Monitoring Global Forest Land-Use and Change

Erik J. Lindquist
South Dakota State University

Follow this and additional works at: http://openprairie.sdstate.edu/etd

Part of the Geographic Information Sciences Commons, Physical and Environmental Geography Commons, and the Remote Sensing Commons

Recommended Citation
MONITORING GLOBAL FOREST LAND-USE AND CHANGE

BY

ERIK J. LINDQUIST

A dissertation submitted in partial fulfilment of the requirements for the

Doctor of Philosophy

Major in Geospatial Science and Engineering

South Dakota State University

2016
MONITORING GLOBAL FOREST LAND-USE AND CHANGE

This dissertation is approved as a creditable and independent investigation by a candidate for the Doctor of Philosophy in Geospatial Science and Engineering degree and is acceptable for meeting the dissertation requirements for this degree. Acceptance of this does not imply that the conclusions reached by the candidate are necessarily the conclusions of the major department.

Matthew Hansen, Ph.D.
Dissertation Advisor Date 29APR2016

Michael Wimberly, Ph.D.
Academic Advisor Date: 29APR2016

Geoffrey M. Henebry, Ph.D.
Co-Director GSCE Date: 29APR2016

Dean, Graduate School Date:
ACKNOWLEDGEMENTS

This work was supported under the NASA Earth and Space Science Fellowship Program (NNX07AO57H), the United States Agency for International Development and the European Commission.

Professionally, I would like to express my most sincere gratitude to my major advisor, Dr. Matthew Hansen, and the members of my graduate committee: Dr. Thomas Loveland, Dr. David Roy and Dr. Michael Wimberly. Your guidance, knowledge and patience are appreciated beyond my ability to express. I thank too my colleagues at the Geographic Information Science Center of Excellence and University of Maryland with whom I worked closely including Dr. Christopher Barber and Ms. Alice Altstatt. I would also like to thank my colleagues at the Food and Agriculture Organization of the United Nations, especially my close friends and collaborators Dr. Remi D’Annunzio, Dr. Anssi Pekkarinen, Dr. Kenneth MacDicken, Ms. Mette Wilkie and Mr. Adam Gerrand, you all provided inspiration, drive, discipline and were amazing travel companions while completing the Global Remote Sensing Survey. I want to especially thank my friends and colleagues in the Democratic Republic of Congo at ERAIFT, OSFAC, OFAC and DIAF. It is impossible to thank the number of people I met in a professional context along the way and in my travels that have influenced me and provided inspiration. Suffice it to say that I have memories of you all and am grateful to have come into contact with you.
Personally I would like to thank Drs. John and Therese Hart and Mr. Didier Devers. You saved me, put me on this path and now I have closed the circle. I owe you each so much. I would also like to thank Mr. Jacob Limmer for being a fantastic friend in Brookings. South Dakota would not have been nearly as much fun without you. And, of course, I thank my parents for the patient, quiet, unending support though everything. I am happy I can look at you both and tell you it’s complete now. Finally, I thank my daughter Ocia for reminding me what’s really important and to never, ever quit.
CHAPTER 4

ASSESSING GLOBAL FOREST LAND-USE CHANGE BY OBJECT-BASED IMAGE ANALYSIS

4.0 Abstract

4.1 Introduction

4.2 Materials and Methods

4.2.1 Image segmentation

4.2.2 Automated Image Classification

4.2.3 Neural network land cover classification

4.2.4 Multivariate Alteration Detection for land-cover change

4.2.5 Assigning class labels to 1990 and 2005 time periods

4.3 Results and Discussion

4.3.1 OBIA, minimum mapping units and land cover classification

4.3.2 OBIA, minimum mapping units and change detection

4.3.3 Landscape metrics and change detection

4.3.4 Automated land-cover classification and change detection results
4.3.5 Land cover v. land use

4.3.6 Effect of land cover or land use definition on forest area calculations

4.4 Conclusions

CHAPTER 5

GLOBAL FOREST LAND-USE CHANGE FROM 1990 - 2005

5.0 Abstract

5.1 Introduction

5.1.1 The FRA 2010 Remote Sensing Survey

5.2 Methods and materials

5.2.1 Land cover and land use

5.2.2 Systematic sample design

5.2.3 Imagery data sources

5.2.4 Image preprocessing

5.2.5 Automated land-cover classification

5.2.6 Training the classification

5.2.7 Land-use classes

5.2.8 Conversion of land cover to land use

5.2.9 Expert interpretation, validation and correction of land cover and land use

5.3 Data analysis

5.3.1 No data

5.3.2 Adjustment for latitude and area weighting

5.3.3 Aggregation for regional and climatic domain analysis

5.3.4 Forest area: gains and losses
5.3.5 Annualizing forest-area change 93
5.3.6 Error 93

5.4 Results and discussion 95
5.4.1 The area in forest land-use declined between 1990 and 2005 98
5.4.2 Global forest loss and gains 100
5.4.3 Regional differences in forest loss and gain 101
5.4.4 Differences in the annual rate of change by region and climatic domain 104

5.5 Comparison with other FAO studies 105
5.5.1 Comparison with FRA 2000 pantropical remote sensing data 105
5.5.2 Comparison with FRA 2010 tabular reports 109
5.5.3 Differences in forest area 110
5.5.4 Differences in net forest area change 113

5.6 Causes of land-use change 115

5.7 Accuracy assessment 117

5.8 Conclusion 117
5.8.1 Integration of coarse resolution satellite imagery to help classification 118
5.8.2 Visual review and revision of classification important 119
5.8.3 The utility of Landsat for global monitoring 119
5.8.4 Establishment of global networks 120

Annex 5-1 Country-specific review-and-revision methodologies 120
Annex 5-2 Survey sites processed vs analysed 124
Annex 5-3 Review-and-revision contributors 125
Annex 5-4 Details of calculation 128
6.0 Summary of Chapters

6.1 Recommendations for future research or work

6.1.1 Decrease uncertainties in estimates of forest cover, use and rates of change.

6.1.2 Improve operational methods of integrating high spatial resolution optical data.

6.1.3 Ensure continuous and frequent acquisition of Landsat-like data.

6.1.4 Increased access to high performance computing facilities.

6.1.5 Exploit the power of the global Landsat image archive via time series analyses to automatically produce land use classifications.
Figure 3-1. The study area (a), and 2005 epoch mosaics generated to illustrate DFCM compositing of (b) one, (c) two, (d) three, (e) four and (f) five Landsat-7 ETM+ acquisitions per path/row. Colour composites of Landsat-7 ETM+ bands 4, 5 and 7 are shown. Missing ETM+ SLC-off pixels are evident as white stripes and residual cloud contaminated pixels are evident in white, particularly in (b). Each mosaic is composed of more than 51 million 57m pixels.

Figure 3-2. Histograms of the percentage of pixels in the four Landsat epochal mosaics for each QA value. QA values in descending order of quality are: 1, no cloud/haze or shadow (i.e. best quality); 2, low likelihood of cloud/haze; 3, low likelihood of shadow; 4, adjacent to high likelihood cloud/haze or shadow; 5, high likelihood of cloud/haze; 6, high likelihood of shadow; 7, missing due to SLC-off gaps. The four Landsat epochal mosaics are labeled in the key as method, number of Landsat acquisitions used per path/row, and epoch period. See text for further details.

Figure 3-3. The percentage of good quality pixels (QA values 1, 2 and 3, i.e. low likelihood of cloud/haze or shadow) and best quality pixels (QA value 1, i.e. no cloud/haze or shadow) occurring at the same pixel locations in both epochal mosaics: grey shows a comparison of the 2000 Geocover (1 acquisition) and the 2005 simulated GLS (3 acquisitions) mosaics; black shows a comparison of the 2000 DFCM (2–3 acquisitions) and the 2005 DFCM (5–7 acquisitions) mosaics.
Figure 4-1. (From Lindquist et al., 2012) The global, degree grid framework, including the Canadian NFI plot scheme. FAO processing methodology was applied to sites (>11,000) in the boreal, temperate and sub-tropical climatic domains (in black). ............ 46

Figure 4-2. (From Lindquist et al., 2012) A multi-date segmentation for a 20km x 20km site in the boreal climatic domain. Landsat imagery from 1990, 2000, 2005 are combined into a single data stack and segmented. The segments (white outlines in image on right) capture areas of spectral similarity as well as areas of change over time. The Landsat image on the right is a color composite of Landsat band 5 from 1990 (red), 2000 (blue) and 2005 (green). Changes in tree cover between time periods are evident.................. 48

Figure 4-3. The multi-date processing scheme showing for a boreal forest sample site in year 1990 (top row), year 2000 (middle row) and year 2005 (bottom row) the following: (A) the 20x20km Landsat imagery (B) the potential training objects (grey) for year 2000, (C) the training dataset after class labeling, (D) the final automated land cover classification, (E) the final, expert revised land cover classification of the central 10x10km (corresponding to white box in D), and (F) the corresponding final, expert revised land use classification. Dark green = Tree Cover / Forest, Light green = Tree Cover Mosaic, Orange = Shrub / Other Wooded Land, Pink = Other Land. ................. 49

Figure 4-4. (Left) A comparison of the tree cover area obtained automatically for level-1 segmentation (x-axis) and the 5 ha MMU (y-axis) and (Right) a comparison of the tree
cover area obtained automatically at the 5 ha MMU (x-axis) and the expert reviewed-revised tree cover at the 5 ha MMU (y-axis). The dotted line represents the slope of the line formed by the linear regression. The solid line represents the one-to-one line.

Figure 4-5. The amount of change detected between 1990 - 2000 using IR-MAD and increasing change thresholds (columns) aggregated to level-1 and 5 ha MMU segments (rows) for a 20x20 km sample site. Columns left to right represent thresholds of 0.91, 0.95 and 0.99 of the IR-MAD change likelihood layer, respectively. Rows top to bottom represent pixel-level, level-1 image segments and 5 ha MMU segments, respectively. Changed areas are in black. Non-changed areas are white.

Figure 4-6. (Left) The mean area detected as change (Y-axis) over all samples using pixel-based (triangles), level-1 segments (+) and level-2 segments (circles) at increasing IR-MAD thresholds (X-axis). (Center) The difference in the total area detected as change between pixel and level-2 segments plotted as a function of the change patch landscape shape index. (Right) The difference in the total area detected as change between pixel and level-2 segments plotted as a function of the mean change patch area.

Figure 5-1. The 15,779 1-degree grid sample site locations used in the survey, with reduced intensity above 60° latitude north and south. Canada samples were spaced on a 20-km grid to match the Canadian National Forest Inventory (inset, see Annex 1 and Annex 2). Sites processed by JRC are in grey and sites processed by FAO are in black.

Figure 5-2. Example of three imagery dates combined to make a single composite image.
with segments that capture reflectance changes in each period.

Figure 5-3. Generalized flowchart of the processing chain.

Figure 5-4. Land-cover and land-use classes and their associated numeric codes. In the conversion from land cover to land use, tree cover is converted to forest, shrub cover is converted to other wooded land, other land cover is converted to other land and water stays as water. Ideally, where there was a change in land use either to or from forest, the subclasses of other land use were to be used to identify the cause of the change.

Figure 5-5. JRC validation tool user interface showing Landsat imagery from each survey period (left) and FAO land-use classes (center).

Figure 5-6. Final 13,066 sample sites used in the 2010 analysis. Note the absence of sites in the Eastern Russian Federation due to a lack of Landsat acquisitions from the 1990s time period.

Figure 5-7. Regions (top) and climatic domains (bottom) used for aggregation and analysis.

