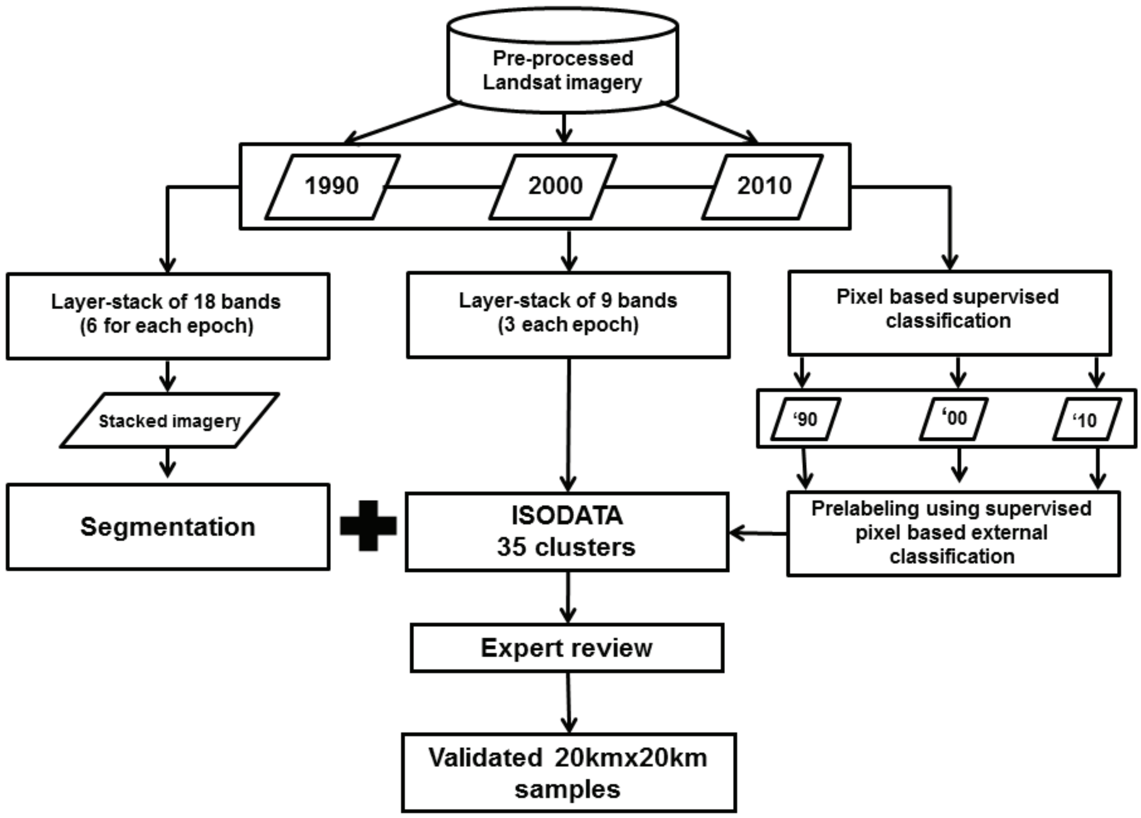


*Figure 1: Classification scheme.*

### Image preprocessing

The satellite data was carefully preprocessed, this comprised: (i) coregistering of the multitemporal images, (ii) calibrating the radiometric data, (iii) masking the clouds and shadow and (iv) correcting the haze (8). Once the data were preprocessed they were stored in a central image repository.

### Image segmentation and automatic pre-labeling

The next step in the development of the processing chain consisted of the creation of a single automatic segmentation approach, which would cope with different landscape patterns and deliver 'robust' objects across the whole region under study.

The available segmentation algorithms have been tested by several authors (9). Based on those findings about technical performance and the visual assessment of the object delineation, the eCognition software (Trimble©) was confirmed to be the most suited for our specific purpose. In particular, the software offers capacity for processing large amounts of data and for classifying objects in one common processing chain. The segmentation algorithm is based on 'region-merging' (10), starting with each pixel forming one image object or region and then merging objects according to user-defined local homogeneity criteria (spectral homogeneity and object shape).