Figure 5-8. The range of dates of satellite imagery used in the study by survey period. The table below the graph lists the earliest, latest, average and median dates for each survey period.
Figure 5-9. Forest area by region for 1990, 2000 and 2005. .................................................. 99

Figure 5-10. Forest area by climatic domain for 1990, 2000 and 2005. ......................... 100

Figure 5-11. Gross gains and losses and net changes in forest area, by FRA region, 1990-2000 and 2000-2005. ........................................................................................................ 102

Figure 5-12. Gross gains and losses and net changes in forest area, by climatic domain, 1990-2000 and 2000-2005. .................................................................................................. 103

Figure 5-13. Distribution of RSS 2000 sample sites in the pantropics. The 117 sampling units of the survey were selected over the entire pantropical zone following a two-stage random sampling method based on geographical divisions (subregions) and forest cover or forest dominance ........................................................................................................ 106

Figure 5-14. Pantropical forest area for year 1990 (top) and 2000 (bottom) as estimated from RSS 2000 (blue bars) and this study (red bars) ......................................................... 107

Figure 5-15. Comparison of pantropical net forest area change and gross forest area loss between 1990-2000 as estimated by RSS 2000 (blue bars) and this study (red bars). ... 108

Figure 5-16. Pantropical forest area for the years 1980, 1990, 2000 and 2005 as estimated
from RSS 2000 (1980) and this study (all other years). ........................................ 109

Figure 5-17. A comparison of forest area, by region, as reported in FRA 2010 tabular reports (blue dots) and the remote sensing-based estimates presented in this paper represented by black bars indicating the 95 % confidence intervals about the mean.... 111

Figure 5-18. Relationship between proportion of dry climatic domains per region and the proportional difference between FRA 2010 and RSS 2010 forest area estimates for that region. ........................................................................................................................................................................ 113

Figure 5-19. A comparison of net change between the remote sensing-based estimates in this paper (bars with 95% confidence interval indicated) and FRA 2010 tabular reports (diamonds) for 1990-2000 (left) and 2000-2005 (right). ......................................................... 114
LIST OF TABLES

Table 3-1. Per-pixel QA values, the relative quality, and the defining characteristics of those values obtained as part of a decadal forest cover change mapping algorithm developed by Hansen et al. (2008), and used to create best-pixel temporally composited Landsat mosaics in central Africa. .......................................................... 34

Table 4-1. Training segment labelling rules based on underlying MODIS VCF, GLOBCOVER, MODIS Water Mask and hard-coded Landsat DN thresholds. .............. 50

Table 4-2. Level-2 segment labelling rules for the automated land cover classification. Level-2 labels were based on the areal proportion of component level-2 segments. ...... 54

Table 4-3. Confusion matrix between land cover classes before and after expert review and revision for a sample of year 2000 sites. Figures are based on area in each class in millions of hectares. Automatically mapped class labels are on left-hand side and the expert reviewed and revised labels are across the top of the table. Tree = Tree cover, Wooded = other wooded land, Other = other land cover.................................................. 61

Table 4-4. Confusion matrix for change classes for the time periods 1990 – 2000 (top)
and 2000 – 2005 (bottom) before and after expert review and revision. Figures are based on area in each class in millions of hectares. Automatically mapped class labels are on left-hand side and the expert reviewed and revised labels are across the top of the table. Tree-Tree = Tree cover in both periods, Tree - Other = Tree cover change to other land, Other - Tree = other land to tree cover, Other - Other = other land cover in both periods.

........................................................................................................................................... 62

Table 4-5. The mean proportion of forest when using a land-use or tree cover definition by Continent/Country and Climatic Domain grouping................................................. 65

Table 5-1. Level-2, 5-ha MMU land-cover labelling scheme based on the percent composition of underlying level-1 segments, listed in descending order of priority....... 80

Table 5-2 Mean forest area (‘000 ha ± standard error) by region (a) and climatic domain (b), 1990, 2000 and 2005. The sum of the forest areas of all regions was used as the global forest area total............................................................................................................................................... 94

Table 5-3. Significance of net annual changes and gross annual gains and losses for regions and climatic domains, a * indicates a value significantly different from zero (p < 0.05) using Welch’s t-test. ........................................................................................................ 96
Table 5-4. P values for the slope of the line formed by a general linear model relating annualized net change and gross gains and losses with survey period by regions and climatic domains. Significant differences (p < 0.05) between survey periods are in green. For net change, the direction of the arrow indicates whether there was a net forest area loss (↓) or gain (↑)................................. 96

Table 5-5. Results of ANOVA test for annual net forest area change, by climatic domain and year................................................................. 97

Table 5-6. REML results for annual net change by climatic domain and survey period (1990-2000 and 2000-2005). ............................................................... 98

Table 5-7. Mean annual net forest area change and 95 % confidence intervals between survey periods for FAO regions and FAO climatic domains. Global net change was calculated by summing estimates for FAO regions.................................. 101

Table 6-1. A comparison of recent global forest cover and land use area and change estimates. All areas in million sq. km. Net figures indicated with (*)......................... 140
ABSTRACT

MONITORING GLOBAL FOREST LAND-USE AND CHANGE

ERIK J. LINDQUIST

2016

Earth’s forests contain nearly three-fourths of the World’s floral and faunal diversity, function as a large carbon sink capable of mitigating the effects of global climate change, affect local and regional physical and chemical cycles and provide wood and non-wood products. However, humans are now capable of modifying their environment in ways more impactful and at rates faster than at any other time in history. Consistent and comparable estimates of global forest land-use and change are critical for monitoring human impacts on the Earth system. International treaties and reporting requirements aimed at safeguarding the delivery of forest-related ecosystem services depend on such estimates for measuring progress against their stated goals. Many existing studies have estimated tree cover and change at a variety of spatial scales from local to global. However, this existing research focuses largely on land cover classification, but generally lacks ecological context for estimating true human land use.
The objective of this dissertation is to address this gap by exploring how forest land use can be mapped and monitored using medium spatial resolution optical satellite imagery in order to estimate forest land use change over time for large geographic areas. First, the effects of clouds, cloud shadows and missing data were analyzed to determine the amount of moderate spatial resolution, optical satellite data needed to detect and map land cover changes over large, spatially continuous areas on frequent time intervals. Second, an alternative method to spatially exhaustive mapping was developed and tested for estimating land cover and land use change globally employing object-based image analysis and a sample-based estimation approach. The method facilitated expert human intervention to identify true land use change in an operational way. Finally, these methods were applied to a globally distributed sample of remotely sensed data for the time periods 1990, 2000 and 2005. The results of this research produced the first consistent and comparable global time-series dataset of forest land-use estimates.
CHAPTER 1

INTRODUCTION
1.0 Introduction

Forests provide vital economic, social and environmental benefits on local, regional and global levels. They supply wood and non-wood forest products valued in the billions of dollars annually (FAO, 2014), over 25 percent of the World’s forests serve protective functions, including safeguarding watersheds for drinking water (Miura et al., 2015), contain two-thirds of the planet’s biodiversity in the tropical forests alone (Gardner et al., 2009) and support human livelihoods through protective, productive, economic and spiritual resources (Vedeld et al., 2007). Forest photosynthesis and respiration combined with the physical properties of trees, such as structure and density affect local hydrologic conditions (Spracklen, Arnold & Taylor, 2012), chemical and energy cycles (Bonan, 2008) and, in combination with large-scale atmospheric circulation patterns, influences regional and global climatic conditions (Avissar and Werth, 2005). Forests, their protection or exploitation, can serve to either mitigate or accentuate the effects of global climate change (Jackson et al., 2008).

The World’s forest resources are clearly important. However, the natural systems of the Earth are increasingly human dominated (Vitousek et al., 2008). Humans are deciding, by purposeful use or non-use of the land resource, the kind and distribution of landscapes existing in all corners of the world. Monitoring land use, or the way humans decide to use (or not use) a particular piece of land, is extremely important for understanding and managing humankind’s relationship with the biosphere, particularly how changes in land use affect terrestrial biogeochemical processes and ecosystem services provided by natural systems (Foley et al., 2005).
Land use and land cover are different. Land use is function based, and can be 'defined in terms of a series of activities undertaken to produce one or more goods or services' (diGregorio, 2005). Land cover refers to the biophysical characteristics of the Earth’s land surface. A land use change typically represents a change in the level of economic productivity and is frequently associated with a change in land cover. In Paraguay, for instance, the clearing of the dense chaco woodlands for pasture (Huang et al., 2009; Vallejos et al., 2014) represents a shift to a higher order of economic productivity and is accompanied by a land cover change from trees to graminoids, forbs or bare ground. In post-Soviet era Russia, however, abandoned farmlands have been slowly re-colonized by forest representing a land cover change from crops to trees and a land use of lower order economic productivity (Kuemmerle et al., 2011). Many land cover changes are natural, however and have no land use dynamics, for example the removal of boreal forest cover due to fire. In this case, the removal of tree cover is only a temporary change in land cover and does not necessarily represent any shift, positive or negative, in economic productivity.

The subsequent effects of land use and land cover changes on the Earth system are also different. In the examples above, the conversion of tropical rain forest to cropland and the burning of boreal forest, both represent a change in land cover from tree to non-tree life forms. However, fire is an integral part of the boreal ecosystem; the flora and fauna are adapted to fire and it may even be required for boreal systems to function properly (Ryan, 2002). After fire, and in the absence of continued disturbance or modification, the trees will re-grow and the system will return to its pre-burn state. In other words, the forest ultimately remains a forest. The conversion of tropical forest to
cropland is altogether a different scenario and is not considered part of the natural cycle of properly functioning forests. Conversion to crops or pasture results in near complete removal of the tree canopy, alters habitat for native flora and fauna (Dunn, 2004; Skole and Tucker, 1993), is associated with increased human and domestic livestock activity and changes the way energy and nutrients are exchanged and cycled within the local ecosystem (Giambelluca et al., 2000).

Remote sensing is commonly used to assess the state and trend of the Earth’s land surface to produce estimates of land cover and change. Remotely sensed data provides consistent, well-calibrated, systematic land surface observations over large areas and multiple time periods. These observations can then be used to characterize the Earth’s surface in terms of land cover and, ultimately, land use. Landsat is one of the most widely used sources of information for land-based remote sensing in the World with a total of nearly 37 million individual scenes downloaded by users as of January, 2016 (USGS, 2016). The combination of Landsat's medium spatial resolution (30m pixel size) and ability to detect electromagnetic radiation across the visible, infra-red and thermal wavelengths allows the discrimination of features on the Earth’s surface important for characterizing vegetation type, condition, shape and areal extent (Williams, Goward and Arvidson, 2006). Landsat’s 40-year archive of well-calibrated image acquisitions makes the sensor’s data invaluable for detecting changes in these vegetation attributes over time (Markham and Helder, 2012; Wulder et al., 2015). Finally, the free and open, web-based distribution policy of Landsat data (Woodcock et al., 2008; Wulder et al., 2012) has made the information collected by the sensor readily accessible in large quantities for classifying, mapping and monitoring the Earth’s surface over very large areas (Roy et al.,
Operational methods for monitoring land use change over large geographic areas from satellite imagery are not common in historic literature (Coulston et al., 2013). Though land cover dynamics are extremely important, a system of true forest land-use monitoring ought also be incorporated into regularly produced updates on the state and trend of the Earth’s forests. Estimating global forest land-use area and changes in area over time is critical for three main reasons: 1) it serves to track the current status and trend of forest land-use in order to assess the potential effects to ecosystem services and human wellbeing, 2) drive management decisions that can reverse or mitigate potential negative impacts of forest loss and, subsequently 3) enable the meaningful evaluation of those decisions against desired outcomes such as the reduction of deforestation and biodiversity loss. The subject of this dissertation is global forest land-use. The chapters are meant to explore how forest land-use may be classified and monitored using medium spatial resolution optical satellite imagery in order to estimate forest land-use extent and change over time for large geographic areas. The proceeding sections of this document explore how such a monitoring system, at least for forest land-use, can be established and a first-ever global assessment of forest land-use and change is presented.

1.1 Research questions

Global forest land-use monitoring, its feasibility with respect to available data, appropriate methodologies and, ultimately, results obtained from an applied approach will be addressed through three main research questions:
Research Question 1. How do per-scene percentages of cloud, cloud shadow, haze and missing data affect the area of Landsat data required to create temporal composites suitable to monitor land cover and land-use changes over time?

Research Question 2. What are the strengths and limitations of object-based image analysis methods used to estimate land-use and change?

Research Question 3. What is the global extent of forest land-use, how does the area of forest land-use differ by geographic region and major climatic zone, and how have these areas changed over time?

Research question one examines the feasibility of spatially exhaustive (e.g. wall-to-wall), large-area monitoring with Landsat or Landsat-type optical remote sensing data. Spatially continuous land surface observations at multiple points in time are essential to produce large-area maps and associated estimates of land cover, land use and change. Wall-to-wall mapping may be complicated, however, due to a number of limitations including the presence of cloud, cloud shadow, missing data and a 16-day revisit period for Landsat that yields only 22 or 23 annual opportunities to capture a cloud-free or otherwise high-quality acquisition. In the Brazilian Amazon, Asner (2001) analyzed Landsat metadata cloud-cover estimates from 1984 to 1997 to conclude that obtaining optical remote sensing data with sufficient cloud-free observations (e.g. < 30%) to produce time-series analyses on a monthly basis was not possible and, on a yearly basis
was possible but difficult. Since cloud cover is a locally variable phenomenon, the probability of successful large-area monitoring depends on the distribution of cloud cover across the study area (Roy et al., 2006). The utility of Landsat acquisitions were further constrained by a failure, in year 2003, of Landsat 7 that rendered 23% of each acquisition unusable (Williams, Goward and Arvidson, 2006).

These complications make accumulating data to completely cover the Earth’s land surface, annually or seasonally, challenging. Kovalskyy and Roy (2013) analyzed the Landsat metadata record to show that the mean probability of obtaining at least one cloud-free acquisition in a twelve month period in each of the three seasons with the highest seasonal probabilities of cloud-free observations was 0.74 in year 2000 and 0.62 in year 2010, when considering only the ETM+ sensor. This proportion increased to 0.79 and 0.73 for 2000 and 2010, respectively, when considering both TM and ETM+ acquisitions together. Ultimately, it required 36 months to achieve probabilities of 0.92 and 0.90 for 2000 and 2010, respectively.

Per-scene cloud cover metadata can serve as a general indication of the overall scene cloudiness, and thus provides utility when selecting data for the purposes of monitoring (Irish et al., 2006) or estimating the likelihood of obtaining enough high-quality Landsat data for large-area land cover monitoring and change detection within a given time period (Ju and Roy, 2008; Kovalskyy and Roy, 2013). Landsat metadata cloud cover records, however, provide only an overall indication of cloud cover for a specific acquisition, either for the scene as a whole or disaggregated by quarter-scene increments. Spatially explicit indications of cloud, haze, cloud-shadow and missing data are necessary and may improve estimates regarding the number of individual acquisitions
and the realistic time interval required to produce imagery composites suitable for land surface characterization and change detection.

The first research question analyzes the effects of spatially explicit cloud, cloud shadow and no-data information on the amount of moderate spatial resolution, optical satellite data needed to map and detect land cover changes over large, spatially continuous areas on frequent time intervals. Given the large amount of data required for producing frequent estimates of land use, land cover and change over large areas, complex image compositing techniques or estimation methods other than wall-to-wall analyses may be advisable, depending on the dates and periodicity of the analysis.

Research question two details an alternative to spatially exhaustive mapping for estimating land cover and land use change globally employing object-based image analysis (OBIA) and a sample-based global survey. The sample-based method is proposed as a solution to estimate the area of the Earth’s surface in forest land-use and the changes in forest land-use over time. The method overcomes the difficulties of wall-to-wall mapping and, by facilitating expert human intervention to identify true land use change, provides an operational method for classifying land use directly from remotely sensed data. The sampling methodology described requires much less data than an spatially exhaustive mapping and, by incorporating object-based image analysis (OBIA) techniques along with a strong reliance on expert image re-interpretation, may be one of the most efficient methods available for land use characterization over large geographic areas.

Blaschke *et al.* (2014) describe OBIA as a new image processing paradigm and a significant improvement over traditional image processing and classification techniques.
OBIA treats image pixels as collections that form identifiable patterns on the earth’s surface. These identifiable units, or ‘objects’ are then defined as a whole, and circumscribed by a polygon. The unit of spatial analysis thus becomes the object, not the pixel. In the case of land use characterization, these objects may correspond to agricultural fields, urban parks or timber stands. All pixels within the defined object are assigned the label identifying the object and collectively define the object properties. Object-based assessments offer particular promise for characterizing land use classes as features can be identified as entire units instead of per-pixel.

There is some question, however, regarding the advantages of using OBIA on medium spatial resolution data when a single pixel can represent more than one object of interest on the ground (Blaschke, 2010; Duro et al., 2012). With regards to detecting biophysical land surface changes, there is very little documentation specifically illustrating the effectiveness of OBIA at detecting small-scale change using medium spatial resolution data. Research question two examines the effect of the change dynamic on classification results by comparing OBIA classification techniques with pixel-based assessments. The use of OBIA is also analyzed for the effect on classification results relative to object size (e.g. minimum mapping unit), human review and revision and, ultimately, conversion to land use.

Research question three addresses the need for consistent and comparable estimates of forest land-use and changes over time for large areas and is the result of applying methodologies tested in research question two. Current global forest land-use estimates are problematic (Grainger, 2008; Matthews, 2001). Operational methods for assessing forest land-use and land-use change over time globally do not currently exist.
and data limitations (from question one) make the prospect of frequently producing such estimates a challenge. Many studies exist for local and large areas both on assessing tree cover extent and change (Hansen et al., 2008; Hansen et al., 2013; Masek et al., 2008; Potapov et al., 2012). Fewer studies exist assessing forest land-use, especially over large areas or globally. This is likely because classifying forest land-use requires additional information than that which can be attained by a single satellite overpass. For instance when land is temporarily un-stocked due to the occurrence of fire or timber management activities these lands will likely be classified according to their condition at the time the image is acquired without respect to the overall ecological conditions or land management practices (Kurz, 2010).

Research question three applies the OBIA methods from research question two to a global sample of Landsat data from 1990, 2000 and 2005 in order to assess the areal extent of tree cover, forest land-use and changes in each between the epochs. In addition to global estimates, results are presented by continental regions and large climatic domains. The OBIA methods provide results more easily reviewable by expert human interpreters and, because the polygons created by the classification correspond to meaningful changes on the earth’s surface in their entirety (e.g. an agriculture field) facilitate the conversion from a traditional land cover classification to allow estimates of forest land-use and changes over time. The results generated from question three, considering the importance of earth’s forest resources, respond to the urgent requirement for true forest land-use assessments.
1.2 Summary of chapters

Chapter three addresses research question one and describes the methods used to assess the number and temporal span of Landsat acquisitions required to reduce the amount of cloud cover, cloud shadow and missing data to levels suitable for land surface characterization and change detection. For a six path/row study area in the humid, tropical forest of the Democratic Republic of Congo, results show that, 80% of pixels are of suitable quality (69% best quality) for change detection using composites with three Landsat acquisitions per path/row. Quality is determined based on a per-pixel assessment of cloud-cover, haze or missing data. The percentage of suitable quality pixels increases to 96% (89% best quality) when compositing five acquisitions per path/row. The results indicate that as many as five or more high-quality individual acquisitions per path/row may need to be acquired within a given year to enable forest change detection in the humid tropical forest. This amount of data may preclude frequent wall-to-wall mapping and change detection with Landsat, especially in the humid tropics. Chapter three was published in the *International Journal of Remote Sensing*.

Chapter four addresses research question two and describes the object-based methods used to characterize global forest land-use and change. The strengths and weaknesses of object-based image analysis processing system are analyzed and the limitations of object-based mapping methods with medium spatial resolution remotely sensed data are explored. Results indicate that the OBIA classification and change detection methodology provides an efficient means of processing a global Landsat sample-based dataset over three epochs. Implementing a relatively large MMU (5 ha) facilitates expert human review and revision. The conversion of the results to land use
was also made easier by OBIA as the objects represented meaningful, identifiable land use units on the ground. However, for change detection, if the change dynamic is very fine-scale, a much smaller or no MMU is advisable as image segmentation with a MMU may systematically under-segment areas of land cover change, some or all of which may be ecologically significant. Chapter four has been submitted for peer-review in the journal *Remote Sensing*.

Chapter five addresses research question three and describes the results of an analysis of global forest land-use change from 1990 to 2005 based on a systematic sample of Landsat imagery. Estimates of forest land-use area and rates of change between time periods are presented globally, by climatic domain and by geopolitical regions. Results show that the gross reduction in global forest land-use was 9.5 million ha per year between 1990 and 2000 and 13.5 million ha per year between 2000 and 2005. This reduction was partially offset by gains in forest area through afforestation and natural forest expansion of 6.8 million ha per year between 1990 and 2000 and 7.3 million ha per year between 2000 and 2005. Thus, the rate of annual net forest loss increased significantly ($p < 0.05$) from 2.7 million ha between 1990 and 2000 to 6.3 million ha between 2000 and 2005. There are significant differences in the rate of forest land-use change by large geographic region and climatic domain. Differences between the results of this study and other similar studies can be attributed to the difficulty some countries have in reporting their national forest statistics and also the difficulty of detecting forest area and changes within dry, sparsely vegetated ecozones. This is the first survey of its kind to assess forest land-use globally. Chapter five was published as FAO Forestry Paper 169.
Chapter six is the conclusion and places the research described in this dissertation within the context of emerging methodologies and immediate future applications.
CHAPTER 2

RESEARCH BACKGROUND
2.0 The significance of forests and forest change

Forests and trees are very important for human wellbeing. At the present time, approximately 3.9 billion hectares, or roughly 30% of the Earth’s land area is estimated to be in forest land-use (Keenan et al., 2015). Forests contain an estimated 861 Pg of carbon, 42% of which is stored in above or below ground live biomass and represent a large carbon sink if left standing (Pan et al., 2011). Of the World’s estimated 450,000 plant species, nearly two-thirds are known to be found in the tropics and largely in the forested zones (Pimm and Joppa, 2015). Tropical forests alone contain between 50 and 75% of all the Earth’s plant and animal species combined. The FAO (2010) estimates that almost 30% of the World’s forests are primarily used for wood production and non-wood products and that harvesting for wood-fuel accounts for around 50% of all wood removal globally. Again, FAO (2010) estimates that revenues from forest products have a 100 billion USD value annually and non-wood forest product revenue amounts to nearly 19 billion USD annually (2005 estimate). In total, between private and public institutions, the forestry sector employs nearly 12 million people globally. Local forest resources are also important for spiritual and cultural values and, indeed, some human communities still rely entirely on the forest for nearly all of their basic necessities. The importance of forests to people is also reflected in the laws and motivations governing the establishment and conservation of biodiverse, forested protected areas, especially in the humid tropics (Naughton-Treves et al., 2005).

Forest change, especially anthropogenic forest loss, is a threat to the immense value of forests. Human ability to alter landscape composition and pattern, largely to
produce goods and services, has evolved and, currently, has a greater impact on Earth’s natural systems than at any time in history (Klein Goldewijk et al., 2010). Human land use directly affects, through complete transformation or degradation, nearly half of the Earth’s land surface (Vitousek et al., 2008). Perhaps no greater change has occurred on the Earth’s land surface in the last 10,000 years than the changes to forests (Williams, 2000) which have decreased by almost 50% of their original extent (Kates and Parris, 2003).

The loss of forest affects the environment in three main and complex ways: through the loss and fragmentation of habitat critical for many plant and animal species, by altering the Earth’s energy balance and through the modification of other environmental services critical to human well-being. Forest loss has devastating effects on habitat and biodiversity by completely changing, fragmenting and/or degrading forest areas required for a large majority of the World’s species. This effect is especially strong in tropical forests (Lewis, 2006). Biodiversity loss is predicted to be even more of a threat to human wellbeing than is climate change (Sala et al., 2000). Species extinction rates in the tropical forests due to land use change are expected to be around 18% of currently existing species by the year 2100 (Pimm et al., 2014). Pimm et al. (2014) continue to predict extinction rates of up to 40% of currently existing species if only the areas currently in protected status are preserved.

Bonan (2008) distributes the effects of forest loss on climate and energy into three main categories, biogeochemical, biogeophysical and biogeographical. Biogeochemical effects include those to the Earth’s carbon cycle. Forest loss currently contributes about 10% of the total, global atmospheric carbon dioxide emissions (Watson et al., 2000;
Houghton, 2005; Pan et al., 2011) and has become the main focus of the global effort to Reduce Emissions from Deforestation and Degradation (REDD). Because carbon dioxide and other greenhouse gasses are well mixed in the atmosphere the conversion of forest to other land use anywhere on the planet contributes to the global warming effects of increased atmospheric carbon dioxide (Pielke et al., 2002). However, as Bonan et al. (2008) also describe, the effects of forest cover change on climate are uncertain and the amplitude and magnitude of these effects vary depending on the geographic location (e.g. boreal, temperate or tropical), the predictive models used to estimate the interactions between the Earth’s surface and climate change and the scenarios considered to project long-term land-use changes.

Biogeophysical changes refer to changes in the color, arrangement and amount of tree cover on the Earth’s surface. Changes in the biophysical properties of forests may have a greater forcing effect on local and regional climate than the effects of increased carbon dioxide (Bala et al., 2007; Jackson et al., 2008). Dense, green forests have a low albedo and absorb more incoming solar radiation and, thus, locally warm the Earth’s surface. Land use types with higher albedo (e.g. grassland or agriculture) reflect the incoming solar radiation away from the Earth’s surface and back into space resulting in local cooling of the Earth’s surface. In the tropics, the humid forest has very high rates of evapotranspiration. The evapotranspiration of the tropical forest more than compensates for the local warming effects of albedo and results in a net cooling effect of the tropical forest. Additionally, the evapotranspiration of the tropical forest is documented to account for between 25 and 50 % of local rainfall (Eltahir et al., 1994) and an increase in cloud cover. Cloud cover is responsible for reflecting a large amount of solar radiation
back into space and so also contributes to the local cooling effect of tropical forests. In boreal forests, albedo also has a local warming effect on the Earth’s surface (Bonan, Pollard and Thompson, 1992). The warming is not offset by the evapotranspiration as in the tropical forests. However, due to extensive snow cover in the winter months, the boreal zones of the Earth generally reflect more sunlight than they absorb and the overall result is a global cooling effect. Biophysical changes to the forestland on the Earth’s surface, though largely a local phenomenon, can alter global environmental conditions through atmospheric tele-connections and other large-scale land-water-atmosphere interactions, referred to as biogeographic effects. For instance, Avissar and Werth (2005) describe how, through the use of climate models, deforestation in the Amazon basin can alter precipitation in the temperate United States. Marchant and Hooghiemstra (2004) describe how vegetation changes and interactions with sea surface temperatures may have combined to create rather large-scale environmental change in the tropics nearly 4000 years ago.