An automated multi-date segmentation has been applied for object delineation. This multi-date segmentation using 18 bands (three Landsat images consisting of bands 1 to 6) is used for each site. This approach, defining the objects in one single operation from the whole set of spectral bands, relies on spatial, spectral and temporal information to delineate the objects.

### Image classification

Image classification is the process of translating pixel values into user defined classes. Over the years several methods have been developed to do this (11,12).

Despite vigorous and often creative efforts to establish new image classification algorithms it was found (12) that the classification results did not improve significantly over a 15 year period. It has since been reported that the use of a combined supervised and unsupervised classification approach produces better results (13,14). Faced with a very extensive study area comprising a wide variety of land cover it was decided to apply such a hybrid classification approach.

### Class definition

For the purposes of our study we defined 14 land cover classes. Given the focus on forests, special attention was devoted to the differentiation of 8 different forest cover types. The complete legend is described in Table 1.

*Table 1: classification legend.*

| | Name of land cover class | Description with dominant/subdominant species |
|---|---|---|
| 1 | Evergreen Dark Needleleaf Forest (ED) | Forest ecosystems consisting of spruce (Picea), fir (Abies) and Siberian pine (Pinus Sibirica) for at least 80% of the forest canopy |
| 2 | Evergreen Light Needleleaf Forest (EL) | Forest ecosystems consisting of pine (Pinus Sylvestris) for at least 80% of the forest canopy |
| 3 | Deciduous Broadleaf Forest (Brd) | Forest ecosystems consisting of birch (Betula), aspen (Populus tremula), oak (Quercus), linden (Tilia), ash (Fraxinus), maple (Acer) and some other deciduous broadleaf tree species for at least 80% of the forest canopy |
| 4 | Mixed Needleleaf Majority Forest (MNM) | Forest ecosystems consisting of the evergreen needleleaf tree species for 60% to 80% and the deciduous broadleaf tree species for 20% to 40% of the forest canopy |
| 5 | Mixed Forest (M) | Proportions of the evergreen needleleaf and the deciduous broadleaf tree species in the forest canopy are approximately equal (40% to 60%) |
| 6 | Mixed Broadleaf Majority Forest (MBM) | Forest ecosystems consisting of the deciduous broadleaf species for 60% to 80% and the evergreen needleleaf species for 20% to 40% of the forest canopy |
| 7 | Deciduous Needleleaf Forest (DN) | Forest ecosystems consisting of larch (Larix) for at least 80% of the forest canopy |

| | | |
|---|---|---|
| 8 | Sparse Deciduous Needleleaf Forest (SD) | Single trees or open stands of larch (Larix) with density of canopy below 20% |
| 9 | Peatlands (Ptlnd) | Permanent mixture of water and vegetation: Sphagnum moss and lichens, or rushes and sedges are dominant. Sometimes sparse tree canopy (up to 20%) can be found. |
| 10 | Recent Burns | Burn scars <5 years old. May contain dead trees, some pioneer vegetation types may be present |
| 11 | Water Bodies | Open water bodies including seas, lakes, reservoirs and rivers |
| 12 | Other land | Lands having a vegetation canopy coverage smaller than 20% |
| 13 | Shrubs | Shrublands or low trees (height is less than 5 m more than 50 cm) |
| 14 | No data | Missing pixels |

**Unsupervised ISODATA classification**

The unsupervised Iterative Self-Organizing Data Analysis Technique (ISODATA) algorithm was selected to classify the images automatically into 35 clusters. This algorithm is solely relying on pixel-based spectral statistics and incorporates no prior knowledge of the characteristics of the themes being studied. The benefit of applying unsupervised classification methods is to automatically convert spectral image data into spatial clusters which can then be used for land cover classification (11). This approach has been proven to be effective and well suited to forest species classification (15), and has also been applied for other global land cover classifications (16).