2.1 Tree land cover and forest land-use

Tree cover and forest land-use are different and the distinction between them has large implications for the ability to measure or monitor each. This is especially true within the context of land management and when changes in each are considered over time. Tree cover is used to describe the Earth’s surface where the biophysical cover is tree life forms. Forest land-use is used to describe the Earth’s surface where the description of the human-ascribed use of the land surface is for forestry (IPCC, 2006).
The FAO definition of forest, for example, is primarily as a land use and is defined as land spanning more than 0.5 ha with trees higher than 5 meters and canopy cover of more than 10 %, or trees able to reach these thresholds in situ. It does not include land that is predominantly under agricultural or urban land use (FAO, 2010) even though this land may have tree cover as the dominant vegetation. Tree cover change is often characterized by ephemeral variations in the surface vegetation whereas changes in forest land-use imply a long-term shift in land cover and, potentially, significant effects to carbon fluxes (IPCC, 2006) and other biogeochemical cycles (Foley et al., 2005).

The differentiation of forest land-use is probably most important for assessing and accounting for management activities within forestland. The case of Canada is a popular example in this context. Hansen et al. (2010) report the country of Canada as having the second-highest amount of gross forest cover loss (GFCL) in the World between the years 2000 and 2005, behind only Brazil. The country of Canada, on the other hand, officially reports that they have almost no forest loss during this same time period (FAO, 2010). The definition of forest in the first study as a land cover and, in Canada’s official statistics, as a land use makes the interpretation of statistics for the same country produce completely opposite results. This is because the large majority of forest loss detected by remote sensing instruments in Canada is due to fire or occurs within areas managed timber harvest, both of which will ultimately re-grow to be again tree-covered forests. Mistaking the land use dynamic in Canada to be similar to that of Brazil based on tree cover loss statistics alone would lead to erroneous conclusions about the effects of forest loss in each.

Distinguishing between tree land cover and forest land-use areas is also important
for assessing post-forest-loss impacts on biogeochemical and biogeophysical cycles. Forest loss in areas managed either for forestry purposes (e.g. wood supply) or natural functions (national parks or forests) can be caused by a number of different agents, some purposeful, and vary in response and recovery depending on the time since forest removal and the ultimate use of the wood removed (Coulston et al., 2013). For example, the effects on the environment immediately following a forest fire are quite different than if those effects are temporally averaged over the course of the immediate post-burn (e.g. one to three years) and long-term recovery period (e.g. 60 to 80 years). Randerson et al. (2006) studied the effects of boreal forest fires on radiative forcing due to carbon dioxide, other greenhouse gas emissions and albedo. Immediately post-burn, radiative forcing was largely positive due to the large pulse of emitted greenhouse gasses. However, on a multi-decadal timeframe of approximately 55 years, the net radiative forcing was negative and thus had an overall atmospheric cooling effect. Howard et al. (2004) found, for a clear-cut stand of boreal forest jack pine (Pinus banksiana) that total ecosystem carbon content on the site increased over time with stand age, though the rate of increase was slower in older stands than younger. Ultimately, at the decadal time-scale and without including the effects of albedo, the stands post harvest fluctuated between carbon sources and sinks but ultimately were a slight net source of carbon dioxide emissions to the atmosphere. Lutz et al. (2015) show how managing a temperate forest in the northeast United States can be optimized for climate change mitigation by selecting optimal rotation periods which account for growing stock, albedo and the provision of timber. The authors found that the optimum rotation varied from 10 to > 200 years, depending on the site and species characteristics. In each of these examples the loss of
tree cover has varying effects, some negative others positive, depending on the reason for loss, the time since loss and the management practices in place. Identifying these areas and changes as forest land-use can go a long way towards understanding the future effects of forest loss and prioritizing forest losses as a detriment or aide in the provision of ecosystem services including mitigating the effects of global climate change.

2.2 Importance of consistent forest land-use estimates

Accurate information on the World’s forests has been critical since human beings began regularly utilizing wood as a resource. A demand for estimates of fuel-wood area spawned the first forest inventories in Europe in the 1500s (Brack, 1997). Since that time changing emphasis, including forest health, the supply of wood and non-wood forest products, habitat loss for critical species and the potential for forests to both contribute to and mitigate the effects of global climate change have driven the need for consistent and comparable estimates of forest and forestry parameters (McRoberts, Tomppo and Naesset, 2010). A constant demand for information and technological improvements have led to very advanced forest information systems containing an ever increasing number of variables (FAO, 2010) critical for informing management decisions that affect the provision of ecosystem services (Foley et al., 2005).

One of the most important variables collected on the World’s forest resources concerns the loss of forest, or deforestation. Deforestation is the long-term (permanent) conversion of forest to a non-forestry land use. Today, many internationally established sustainability targets depend on accurate forest land-use change statistics. The Convention on Biological Diversity’s (CBD) Target 5 of the Strategic Plan for
Biodiversity 2011-2020 specifically states that the rate of forest loss should be at least halved and, ideally, brought to zero by the year 2020 (CBD, 2011). Target 7 of the United Nations Millennium Development Goals, similarly, strives to reduce the loss of forest and biodiversity globally (United Nations, 2011). Quantifying atmospheric carbon fluxes associated with forest land-use change accurately and in a timely manner is critical for understanding and reducing the magnitude of global climate change and the central tenet of the United Nations Collaborative Programme on Reducing Emissions from Deforestation and Degradation in Developing Countries, or UN-REDD. UN-REDD works to implement the Kyoto Protocol’s call for reducing atmospheric carbon dioxide concentrations especially through forestry-related land-use change.

Currently, global estimates of forest land-use and change collected and provided by the United Nations Food and Agriculture Organization (FAO) serve as a major source of global forest land-use data. Since 1948, the FAO has analyzed and compiled data on the extent and state of the world’s forests through a process called the Global Forest Resources Assessment (FRA). Published every 5–10 years (in recent decades), the FRA report reflects the major issues of concern prevalent at the time of reporting. In response to post-Second World War needs, early FRAs focused on timber stocks, while more recent editions, including FRA 2010 (FAO, 2010), have addressed topics such as forest biodiversity, forest carbon stocks and the social benefits of forests. The FRA is an important information source for global efforts to sustainably manage forests, reduce the concentration of atmospheric greenhouse gases and advance other international initiatives. Data from FRA drive many of the models estimating carbon stocks, atmospheric fluxes and associated climate feedbacks (Houghton, 2010) and, according to
guidelines for national greenhouse gas inventories published by the Intergovernmental Panel on Climate Change (IPCC) (Paustian, Ravindranath and van Amstel, 2006), in the absence of detailed country-specific data, aggregate information can be obtained from international data sources such as the FRA (Penman et al., 2003). FAO collected data also provide a baseline and metrics of progress for many of the international agreements (Walpole et al., 2009).

Global forest area reports have, historically, proven challenging to use in the manner intended for international reporting requirements. For example, difficulty in assessing the long-term trend in forest area and change from the FRA reports (Grainger, 2008) contributes to uncertainty in the models used to estimate the effects of land-use change on atmospheric carbon flux (Houghton, 2010; Ramunkutty et al., 2007). Houghton (2010) documents that fluctuations in reported forest land area and rates of change over time, due largely to changes in reporting methods and national forest definitions, in FAO FRA reports between 2000 and 2005 result in a 32% difference in estimated net global carbon emissions. Clearly, a more consistent time series of estimates that can be reliably updated is important.

2.3 Estimating forest land-use from Landsat remote sensing data

The use of remotely sensed data affords two main methods of generating area estimates of land cover, land use and change: spatially exhaustive, or wall-to-wall, maps and sample-based. Passive optical remote sensing by the Landsat sensor detects the
Earth’s reflected electromagnetic energy in wavelengths from the visible to infrared (0.43 μm – 2.3 μm visible and infra-red) and emitted thermal energy (10.3 μm – 12.5 μm thermal). The spectral information detected by the sensor at the time of image acquisition is an instantaneous depiction of the Earth’s biophysical surface, acquired once every 16 days for any given location, from which land cover can be characterized. Exhaustive surveys of Earth’s surface with remote sensing are difficult, costly, time consuming and often, due to missing data (e.g. non-response), are not truly exhaustive. However, exhaustive surveys produce maps, which are easy to interpret and are useful for depicting the spatial arrangement of land cover or use types on a landscape for the purposes of spatially explicit forest management. Sampling allows inference about the characteristics of the Earth’s surface to be made without having to account for complete areal coverage. Sample-based estimates of forest statistics have a long history dating from the first forest inventories where plot-level information collected in the field is used to make inference about the condition of the entire stand (McRoberts, Tomppo and Naesset, 2010; Zon and Sparhawk, 1923). The key to sampling is the sample design and the careful interpretation of the sample to produce a set of reference data. Reference data collected through sampling are used in two main ways: (i) generating direct estimates of forest parameters or (ii) for assessing the accuracy and generating area estimates from maps. For a thorough explanation of using sample-based reference data for map accuracy assessment and area estimation see (Olofsson et al., 2013, Foody, 2002; Stehman, 2000; Stehman, 1998, Foody and Arora, 1997).

Sampling approaches for estimating land cover were once considered inadequate. The Food and Agriculture Organization (FAO) utilized a sampling approach with
remotely sensed satellite imagery to estimate the area and area change of forest and other land uses across the pan-tropics for the years 1980, 1990 and 1995 as part of two previous Global Forest Resource Assessments (FRA) (1996, 2001). The stratified random sampling approach yielded statistically valid results for tree cover area and change in area at the continental scale and, at the time, represented one of the first large-area estimates of tropical tree cover and tree cover change. The publication of the FRA reports in 1990 and 1995 sparked debate about whether or not a sample of remotely sensed data could actually be used to produce statistics with uncertainties low enough to be useful for generating area estimates. Tucker and Townshend (2002) suggested that only exhaustive classification of full Landsat scenes could provide adequate estimates. Czaplewski (2003) countered this result by proving that, indeed, a probability sampling rate of 10% could produce estimates with a high probability of equaling the ‘true’ amount of deforestation over large geographic areas. Czaplewski also showed that, for any given area and sampling rate, more but smaller samples were more likely to produce ‘true’ estimates of deforestation area than results obtained from fewer, larger samples.

Generating area and change estimates of tree cover from remotely sensed data using sample-based approaches are now relatively common at all spatial scales. Numerous studies exist which highlight the utility of sample-based estimates and the effects of sample design on precision (Stehman, 2005), the effect of sampling rate on precision at various spatial scales (Eva et al., 2010; Steinenger et al., 2010), the improvement in precision possible using stratification (Broich, 2009; Hansen, 2009; Stehman, 2003) and the improvement in precision possible when comparing design-based with model-based estimates (McRoberts, 2010).
Sampling to generate estimates of forest land-use and change has some critical advantages over spatially exhaustive mapping, especially when spatial explicitness is not a major priority. Land use is more difficult to discern than land cover when using satellite imagery to create maps or estimates as land use is not necessarily directly associated with land cover for any given site at the time of observation by the remote sensing instrument (IPCC, Consistent Representation of Lands) and, thus, demands additional inputs that cannot be automatically derived from many remote sensing-based analyses (Coulston et al., 2013). Given finite resources (e.g. time, money) sampling is generally cheaper and more careful observations can be made per-sample decreasing effects such as bias from measurement errors and producing theoretically more accurate parameter estimates than if one were surveying the entire population. Most importantly, the conversion of the Earth’s biophysical properties as detected from remotely sensed data sources into correct land use categories is complicated (Kurz, 2010; Lambin, 2001; Rogan and Chen, 2004). Sampling, which provides individual unit areas small enough to be carefully interpreted by human experts, may be the most effective method to facilitate conversion of land cover data into correct land use classes.

2.4 Summary

Characterizing and monitoring the amount of Earth’s area in forest land-use is important for estimating the consequences of and informing the decisions that affect land use change. This is especially true in an ever more human-dominated environment, where human beings decide what to do with and how to manage landscapes. The
chapters in this dissertation describe a sample-based approach to forest land-use area estimation, change detection and monitoring. Forest land-use extent and change are more difficult to characterize than tree land cover as they represent the associated functions of the land in terms of economic activity. Since land cover is the actual biophysical surface properties of the land surface, it is a more directly mapped phenomenon compared to land use. Determining land use from satellite imagery often requires interpretations from experts taking into account landscape pattern, context and other ancillary data sources. This fact has implications for methods aimed at quantifying land use extent and change. First, sample-based methods that provide an interactive opportunity for expert input, something not easily achieved in large area mapping exercises, offer a comparative advantage over mapping approaches. Samples can be analyzed in isolation and landscape-specific interpretations applied, unlike all-at-once mapping methods. Second, OBIA, which group pixels into homogenous units with some interpretable, ground-based significance (e.g. an agricultural field or a timber stand) are wholly appropriate as land use is a more generalized spatial theme than land cover. Land use typically does not vary at a per pixel scale, as does land cover. Interpreting landscape elements using OBIA methods is a promising approach for land use categorization. Finally, the classification and production of a global estimate of forest land-use, using the aforementioned techniques, is an extremely important first step in enabling the critical distinction between tree land cover and forest land-use; a distinction that allows more detailed assessment of the effects of forest area losses on biodiversity, biogeochemical and biogeophysical systems and, ultimately, on human wellbeing.
CHAPTER 3

THE SUITABILITY OF DECADAL IMAGE DATA SETS FOR MAPPING TROPICAL FOREST COVER CHANGE IN THE DEMOCRATIC REPUBLIC OF CONGO: IMPLICATIONS FOR THE GLOBAL LAND SURVEY.

This chapter was published as:

3.0 Abstract

Landsat remote sensing of the central African humid tropics is confounded by persistent cloud cover, and, since 2003, missing data due to the Landsat-7 Enhanced Thematic Mapper Plus (ETM+) scan line corrector (SLC) malfunction. To quantify these limitations and their effects on contemporary forest cover and change characterization, a comparison is made of multiple Landsat-7 image mosaics generated for a six Landsat path/row study site in central Africa for 2000 and 2005. Epoch 2000 mosaics were generated by compositing (i) two to three Landsat acquisitions per path/row, (ii) using the best single GeoCover 2000 acquisition for each path/row. Epoch 2005 composites were generated by compositing SLC-off data using (iii) five to seven acquisitions per path/row, (iv) three acquisitions per path/row. Eighty % of pixels were of suitable quality for change detection between (ii) and (iv), emulating that which is possible with current GeoCover and planned Global Land Survey inputs. In a more data intensive change detection analysis using mosaics (i) and (iii), 96% of pixels had suitable quality. Compositing more acquisitions per path/row for the study area systematically reduced the percentage of SLC-off gaps and, when more than three acquisitions were compositing, reduced the percentage of pixels with high likelihood of cloud, haze, or shadow. The results indicate that additional input imagery to augment both the Geocover and Global Land Survey data may be required to enable forest cover and change analyses for regions of the humid tropics.
3.1 Introduction

Quantifying the rates and spatial pattern of the fine-scale, tropical forest cover change observed in central Africa (Wilke and Laporte, 2001) is important for making land management decisions that affect biodiversity, biogeochemical processes and human health (IPCC, 2001; CBFP, 2005).

Remotely sensed regional land cover characterization at fine spatial resolution has typically utilized the Landsat sensor series (Townshend and Justice, 1988; Goward et al., 2001; Williams et al., 2006). The GeoCover global decadal Landsat data set is composed of single date acquisitions selected for each path/row from the 1970s, 1990s and 2000s but does not provide complete land surface observations due to persistent cloud (Tucker et al., 2004).

The humid tropics are particularly cloudy at the time of Landsat overpass (Ju and Roy, 2008) and cloud is limiting for many Landsat applications (Asner, 2001). For example, in the Congo Basin study area considered in this letter, 16% of the 2000 GeoCover data were cloud and cloud shadow contaminated. Contemporary studies using Landsat-7 ETM+ are further complicated by the May 2003 failure of the scan line corrector (SLC) that decreased the usable Landsat data by 22% without respect to clouds or other atmospheric contamination (Markham et al., 2004; Trigg et al., 2006). Consequently, unobscured, remotely-sensed observation of the humid tropics often requires multiple Landsat acquisitions.

For these reasons, the planned Global Land Survey (GLS) 2005 data set being developed by NASA and the USGS will generate a circa 2005 GeoCover-like data set by
compositing up to three low cloud cover Landsat acquisitions per path/row (Masek, 2007, Gutman et al., 2008). The Geocover and planned GLS data sets are of unquestionable value for many monitoring applications (Laporte et al., 2008). However, for exhaustive characterization of forest cover and change within the Congo Basin, improved data sets may be required (Hansen et al. 2008).

Recently, a decadal forest cover change mapping (DFCM) approach was developed to temporally composite best available pixels from multiple Landsat acquisitions and generate mosaics for forest classification and change detection analyses (Hansen et al., 2008). This letter examines the suitability of image inputs from the Geocover and planned GLS data sets compared to more intensive compositing methods using the DFCM approach.

3.2 Study Area and Data

The study area, defined by six adjacent Landsat path/rows (WRS PR 179059, 179060, 178059, 178060, 177059 and 177060, each about 185x185km), is situated in the north-central Democratic Republic of the Congo (DRC) within the ‘cuvette centrale’ of the Congo River Basin (Figure 3-1a). The area is characterized by low relief, meandering rivers and continuous dense Guineo-Congolian lowland tropical evergreen rain forest (White, 1983) with settlements connected by unsurfaced roads.
Figure 3-1. The study area (a), and 2005 epoch mosaics generated to illustrate DFCM compositing of (b) one, (c) two, (d) three, (e) four and (f) five Landsat-7 ETM+ acquisitions per path/row. Colour composites of Landsat-7 ETM+ bands 4, 5 and 7 are shown. Missing ETM+ SLC-off pixels are evident as white stripes and residual cloud contaminated pixels are evident in white, particularly in (b). Each mosaic is composed of more than 51 million 57m pixels.

The six GeoCover 2000 Landsat-7 acquisitions were obtained for the study area. In addition, Landsat-7 ETM+ data from two epochs (2000 and 2005) were selected based on minimal cloud cover as determined by metadata and browse imagery downloaded from GLOVIS (WWW1). A total of 14 Landsat-7 SLC-on acquisitions were selected from 2000 to 2003, providing two to three acquisitions per path/row for the 2000 epoch. A total of 36 SLC-off acquisitions were selected from 2004 to 2006, providing five to seven acquisitions per path/row for the 2005 epoch. These data were geometrically
registered to the Geocover data using an automated ground control point matching algorithm and bilinear resampling (Kennedy and Cohen, 2003). Visual examination of the coregistered data, focusing on regions containing distinct features, indicated that the data were misregistered by less than half a pixel, which did not significantly impact the subsequent compositing. Water bodies were removed from the analysis by applying an existing water mask (Hansen et al. 2008).

3.3 Methods

Epochal Landsat mosaics were created for both the 2000 and mid-decadal (2005) epochs using the DFCM compositing method described in Hansen et al. (2008). Every Landsat-7 pixel was assigned one of seven quality assessment (QA) values defined using classification trees and extensive training data applied to Landsat-7 bands 4 (0.78-0.90 mm), 5 (1.5-0.75 mm), 6 (10.4-12.5 mm), 7 (2.09-2.35 mm) and all combinations of possible two band simple ratios (Table 3-1). Composites were created by selecting pixels with the best quality as defined by the QA values. Pixels with no, or low likelihood of cloud/haze or shadow (i.e. QA values less than 4) are suitable for land remote sensing applications.
Table 3-1. Per-pixel QA values, the relative quality, and the defining characteristics of those values obtained as part of a decadal forest cover change mapping algorithm developed by Hansen et al. (2008), and used to create best-pixel temporally composited Landsat mosaics in central Africa.

<table>
<thead>
<tr>
<th>QA Value</th>
<th>Quality</th>
<th>Characteristic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Best</td>
<td>No cloud/haze/shadow</td>
</tr>
<tr>
<td>2</td>
<td>Good</td>
<td>Low likelihood of cloud/haze</td>
</tr>
<tr>
<td>3</td>
<td>Good</td>
<td>Low likelihood of shadow</td>
</tr>
<tr>
<td>4</td>
<td>Buffer/Poor</td>
<td>Adjacent to QA 5 or 6</td>
</tr>
<tr>
<td>5</td>
<td>Poor</td>
<td>High likelihood of cloud</td>
</tr>
<tr>
<td>6</td>
<td>Poor</td>
<td>High likelihood of shadow</td>
</tr>
<tr>
<td>7</td>
<td>Missing</td>
<td>Missing due to SLC-off</td>
</tr>
</tbody>
</table>

Epoch 2000 mosaics were compiled (i) using single image GeoCover data, (Tucker et al., 2004) and (ii) by compositing two to three Landsat-7 pre-SLC-off acquisitions per path/row. Epoch 2005 composites were generated (i) simulating the planned GLS approach using three Landsat-7 SLC-off acquisitions per path/row and (ii) by compositing five to seven Landsat-7 SLC-off acquisitions per path/row. Selection of the three acquisitions for the simulated GLS approach was undertaken by exhaustively computing which combination of three out of the available five to seven acquisitions provided the greatest number of composited pixels with QA values less than four.

The quality of the individual epochal mosaics was quantified by computing the percentage of pixels with different QA values. In order to assess the utility of the epochal mosaics for mid-decadal change detection, the percentage of pixels with QA values less than four occurring at the same pixel locations was compared for the 2000 and 2005 mosaics.
3.4 Results

Figure 3-1 (b-f) illustrates the results of the per-pixel DFCM compositing process, for the 2005 epoch mosaic generated by using one (b), two (c), three (d), four (e) and five (f) Landsat-7 SLC-off acquisitions per path/row. In each case, the Landsat acquisitions were selected by exhaustively computing which combination of the available five to seven acquisitions per path/row provided the greatest number of composited pixels with QA values less than four. For example, the mosaic produced using one acquisition (Figure 3-1b) was generated using the single Landsat acquisition for each path/row that had the least number of cloud, haze, and shadow pixels. As the number of acquisitions composited is increased, the percentage of SLC-off gaps decreases monotonically (from 22.5%, 3.8%, 0.5%, 0.2% to 0.07% respectively). This is because the spatial phase of the SLC-off gaps is not constant between acquisitions of a path/row (USGS, 2004). The percentage of cloud, haze, or shadow pixels increases from 2.1% to 4.6% and then decreases to 3%, 1.4%, and 0.7% as the number of acquisitions composited is increased from one to five respectively. This pattern occurs because clouds, haze, and shadows may persist at certain locations and pixels from additional acquisitions that fill SLC-off gaps may be atmospherically contaminated.

Figure 3-2 shows histograms of the QA values in each of the four epochal mosaics. The epoch 2000 mosaics were generated with Landsat data acquired prior to the SLC-off failure and so have no SLC-off gaps (QA value 7). Best quality pixels comprised 77% and 93% of the epoch 2000 GeoCover and DFCM mosaics, respectively. The percentage of best quality pixels increased to 86% and 96% of the epoch 2005 simulated
GLS and DFCM mosaics, respectively. The percentage of poor quality pixels totaled 16% and 3% in the 2000 Geocover and DFCM composites and 6% and 2% in the 2005 GLS and DFCM composites respectively.

Figure 3-2. Histograms of the percentage of pixels in the four Landsat epochal mosaics for each QA value. QA values in descending order of quality are: 1, no cloud/haze or shadow (i.e. best quality); 2, low likelihood of cloud/haze; 3, low likelihood of shadow; 4, adjacent to high likelihood cloud/haze or shadow; 5, high likelihood of cloud/haze; 6, high likelihood of shadow; 7, missing due to SLC-off gaps. The four Landsat epochal mosaics are labeled in the key as method, number of Landsat acquisitions used per path/row, and epoch period. See text for further details.

Figure 3-3 shows the percentage of good quality (QA value <4) and best quality (QA value 1) pixels occurring at the same pixel locations in the 2000 and the 2005 epochal mosaics. Specifically, the 2000 Geocover mosaic is compared with the 2005 simulated GLS mosaic (grey), and the 2000 and 2005 DFCM mosaics are compared
Increasing the number of image inputs in both epochs affects the amount of best and good quality data available for change detection. In a simulated GLS approach to change detection, utilizing the single acquisition Geocover 2000 mosaic and the three acquisition simulated GLS 2005 mosaic, 80% of pixels were of suitable quality (69% of which were best quality) for change detection. The percentage of suitable quality pixels increased to 96% (89% of which were best quality) using the two to three and five to seven DFCM composited 2000 and 2005 mosaics.

Figure 3-3. The percentage of good quality pixels (QA values 1, 2 and 3, i.e. low likelihood of cloud/haze or shadow) and best quality pixels (QA value 1, i.e. no cloud/haze or shadow) occurring at the same pixel locations in both epochal mosaics: grey shows a comparison of the 2000 Geocover (1 acquisition) and the 2005 simulated GLS (3 acquisitions) mosaics; black shows a comparison of the 2000 DFCM (2–3 acquisitions) and the 2005 DFCM (5–7 acquisitions) mosaics.
3.5 Discussion and conclusion

The results reported in this letter demonstrate that multiple Landsat acquisitions per path/row are needed to generate high quality composites for the humid tropical forests of central Africa. It is generally difficult however to establish the number of acquisitions required, due to spatio-temporal variation in cloud at the time of satellite overpass and the selective availability of Landsat acquisitions in many parts of the world, including the Congo (Ju and Roy, 2008). Due to this lack of knowledge and the reality of resource constraints, generic approaches have been suggested and applied to the processing of tropical decadal datasets: best single-date imagery in the case of GeoCover (Tucker et al., 2004) and best two to three date imagery in the case of the planned Global Land Survey Landsat-7 SLC-off processing (Masek, 2007).