For each sample site one multi-temporal Landsat imagery stack is created by stacking the Red, NIR and SWIR bands (0.63-0.69 µm, 0.75-0.9 µm and 1.55-1.75 µm) of the selected images for three reference years (1990, 2000 and 2010). This creates a nine layer input file for the unsupervised classification. The choice fell on this method as it is apt at describing changes by way of creating clusters of spectrally similar pixels. This enables a straightforward and relatively fast validation.

**Automatic prelabeling**

The automatic image classification algorithm proposed by (17) has been used to pre label the three images (1990/2000/2010) independently. The algorithm was developed in IDL and operates on a single date TOA reflectance Landsat TM/ETM+ imagery. The output is a set of three raster files (one for each epoch) containing the most probable land cover class for each pixel.

**Pre-classification: from pixels to objects**

The ISODATA unsupervised classification is combined with the shapefile derived from the segmentation step using a straightforward majority rule. As output, each individual object carries information regarding the cluster ID. Similarly, the predominant preliminary pixel based classification label within each object and epoch represents the land cover class that will be further confirmed or changed by the expert. The majority filter is a simple and reliable approach although some complex land cover classes as peatland or temporary vegetation cover status as burned might be better identified using a more complex rule set implying class proportions.

**Expert interpretation**

The use of expert knowledge in forestry mapping has been demonstrated in several studies as being a viable solution to improve or correct the final product (18,19,20).

The phase of visual control and refinement has been introduced in the workflow as a crucial component for correcting classification errors and for implementing the change assessment. Only at this stage the three individual digital classifications derived from an image triplet were linked in the graphical interface for a visual interpretation and consistency control. Up to this point the automated classification of the imagery has been done independently for each reference year, in spite of using a common spectral signature database.

A dedicated software solution was developed to optimize the expert interpretation task (21). The software presents the expert with a user interface displaying simultaneously the satellite imagery and the classifications for one specific site. Thus it allows the interpreter to verify and potentially re-assign land cover labels for individual objects or cluster of objects. The tool is highly interactive and allows the expert to zoom in and out, apply image enhancements and offers selective object selection options. With the use of this dedicated tool the expert interpretation achieves a twofold result:

1. Obvious mistakes from the unsupervised classifications are corrected
2. A thorough revision of the mapping results is carried out with in depth knowledge of local and regional specificities

## RESULTS

### Segmentation and isodata

The multi date segmentation method was judged by the expert interpreter to deliver homogeneous landscape units. Where land cover change occurred, the area was captured as an individual 'change' polygon. Since the object boundaries are constrained to fit to a multi-data dataset, slivers are avoided. Such small areas are integrated within the surrounding and spectrally closest larger polygon, similarly to the findings of Raši et al. (8).

For the subset of 502 sample sites selected across Russia and Mongolia the segmentation produced on average 3353 objects per site. There is a clear latitudinal gradient within the distribution of the amount of objects per site. The sites between 50°N and 58°N have on average 4595 objects, a significantly higher number. This corresponds to the latitudes mainly associated with more heterogeneous and fragmented landscapes. A significant longitudinal gradient is non-existent.

The average mean size for objects is 15.41 ha (with a standard deviation of 8.24 ha). For about 84.9% of the sample sites the mean size of objects is smaller than 25 ha. There was no specification of a minimal mapping unit (MMU) during the segmentation process. Indeed, the segmentation software used in our study does not even support this concept. Despite several examples of study's implementing MMU's of 5 ha for forest monitoring (8), 5 ha for the CORINE Land cover mapping in Europe and 6.25 ha for operational deforestation monitoring in Amazonia (22). Even though (8) used the same input data, they have had to implement a MMU by using the quantile-5 object size as approximation. By a combination of applying a multi date segmentation on three Landsat images simultaneously (as opposed to two images), and the differences in homogeneity of boreal and moderate latitude landscapes (as compared to tropical landscapes) our segmentation results have yielded segments that in 99.15% are bigger than 5 ha. Thus, the devised method effectively brings about a MMU of 5 ha.