This letter has demonstrated, for a limited study over the central African humid tropics, that the planned Global Land Survey (GLS) approach of compositing up to three low cloud cover Landsat-7 ETM+ acquisitions for each path/row will produce composites that have minimal cloud, haze and shadow contamination and minimal SLC-off gaps. However, when comparing composites from different epochs for change detection purposes, SLC-off gaps and cloud, haze and shadow contaminated pixels may not occur at the same locations, and the resulting number of pixels useful for change detection will be reduced. The current planned option for contemporary mid-decadal change detection by comparison of GeoCover and GLS data sets, for the study area, leaves 20% of the pixels unsuitable for change detection, which may preclude meaningful analysis of humid tropical forest cover change. This limited study shows that use of 2000 and 2005 epochal
mosaics composited using more Landsat acquisitions per path/row (two to three in 2000 and five to seven in 2005) results in only 4% of the pixels remaining unsuitable for change detection. This reinforces the concept that, ideally, all of the data in the Landsat archive should be used to overcome the prevalence of cloud contamination, SLC-off gaps and other deleterious remote sensing variations (Hansen et al., 2008; Roy et al., 2008). Instead of using a generic number of image inputs per path/row, a minimum data quality threshold could be defined and the Landsat data archives mined until the threshold is met. This approach may become feasible when the current USGS plans to open up the Landsat-7 archive for free digital download are realized and when this data policy is more broadly adopted by other international satellite data providers.

**Acknowledgements**

Funding from the National Aeronautics and Space Administration supported this research under grants NNX07AO57H and NNG06GC41G. The comments of three anonymous reviewers are also appreciated and greatly improved this manuscript.
CHAPTER 4

ASSESSING GLOBAL FOREST LAND-USE CHANGE BY OBJECT-BASED IMAGE ANALYSIS

This chapter has been submitted for peer review in the journal *Remote Sensing* as Lindquist, E.J., D’Annunzio, R., Assessing Global Forest Land-use Change by Object-based Image Analysis.
4.0 Abstract

Consistent estimates of forest land-use and change over time are important for understanding and managing human activities on the Earth’s surface, parameterizing models used for global and regional climate change analyses and a critical component of reporting requirements faced by countries as part of the international effort to Reduce Emissions from Deforestation and Degradation (REDD). In this study, object-based image analysis methods were applied to a global sample of Landsat imagery from years 1990, 2000 and 2005 to produce a land cover classification suitable for expert human review, revision and translation into forest and non-forest land-use classes. We describe and analyse here the derivation and application of an automated, multi-date image segmentation, neural network classification method and independent, automated change detection procedure to all sample sites. The automated results were compared against expert human interpretation and found to have an overall agreement of ~76% for a 5-class land cover classification and ~88% agreement for change / no-change assessment. The establishment of a 5 ha minimum mapping unit affected the ability of the segmentation methods to detect small or irregularly-shaped land cover change and, combined with aggregation rules that favor forest, added bias to the automated results. However, the OBIA methods provided an efficient means of processing over 11,000 sample sites, 33,000 Landsat 20x20 km sample tiles and more than 6.5 million individual polygons over three epochs and adequately facilitated human expert review, revision and conversion to a global forest land-use product.
4.1 Introduction

In an effort to produce a set of spatially consistent and comparable statistics on global tree cover, forest area and change, the United Nations Food and Agriculture Organization (FAO) in collaboration with the Joint Research Centre of the European Commission (JRC) used object-based image analysis (OBIA) techniques and remotely sensed satellite imagery to implement a sample-based survey of the Earth’s land surface called the Global Forest Resource Assessment (FRA) 2010 Remote Sensing Survey. Forest land-use change was estimated at global, regional and ecological domain scales for the time period 1990 – 2005 (Lindquist et al., 2012).

OBIA is increasingly used to classify remotely sensed data (Blashke et al., 2014) in a process that combines image segmentation techniques as an integral part of the classification. Image segmentation is the process of combining the individual picture elements (pixels) of raster data into meaningful objects for identification purposes (Blashke et al., 2010). Merging pixels of similar spectral and proximal spatial properties together and assigning a common label accomplish this.

Object-based methods have been frequently used in land cover classification. Classic methods of achieving this result include the unsupervised clustering algorithms of ISODATA (Ball and Hall, 1965) or k-means (MacQueen, 1967), in which pixels with similar spectral qualities are assigned a common label depending on rules specified by the user including the total number of classes desired in the output image. Though segmentation algorithms have advanced since, the basic premise remains the same; to delineate areas on the Earth’s surface meaningful to the purposes of analysis. Hussain et
al. (2013) provide a thorough review of object-based classification and change
detection methods and applications. Viera et al. (2012) used image segmentation and
data mining algorithms to successfully classify patterns of agriculture in Brazil; Eva et al.
(2010, 2012) used image segmentation techniques to assess forest cover, forest cover
change and associated carbon dioxide emissions in South America; Duveiller et al. (2008)
also used image segmentation of medium spatial resolution imagery to classify land cover
and change for a sample-based assessment in the Congo Basin; Ernst et al. (2010) also
assessed forest cover and change for the Congo Basin using medium spatial resolution
data; Mayaux et al. (2013) describe the results of a forest cover change assessment, also
for the humid tropics of Central Africa, obtained from a sample-based assessment of
segmented medium resolution imagery; Rasi et al. (2011) describe the use of OBIA
techniques for classifying medium spatial resolution imagery pan-tropically; Brink et al.
(2009) describe the results of a land cover change detection for parts of Africa over 25
years using OBIA methods; Bodart et al. (2013) estimated tree cover change in dry
Africa from 1990 to 2000 using OBIA and medium spatial resolution imagery.

Image segmentation of medium spatial resolution imagery for the purposes of
delineating forest cover and forest cover changes, however, may be subject to problems
associated with spatial scale and associated effects on detectability of change (Woodcock
and Strahler, 1987). Landsat pixels represent an arbitrary grid with 30 x 30 meter cells in
which spectral reflectance from the surface of the Earth is recorded. Thus, pixels by
themselves are not necessarily optimally placed to represent identifiable objects (Marceau
and Hay, 1999; Yu et al., 2006) and, depending on the spatial resolution of the input
satellite imagery, a single pixel can represent anything from a portion of a tree crown (e.g.
for very fine spatial resolution imagery) to many trees (e.g. for medium spatial resolution imagery) to entire stands or stand complexes (e.g. for coarse spatial resolution imagery).

Within-class spectral heterogeneity leads to difficulties in both land cover classification and change detection (Woodcock and Strahler, 1987; Marceau and Hay, 1999). Ideally, image segmentation will minimize spectral heterogeneity within objects. However, as object size increases so too does heterogeneity (Blashke et al., 2010). In the case of detecting forest cover changes in Landsat imagery using segmentation techniques, it may be that the scale of the change occurs at finer levels than the minimum size of the image segment, especially if the segments are coerced to an arbitrary minimum mapping unit.

This paper describes and analyses the OBIA methodology used by the FAO to classify land cover and land use for over 11,000 globally distributed, satellite image-based sample sites. We provide an analysis of the image segmentation and classification and we revisit the problem of scale in satellite image classification (Woodcock and Strahler, 1987), namely the selection of a minimum mapping unit, its potential addition of bias to the results and the effectiveness of image segments coerced to a pre-defined minimum mapping unit at detecting land cover changes of different shapes and sizes when using medium spatial resolution imagery. Finally, we examine the differences between tree cover and forest land-use, the practicalities of classifying both and the difference in global forest area when they are considered separately.

4.2 Materials and Methods
The United States Geological Survey’s (USGS) Global Land Survey dataset (GLS) was used as input imagery in this assessment. The GLS dataset covers most of the Earth’s land surface and is composed of the single best Landsat 4 or Landsat 5 image acquisition for the years 1990, the single best Landsat 5 acquisition for year 2000 and the single best Landsat 7 acquisitions for the years 2005 and 2010 (Gutman et al., 2008). This study considers GLS data from 1990, 2000 and 2005 only.

A sampling design with a site at each 1-degree intersection of latitude and longitude was employed, except in Canada (see Lindquist et al. 2012 for details). Sampling intensity was reduced above 60 degrees latitude north and south to include only even degrees of longitude. No sample sites were located higher than 75 degrees north or south latitude. At each sample site, Landsat 30 x 30 m optical bands 1-5 (0.45 – 1.75 µm) and 7 (2.09 – 2.35 µm) were subset to a central 20km by 20km box (Beuchle et al., 2011; Potapov et al., 2010). The GLS data were assumed to be the best data available for each site. If more than one GLS acquisition was available for a given site and date the GLS acquisition with the least cloud cover was selected for classification.

The JRC, as part of its ongoing Tropical Ecosystem Environment Monitoring by Satellites or TREES (Achard et al., 2002) and FOREST (JRC, 2012) forest monitoring programmes, processed approximately 4,700 pan-tropical and western European survey sites, respectively (Rasi et al., 2011, Bodart et al., 2011). The FAO processed sites located in the sub-tropical, temperate and boreal regions of the Americas, Asia, Europe including Russia and Oceania (Figure 4-1). The entire global sample grid of 15,770 sites is equivalent to a 1% sample of the Earth’s land surface.
4.2.1 Image segmentation

Image segmentation is the process of partitioning an image by grouping pixels into clusters, called objects, based on intra-object spectral similarity and inter-object spectral differences. In this study, a region-growing multi-resolution image segmentation algorithm was used in which the criteria for creating image objects from individual pixels can be adjusted by specifying values for a series of parameters which control (i) object average size, called the scale factor, (ii) object spectral homogeneity, called the shape factor and (iii) object boundaries, called the compactness factor (Baatz and Schape, 2000).

Landsat image bands 3, 4 and 5 (0.63 – 1.75 µm) from all three time periods were used in a multi-date segmentation routine to create segments capturing spatially
contiguous areas of similar spectral response and areas with unique temporal spectral signatures (Desclee, Bogaert and Defourny, 2006). These particular Landsat bands were selected for two main reasons: their ability to discriminate differences in surface reflectance caused by changes in vegetation cover (Desclee, Bogaert and Defourny, 2006; Duveiller et al., 2008). A parsimonious selection of image bands for segmentation generally benefits the quality of the segmentation and reduces the chances of over-segmentation (e.g. creating many more segments than are necessary to define objects of interest) (Mesner and Ostir, 2014).

Image segmentation was a two-step process, referred to in this paper as level-1 and level-2. Parameters for the level-1 segmentation were selected to create relatively small, spectrally homogenous objects. The same segmentation parameters (scale = 15, shape = 0.1, compactness = 0.5) were fixed for all sample sites as no a priori information about the sample sites was assumed. Level-2 objects were created to meet a minimum mapping unit (MMU) requirement of five hectares; chosen to allow resolution of relatively small forest cover changes and maintain a manageable number of image segments (Ridder, 2007). To meet the MMU requirement, level-1 segments less than five hectares in size were merged with adjacent objects with the most similar average Landsat band 5 (1.55 – 1.75 µm) short-wave infrared reflectance (Horler and Ahern, 1986; Hoffhine and Sader, 2002) until the minimum size requirement was met.
Figure 4-2. (From Lindquist et al., 2012) A multi-date segmentation for a 20km x 20km site in the boreal climatic domain. Landsat imagery from 1990, 2000, 2005 are combined into a single data stack and segmented. The segments (white outlines in image on right) capture areas of spectral similarity as well as areas of change over time. The Landsat image on the right is a color composite of Landsat band 5 from 1990 (red), 2000 (blue) and 2005 (green). Changes in tree cover between time periods are evident.

4.2.2 Automated Image Classification

For each site and date, a land-cover classification was produced with five classes: tree cover, shrub cover, other land, water and no data. Tree cover was defined by woody vegetation greater than five meters in height. Shrub cover was defined as woody vegetation less than five meters in height. Other land cover was defined as all other land cover types, including herbaceous and bare land. Clouds, cloud shadow and missing data were given a ‘no data’ label. Figure 4-3 illustrates the processing method applied to a
Figure 4-3. The multi-date processing scheme showing for a boreal forest sample site in year 1990 (top row), year 2000 (middle row) and year 2005 (bottom row) the following: (A) the 20x20km Landsat imagery (B) the potential training objects (grey) for year 2000, (C) the training dataset after class labeling, (D) the final automated land cover classification, (E) the final, expert revised land cover classification of the central 10x10km (corresponding to white box in D), and (F) the corresponding final, expert revised land use classification. Dark green = Tree Cover / Forest, Light green = Tree Cover Mosaic, Orange = Shrub / Other Wooded Land, Pink = Other Land.

Level-1 image objects were classified first for year 2000 using a supervised approach with training data collected from the year 2000 Landsat data. Training data for all land cover classes were selected automatically utilizing temporally and spatially coincident land cover data from the Moderate Resolution Imaging Spectroradiometer (MODIS) 250 m spatial resolution Vegetation Continuous Fields (VCF) product (Hansen, 2003), the year 2005 300 m spatial resolution GLOBCOVER global land-cover product (Arino et al., 2008) and the MODIS global 250m water mask (Carroll et al., 2009).

A set of potential training objects were identified from all possible level-1 image
objects by first flagging those most likely to represent either unambiguous tree cover or unambiguous non-tree cover (e.g. other land). Per-object class homogeneity was defined in relation to the range of VCF percent canopy-cover values circumscribed by the object (Figure 4-3, column B). An initial VCF range threshold value of 10 % (e.g. Max VCF – Min VCF <10) was chosen. To ensure the availability of ample training data, if less than 30 % of the total number of image objects were flagged using the initial VCF threshold, the threshold was increased by a value of five and the selection process repeated. Iterations continued until at least 30 % of the total image objects were identified as the potential training dataset.