### Interpretation and validation

All the 502 20 km $\times$ 20 km boxes have been used to assess the reliability of the pre-classification. By comparing the pre-classification and the final label for each object it is possible to get an idea of the performance of the system.

The Water class has the highest accuracy rate (94%), Clouds & Shadows are well identified (82.6%). Other Wooded Land and Shrubs have a pre-labeling accuracy of more than 70%. Forest classes show the highest variability spanning from more than 85% accuracy for Needleleaf to 37% corresponding to Sparse Deciduous forest. Despite the low values of the latter (mainly due to its seasonality effect, making it difficult to catch with a single date image classification), the entire Forest classes accuracy is more than 74%.

Noteworthy is the low accuracy of the class Other Land. This can be attributed to the heterogeneity of the class itself: according to the definition it comprises bare soil, grass and agricultural fields; the latter, during the growing season, are often labelled as wooded vegetation or forest by the preliminary classification.

*Table 2: Accuracy of the pre-labeling.*

| Classes | Pre-labeling accuracy % |
|---|---|
| All Forests | 74.41 |
| Evergreen Dark Needleleaf | 84.74 |
| Evergreen Light Needleleaf | 72.71 |
| Deciduous Broadleaf | 63.43 |
| Deciduous Needleleaf | 55.85 |
| Sparse Deciduous | 37.60 |
| Mixed Needleleaf Majority (80%) | 82.74 |
| Mixed Forests | 79.56 |
| Mixed Broadleaf Majority (80%) | 70.05 |
| Water Bodies | 93.95 |
| Other Wooded Land & Shrubs | 71.15 |
| Clouds and Shadow | 82.70 |

## CONCLUSION

The proposed method answers the current political and scientific need for capturing fine scale forest change dynamics across huge stretches of land. Based on advanced image processing, the developed methodology efficiently combines different approaches including systematic sampling, multi-date image segmentation, automated processing and expert correction to produce robust forest change estimates.

The sampling scheme is a based on a randomized extract of a systematic distribution of image extracts in order to preserve a similar sampling rate as in the Tropics (4% of land mass). Where available, high quality Landsat imagery was acquired for the epochs 1990, 2000 and 2010. The developed approach to characterise the land cover illustrates how the combination of multi-date segmentation, unsupervised classification and automatic pre-labeling can be combined to enable land cover classification of a substantial amount of imagery.

Not only does this method provide a rapid, cost-efficient forest monitoring tool, but above all it allows a high classification detail with regards to forest composition. This is of relevance for obtaining quantitative assessments of change type classifications.

## REFERENCES

1 IPCC. 2000. Emission scenarios. Nebojsa Nakicenovic and Rob Swart (Eds.) (Cambridge University Press, UK) 570 pp.

2 Groisman P & A J Soja, 2009. Ongoing climatic change in Northern Eurasia: justification for expedient research. Environmental Research Letters, 4, 045002, 7 pp.

3 Filiptchouk A N, V V Strakhov & V A Borisov, 2001. Forest and forest products country profile: Russian Federation. Geneva Timber and Forest Study Papers No. 19, UNECE, New York

4 FAO. 2012. The Russian Federation Forest Sector: Outlook study to 2030 (Food and Agriculture Organization of the UN, Rome, Italy)) 93 pp.

5 Gallego FJ. 2005. Stratified sampling of satellite images with a systematic grid of points. ISPRS Journal of Photogrammetry and Remote Sensing, 59: 369-376.