Next, class labels were assigned to the training dataset using a simple rule set consisting of a combination of 2000 MODIS VCF, 2005 GLOBCOVER, MODIS Surface Water Bodies land-cover products and hard-coded Landsat digital number thresholds (Table 4-1, and Figure 4-3, column C). Segments indicated as bare ground or cloud cover were labelled as such and removed from the training data set. Clouds were hard-coded as no data and bare ground was hard-coded as other land in the final land cover classification.

Table 4-1. Training segment labelling rules based on underlying MODIS VCF, GLOBCOVER, MODIS Water Mask and hard-coded Landsat DN thresholds.

<table>
<thead>
<tr>
<th>Land Cover Class</th>
<th>Segment Characteristic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>Segment proportion of Global Water Mask &gt; 0.7</td>
</tr>
<tr>
<td>Tree Cover</td>
<td>Mean segment VCF % tree cover &gt; 11</td>
</tr>
<tr>
<td>Shrub Cover</td>
<td>Mean segment VCF % tree cover &lt; 11 AND Globover Class = Shrubland</td>
</tr>
<tr>
<td>Herbaceous Cover</td>
<td>Mean segment VCF % tree cover &lt; 11 AND Globover Class = Herbaceous OR Landsat Band 5 DN &gt; 120</td>
</tr>
<tr>
<td>Bare ground</td>
<td>Mean segment Landsat Band 2 DN &gt; 200 AND Landsat Band 5 DN &gt; 200</td>
</tr>
<tr>
<td>Cloud</td>
<td>Mean segment Landsat Band 4 DN + Band 5 DN + Band 7 DN &gt; 500</td>
</tr>
</tbody>
</table>
4.2.3 Neural network land cover classification

A series of artificial neural network (ANN) classifiers were used to create the final land-cover classification. Neural networks consist of input layers, a layer of processing nodes, referred to as the ‘hidden layer’, and output layers. The generalization power of neural networks is well suited for supervised vegetation classification with non-parametric data and sub-optimal training (Foody, 2000; Foody and Arora, 1997; Mas and Flores, 2008; Yuan, VanDerWiele and Khorram, 2009). A simple feed-forward ANN with back-propagation is implemented within the E-Cognition software through the Neural Network Plugin from Freiberg University (Bachmann, 2009).

Input layers to the classification consisted of training class labels, Landsat image bands 3 (0.63 – 0.69 µm), 4 (0.77 – 0.90 µm), 5 (1.55 – 1.75 µm) and 7 (2.09 – 2.35 µm) and all possible simple 2-band ratios. A simple ratio of Landsat bands 2 (0.52 – 0.60 µm) and 7 (2.09 – 2.35 µm) was also included, making a total of 11 input layers. Band ratios were used to improve discrimination of forest, non-forest and deforestation. Hansen et al. (2008) used simple 2-band ratios to aid classification of forest cover and deforestation in Central Africa. The number of nodes in the hidden layer was arrived at empirically and numbered 16.

ANNs were applied sequentially to the year 2000 image objects first to distinguish water from other land. Other land was then divided into herbaceous vegetation and woody vegetation. The woody vegetation objects were split into tree cover and other wooded land classes. In each model, two output nodes were established,
corresponding to the number of desired land-cover classes (e.g. water or other, herbaceous or woody, tree or other wooded land). Each time, the network was allowed to train itself 1000 times per class using a back-propagation learning method. Once the training of the network was complete (i.e. when the error was minimized or the specified number of iterations reached) the model was applied to all level-1 image objects within the scene.

ANN performance can be affected by the quality and amount of data used to train the model (Foody, 2000; Yuan et al., 2009). To ensure neither over nor under-training the model, the total number of training observations per site was restricted to approximately 2000 (e.g. roughly half the total average number of level-1 image segments per sample). Each training class was reduced and harmonized by proportional random sampling until the total number of samples reached the total allowable number of training observations. If a training class had less than 100 observations to begin with, all observations were used to construct the final model (Yuan, VanDerWiele and Khorram, 2009; Foody and Arora, 1997).

4.2.4 Multivariate Alteration Detection for land-cover change

Change between time periods was assessed independently from the land cover classification by employing the iteratively re-weighted multivariate alteration detection (IR-MAD) algorithm (Nielsen, Conradsen and Simpson, 1998; Nielsen, 2007). The IR-MAD algorithm is an extension of canonical correlation analysis that maximizes the information on change over all variables (i.e. satellite image bands) considered in the analysis. The algorithm is insensitive to spectral changes that occur between co-located,
bi-temporal data due to changing sensor characteristics like gain and bias, linear data
normalization or calibration and affine transformations. IR-MAD improves on previous
multivariate change detection algorithms (Nielsen, 2005) by iteratively increasing the
weight of non-change observations and theoretically improving the discrimination of
substantive spectral changes between time periods. IR-MAD was implemented within
the Ecognition 8.0 software as the MAD-Transformation plugin (John and Bachman,
2009).

Landsat bands 3 (0.63 – 0.69 µm), 4 (0.77 – 0.90 µm), 5 (1.55 – 1.75 µm) were
used in an IR-MAD analysis to label image objects with spectral changes between time
periods. The first order MAD variate was computed from the selected Landsat bands and
their differences between 1990-2000 and 2000-2005. Image objects in year 1990 and
year 2005 more than 0.25 standard deviations away from the overall site-wide mean for
the MAD variate were flagged as having a high likelihood of land cover change between
the survey periods.

4.2.5 Assigning class labels to 1990 and 2005 time periods

Land cover labels were assigned to 1990 and 2005 image objects based on
whether or not the object represented a potential land cover change. Unchanged 1990
and 2005 image objects were directly assigned the land cover label from year 2000. Year
1990 and 2005 image objects exhibiting a likely change in land-cover, relative to year
2000, were assigned a land cover class by supervised classification using training data
taken from the unchanged image objects from each time period, respectively (Figure 4-3,
column C). Training set reduction and ANN application were as for year 2000.

Level-2 image objects (with a 5-hectare MMU) were assigned land-cover labels based on the underlying level-1 segments according to the rules in Table 2.

Table 4-2. Level-2 segment labelling rules for the automated land cover classification. Level-2 labels were based on the areal proportion of component level-2 segments.

<table>
<thead>
<tr>
<th>Level 1</th>
<th>% composition</th>
<th>Level-2 land-cover</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tree cover</td>
<td>≥ 30</td>
<td>Tree cover</td>
</tr>
<tr>
<td>Other wooded land</td>
<td>≥ 70</td>
<td>Other wooded land</td>
</tr>
<tr>
<td>Other land cover</td>
<td>≥ 70</td>
<td>Other land cover</td>
</tr>
<tr>
<td>Water</td>
<td>≥ 70</td>
<td>Water</td>
</tr>
</tbody>
</table>

The final land cover classification for the central 10x10km portion of each sample was subject to manual expert review and revision by image interpreters with local knowledge to correct classification errors caused by the automated procedure (Figure 4-3, column D). The final, revised land cover labels (Figure 4-3, column E) were ultimately converted to land use categories and the experts were again employed to modify the land use labels where necessary (Figure 4-3, column F) (See Lindquist et al., 2012 for conversion details).

4.3 Results and Discussion

4.3.1 OBIA, minimum mapping units and land cover classification

The segmentation routine and enforcement of a 5 ha MMU represents a trade-off between segment homogeneity, data volume and ease of use by expert image interpreters.
(Blashke et al., 2014; Blashke, 2010). However, increasing the heterogeneity of image segments to meet an *a priori* MMU can make classification difficult and introduce bias in the estimates of area and area changed.

A sub-sample of sites was analysed to explore the effect of the segmentation routine on object size and data volume. The initial, level-1 segmentation produced image objects from 900 sq. m. (a single pixel) to > 125 ha in size. The mean area of level-1 polygons was 11.8 ha. The level-2 segmentation, with a MMU of 5 ha, had a mean polygon size of 17 ha. The merging of level-1 to level-2 polygons did not change the relative size distribution, with a large portion of the polygons remaining < 15 ha in size. The level-2 segmentation, in which small, spectrally homogenous areas were merged with larger, neighboring polygons, decreased the total number of polygons by 30%.

Increasing the size of the image objects to meet the MMU combined with the MMU segment labelling rules in which any level-2 object containing > 29 % tree cover resulted in a small bias amounting to a 3% increase in estimated area of tree cover when compared against the level-1 segment results. This bias was removed after the expert review and revision, however the expert review increased the total area of tree cover by 3.5% over the initial level-1 classification. In this sub-sample, expert revision revealed that the automated algorithm produced a slight overestimation of tree cover where tree cover is low and a slight underestimation of tree cover where tree cover is high (Figure 4-4).
Figure 4-4. (Left) A comparison of the tree cover area obtained automatically for level-1 segmentation (x-axis) and the 5 ha MMU (y-axis) and (Right) a comparison of the tree cover area obtained automatically at the 5 ha MMU (x-axis) and the expert reviewed-revised tree cover at the 5 ha MMU (y-axis). The dotted line represents the slope of the line formed by the linear regression. The solid line represents the one-to-one line.
Figure 4-5. The amount of change detected between 1990 - 2000 using IR-MAD and increasing change thresholds (columns) aggregated to level-1 and 5 ha MMU segments (rows) for a 20x20 km sample site. Columns left to right represent thresholds of 0.91, 0.95 and 0.99 of the IR-MAD change likelihood layer, respectively. Rows top to bottom represent pixel-level, level-1 image segments and 5 ha MMU segments, respectively. Changed areas are in black. Non-changed areas are white.

4.3.2 OBIA, minimum mapping units and change detection

To analyse the effects of increasing segment size on amount of area detected as
change, an automated, per-pixel change detection was performed for a sub-sample of sites employing the IR-MAD algorithm and the Landsat samples from 1990 and 2000. A series of nine change likelihood thresholds in one % increments, ranging from 91 to 99 % (Nielsen, 2007; Canty and Nielsen, 2006), were used in the iterations. Pixels with change likelihoods above the threshold in each run were labelled as change, all other pixels were labelled as no-change (Figure 4-5 top-row). Changed pixels were then aggregated by level-1 and 5 ha MMU segments using a simple majority rule (i.e. if > 50 % of a segment contained change, the segment was labelled as change) (Figure 4-5, middle and bottom row, respectively). The difference in area changed for each iteration of the IR-MAD algorithm and for each segmentation level was analysed.

Logically, as the threshold of the IR-MAD algorithm increases and change pixels are more selectively labelled, the mean area of change decreases. The difference in the amount of change detected between the pixel-based and segmentation aggregations also decreases. However, the aggregation to segments results in an overall decrease in the amount of change detected between 25 and 50 % when compared to the pixel-level analyses (Figure 4-6).

4.3.3 Landscape metrics and change detection

The difference in area change per level of aggregation could be explained by the spatial arrangement and areal extent of the detected change in each of the iterations. Landscape metrics were generated for the pixel-based change layer using the R statistical package SDMTools (VanDerWal et al., 2011). Landscape metrics describe the pattern of
land cover classes on a landscape and are commonly applied in ecological studies (McGarigal and Marks, 1995). Though Riitters et al. (1995) emphasized that landscape metrics must be cautiously applied as many are co-variates and suggested six main landscape metrics that are the most suitable for describing landscapes with real implications.

To determine if the size and shape of areas detected as change had an effect on the amount and type of change detected in the image segments, we used two of Riitters et al. (1995) recommended landscape metrics, mean patch size and mean landscape shape index, and calculated these based on the pixel-level change detection. Then, we compared the difference in area between change detection iterations and aggregations as a function of each of the calculated metrics. Mean patch size is the mean area of all change patches within the sample site. Mean landscape shape index is the relative perimeter-to-area ratio of the landscape (i.e. values close to 1 = compact, values increase as patches become disaggregated and amount of edge increases). The difference in area detected as change between the iterations and aggregation levels varies inversely with mean change patch area (e.g. as the size of change patch increases, the difference in area detected at aggregated spatial levels decreases) (Figure 4-6 - right). Figure 4-6 (center) shows there is a strong linear relationship (R2 = 0.9) between increasing landscape shape index and increased differences in the total area of change detected between the pixel-based and object-based estimates, indicating that as change patches become more disaggregated and have a higher perimeter-to-area ratio, they are more difficult to detect in segments meeting the MMU requirement.
Figure 4-6. (Left) The mean area detected as change (Y-axis) over all samples using pixel-based (triangles), level-1 segments (+) and level-2 segments (circles) at increasing IR-MAD thresholds (X-axis). (Center) The difference in the total area detected as change between pixel and level-2 segments plotted as a function of the change patch landscape shape index. (Right) The difference in the total area detected as change between pixel and level-2 segments plotted as a function of the mean change patch area.