6 Beuchle, R, H Eva, H J Stibig, C Bodart, A Brink, P Mayaux, D Johansson, F Achard & A S Belward. 2011. A satellite data set for tropical forest area change assessment. International Journal of Remote Sensing, 32: 7009-7031

7 Bodart C, H R Eva, R Beuchle, R Raši, D Simonetti, H J Stibig, A Brink, E Lindquist & F Achard, 2011. Pre-processing of a sample of multi-scene and multi-date Landsat imagery used to monitor forest cover changes over the tropics. ISPRS Journal of Photogrammetry and Remote Sensing, 66: 555-563

8 Raši R, C Bodart, H-J Stibig, H Eva, R Beuchle, S Carboni, D Simonetti & F Achard, 2011. An automated approach for segmenting and classifying a large sample of multi-date Landsat imagery for pan-tropical forest monitoring. Remote Sensing of Environment, 115: 3659-3669

9 Meinel G & M Neubert, 2004. A comparision of segmentation programs for high resolution remote sensing data. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXV-B4: 1097-1102

10 Baatz M & A Schäpe, 2000. Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation. In: J Strobl, T Blaschke & G Griesebner (Eds), Angewandte Geographische Informationsverarbeitung, XII, p. 12-23 (Heidelberg: Wichmann-Verlag)

11 Xie Y, Z Sha & M Yu, 2008. Remote sensing imagery in vegetation mapping: a review. Journal of Plant Ecology, 1(1): 9-23

12 Wilkinson G G, 2005. Results and implications of a study of fifteen years of satellite image classification experiments. IEEE Transactions on Geoscience and Remote Sensing, 43: 433-440

13 Rozenstein O & A Karnieli, 2011. Comparison of methods for land-use classification incorporating remote sensing and GIS inputs. Applied Geography, 31: 533-544

14 Lo C P & J Choi, 2004. A hybrid approach to urban land-use/cover mapping using Landsat 7 enhanced Thematic Mapper Plus images. International Journal of Remote Sensing, 25: 2687-2700

15 Reese H, T Lillesand, D Nagel, J Stewart, R Goldmann, T Simmons, J Chipman & P Tessar, 2002. Statewide land cover derived from multi-seasonal Landsat TM data: a retrospective of the WISCLAND project. Remote Sensing of Environment, 82(2-3): 224-237

16 Hansen M C & B Reed, 2000. A Comparison of the IGBP DISCover and University of Maryland 1 km Global Land Cover Products. International Journal of Remote Sensing 21(6&7): 1365-1373

17 Szantoi Z & D Simonetti, D. 2013. Fast and robust topographic correction method for medium resolution satellite imagery using a stratified approach. IEEE Journal of Selected Topics in App. Earth Observations and Remote Sensing, 6(4): 1921-1933

18 Lepers E, E F Lambin, A C Janetos, R DeFries; F Achard, N Ramankutty & R J Scholes, 2005. A synthesis of information on rapid land-cover change for the period 1981–2000. Bioscience, 55: 115-124

19 Hansen MC, S V Stehman, P V Potapov, T R Loveland, J R G Townshend, R S DeFries, K W Pittman, B Arunarwati, F Stolle, M K Steininger, M Carroll & C DiMiceli, 2008. Humid tropical forest clearing from 2000 to 2005 quantified by using multitemporal and multiresolution remotely sensed data. Proceedings of the National Academy of Sciences of the Unites Staates off America, 105(27): 9439-9444

20 FAO. 2009. The 2010 global forest resources assessment remote sensing survey: An outline of the objectives, data, methods and approach. Forests Resources Assessment Working Paper – 155, Food and Agriculture Organization of the UN (Rome, Italy) 16 pp.

21 Simonetti D, R Beuchle & H Eva, 2011. User Manual for the JRC Land Cover/Use Change Validation Tool, Office des Publications de l'Union Européenne, EUR 24683 EN, 23 pp.

22 INPE, 2013. Projeto PRODES. Monitoramento da Floresta Amazônica Brasileira por Satélite (last date accessed: 13 Nov 2014)