Small or complex shapes are more difficult to detect in the 5 ha MMU image segments. Whether or not these potential omissions have significant meaning on the landscape, however, is unknown. It is intuitive that patches of change smaller than the MMU are largely lost when the analysis shifts from pixel-based to segment-based. Also true, however, is that patches of change smaller than the MMU remain detectable at the segment level, indicating that some segments labelled as ‘change’ actually overestimate the actual area changed. The figure shows, then, that the total area of a change patch is not the only indicator of whether it is detectable or not at aggregated spatial scales. There is a strong relationship between the landscape shape index and the difference in total area changed between the iterations. This indicates that the complexity of the patch shape has a high degree of influence on its detectability. More compact shapes (i.e. clear-cut timber harvest) are thus more easily detectable than more complex shapes, which are often characterized by large perimeter to area ratios (i.e. logging roads, selective logging,
small-scale agriculture).

4.3.4 Automated land-cover classification and change detection results

To determine the effectiveness of the classification and change detection routine at facilitating the efficient manual review and revision of the samples, the automated results were compared against the expert reviewed and revised land cover labels. For a randomly selected subset of segments, the automated land cover classification achieved an overall agreement of 76% for all five classes in year 2000 (Table 4-3). Similar agreements were obtained for years 1990 and 2005.

Table 4-3. Confusion matrix between land cover classes before and after expert review and revision for a sample of year 2000 sites. Figures are based on area in each class in millions of hectares. Automatically mapped class labels are on left-hand side and the expert reviewed and revised labels are across the top of the table. Tree = Tree cover, Wooded = other wooded land, Other = other land cover.

<table>
<thead>
<tr>
<th></th>
<th>Tree</th>
<th>Wooded</th>
<th>Other</th>
<th>Water</th>
<th>No Data</th>
<th>Total</th>
<th>User's</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tree</td>
<td>6957</td>
<td>802</td>
<td>1599</td>
<td>61</td>
<td>33</td>
<td>9451</td>
<td>0.74</td>
</tr>
<tr>
<td>Wooded</td>
<td>564</td>
<td>1609</td>
<td>2336</td>
<td>10</td>
<td>2</td>
<td>4521</td>
<td>0.36</td>
</tr>
<tr>
<td>Other</td>
<td>580</td>
<td>793</td>
<td>12684</td>
<td>28</td>
<td>21</td>
<td>14107</td>
<td>0.90</td>
</tr>
<tr>
<td>Water</td>
<td>17</td>
<td>24</td>
<td>92</td>
<td>705</td>
<td>3</td>
<td>840</td>
<td>0.84</td>
</tr>
<tr>
<td>No Data</td>
<td>1</td>
<td>3</td>
<td>40</td>
<td>2</td>
<td>103</td>
<td>149</td>
<td>0.69</td>
</tr>
<tr>
<td>Total</td>
<td>8120</td>
<td>3230</td>
<td>16751</td>
<td>805</td>
<td>162</td>
<td>29069</td>
<td></td>
</tr>
<tr>
<td>Producer's</td>
<td>0.86</td>
<td>0.50</td>
<td>0.76</td>
<td>0.87</td>
<td>0.63</td>
<td>0.76</td>
<td></td>
</tr>
</tbody>
</table>
The automated change / no-change mapping achieved an overall agreement of 88% between 1990 and 2000 and 87% between 2000 and 2005 compared to the expert reviewed and revised classification. The relatively high overall agreement is due to the dominance of unchanging classes. Producer’s agreement for actual change classes ranged from 68% to 77% in both comparisons. User’s agreement for actual change classes ranged from 11% to 36%.

Table 4-4 shows the confusion matrix generated for each survey period by comparing the results of the automated classification (original) and the expert reviewed and revised land cover dataset (validated).

Table 4-4. Confusion matrix for change classes for the time periods 1990 – 2000 (top) and 2000 – 2005 (bottom) before and after expert review and revision. Figures are based on area in each class in millions of hectares. Automatically mapped class labels are on left-hand side and the expert reviewed and revised labels are across the top of the table. Tree-Tree = Tree cover in both periods, Tree - Other = Tree cover change to other land, Other - Tree = other land to tree cover, Other - Other = other land cover in both periods.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total</strong></td>
<td><strong>User's</strong></td>
<td><strong>Producer's</strong></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>5934</td>
<td>16273</td>
</tr>
<tr>
<td><strong>User's</strong></td>
<td>0.75</td>
<td>0.98</td>
</tr>
<tr>
<td><strong>Producer's</strong></td>
<td>0.88</td>
<td>0.73</td>
</tr>
<tr>
<td><strong>Tree – Tree</strong></td>
<td>4447</td>
<td>4185</td>
</tr>
<tr>
<td><strong>Tree – Other</strong></td>
<td>63</td>
<td>49</td>
</tr>
<tr>
<td><strong>Other – Tree</strong></td>
<td>33</td>
<td>37</td>
</tr>
<tr>
<td><strong>Other – Other</strong></td>
<td>1391</td>
<td>1449</td>
</tr>
<tr>
<td><strong>Other – Other</strong></td>
<td>922</td>
<td>5720</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>5934</td>
<td>16273</td>
</tr>
<tr>
<td><strong>User's</strong></td>
<td>0.75</td>
<td>0.98</td>
</tr>
<tr>
<td><strong>Producer's</strong></td>
<td>0.88</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td></td>
<td>69</td>
<td>226</td>
</tr>
<tr>
<td>Tree – Other</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other – Tree</td>
<td>119</td>
<td>2</td>
</tr>
<tr>
<td>Other – Other</td>
<td>182</td>
<td>19</td>
</tr>
<tr>
<td>Total</td>
<td>4555</td>
<td>296</td>
</tr>
<tr>
<td>Producer's</td>
<td>0.92</td>
<td>0.76</td>
</tr>
</tbody>
</table>

Land cover and land-use change are statistically ‘rare’ phenomena (Stehman, Sohl and Loveland, 2005) and, as discussed previously, can be difficult to characterize. However, the low levels of agreement found in the study are lower than desired for the automated classification routine, which in all cases overestimated the amount of change. For a sub-sample of FAO-processed sites in boreal, temperate and subtropical regions, automated results overestimated tree cover loss by 11% and tree cover gain by 13.5% compared to the final expert reviewed and revised results for the time period from 1990 to 2005. The lowest user’s agreement was consistently found in the ‘other land to forest’ category suggesting that forest re-growth was the most difficult land cover change to accurately detect.

The consistent overestimation of tree cover change in this study is likely caused by several factors including the tendency of the classification rules to favor tree cover over other land cover classes and a rather liberal threshold of the IR-MAD variable used to indicate whether or not an image object represented land cover change. The results suggest that a stricter rule can safely be applied to flag potentially changed image objects in future iterations of the survey. It may also be advisable, in areas where finer-scale changes are known to occur or are suspected, to selectively ease the minimum mapping unit restriction in order to avoid over or under-estimating change.
4.3.5 Land cover v. land use

The conversion from land-cover class to land-use class is a two-step process and required the input of expert human interpretation. First the land-cover classes were automatically converted into one of four land-use classes; (i) forest, (ii) other land use with tree cover, (iii) other wooded land and (iv) other land use. For example, all objects with the ‘tree cover’ land cover were automatically re-labeled as ‘forest’ land use. The other land cover classes were also translated directly to land use but their class label remained the same. This direct translation into land use classes correctly accounted for a large proportion of the objects within the sample sites. Accurate labelling of land use, however, must be examined in a functional and ecological context that includes determining not only what human activity is taking place and what vegetation is there at the time of the satellite image acquisition but also how that land area will respond in the future (e.g. through regeneration, afforestation, or permanent deforestation) (Coulston et al, 2013). Thus many exceptions to the direct and general rules concerning land cover to land use transitions exist. To account for these exceptions, land-use labels were re-coded manually wherever these exceptions were present.

Tree land cover corresponds to forest land-use, except when the predominant land use is not forestry. This includes, for example, all urban areas, orchards, oil palm plantations, agricultural land with trees and areas under agroforestry. Such lands are classed as other land-use with tree cover.

Other wooded land cover corresponds to other wooded land use, except when the
object represents tree cover regeneration in which the woody vegetation will ultimately achieve height and density requirements to meet the definition of forest.

Other land-cover corresponds to other land-use except when it is a forest stand which is cleared and prone to regenerate or to be replanted, in which the object is re-labeled as forest land-use.

4.3.6 Effect of land cover or land use definition on forest area calculations

We examined the mean proportion of tree cover and forest land-use by continent/country and climatic domain for year 2000 in order to assess the effect of using a land cover or land use definition on the calculation of ‘forest’ area (Table 4-5). For many continental/climate groups, the difference between land cover and land use are minimal. However, for other continental/climate groups there are differences between land cover and land use that, depending on which one is used to calculate ‘forest’ area, could have an impact on global figures.

Table 4-5. The mean proportion of forest when using a land-use or tree cover definition by Continent/Country and Climatic Domain grouping.

<table>
<thead>
<tr>
<th>Continent / Country</th>
<th>Climatic Domain</th>
<th>Forest Land-Use (proportion)</th>
<th>Tree Cover (proportion)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canada</td>
<td>boreal</td>
<td>0.72</td>
<td>0.69</td>
</tr>
<tr>
<td>Europe</td>
<td>boreal</td>
<td>0.69</td>
<td>0.69</td>
</tr>
<tr>
<td>Africa</td>
<td>subtropical</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>Europe (ex. Russian Fed.)</td>
<td>subtropical</td>
<td>0.20</td>
<td>0.20</td>
</tr>
</tbody>
</table>
North and Central America (excluding Canada), in both the subtropical and temperate climatic domain, shows a 5% higher proportion of tree cover than forest land-use. This is probably due to the extensive area of tree cover located in urban settings within the United States (Nowak et al., 2013). Agroforestry in Central America may also contribute to this difference (Garrity, 2012).

Tropical Asia also exhibits a higher proportion of tree cover than forest land-use (2%). This is likely an illustration of the effect of palm oil plantations that are classified as tree cover from a biophysical standpoint but do not meet the requirement for forest
Conversely, boreal and temperate zones of Canada indicate 6% more forest land-use than detected tree cover. This is most probably due to the large amount of forest burned and logged annually within the boreal and temperate zones of Canada (Natural Resources Canada, 2014). Though these areas may be classified from a biophysical standpoint as devoid of tree cover and non-forest at the time of the satellite image acquisition, tree cover will regenerate and the areas remain considered as forest land-use.

Temperate Asia shows 2% more forest land-use than tree cover. This is likely due to the large areas designated as tree plantations in China (Liu and Tian, 2010) that may not be fully planted or fully matured to the point of being detectable by a moderate spatial resolution sensor such as Landsat.

Interestingly, tropical Africa indicates 4% more forest land-use than tree cover. We hypothesize that this may be due to the visual inspection process by local, national experts who could more accurately identify a site as forest in areas where tree cover was sparse and not adequately detected by the medium spatial resolution Landsat sensor (Lindquist et al., 2012).

4.4 Conclusions

The OBIA classification and change detection methodology described in this paper provided an efficient means of processing over 11,000 sample sites, 33,000 Landsat 20x20 km sample tiles and more than 6.5 million individual polygons over three epochs. Comparisons of automated classification results with expert-corrected results yielded agreements of approximately 80%. The automated land cover classification
methods thus provided suitable results and decreased the amount of time necessary for expert review and revision.

Implementing a relatively large MMU (5 ha) facilitated expert human interpretation, review and revision of the results and overcame issues related to fine-scale spectral heterogeneity. However, if the local change dynamic is very fine-scale, a large MMU applied to medium spatial resolution imagery may be incapable of adequately capturing this. In these areas, a strict adherence to a MMU may not be helpful and employing a much smaller or no MMU is advisable. And while pixel-based classification methods may indeed suffer from the ‘salt-and-pepper’ effect of individual pixels being miss-classified [18], image segmentation with a MMU may suffer a similarly confounding fate by systematically under-segmenting areas of land cover change, some or all of which may be ecologically significant.

Finally, land use classifications provide a critical but largely absent component of current land surface change studies. They provide the only information that can elucidate drivers of forest cover change and, subsequently, suggest pathways for land managers seeking to promote or prevent such change. The conversion of land cover to land use, and subsequent expert human review and revision in this study enabled differentiation of relatively temporary changes in the biophysical properties of the land surface from longer-term permanent land use conversions. This has important implications for determining net forest area change, not only gross area losses and may help to explain some of the differences in current estimates of global ‘forest’ area and change.

Acknowledgments: The authors would like to thank the research team at the JRC
including Frederic Achard, Hans-Jurgen Stibig, Catherine Bodart, René Beuchle, Andreas Brink, Silvia Carboni, François Donnay, Hugh D. Eva, Desirée Johansson, Ouns Kissiyar, Andrea Lupi, Philippe Mayaux, Rastislav Raší, Dario Simonetti, Hans-Jürgen Stibig, Michael Vollmar, Roman Seliger, Jesús San-Miguel-Ayanz, Pieter Kempeneers, Fernando Sedano, Lucia Seebach, Peter Strobl, Peter Vogt. The authors also wish to thank the national forestry experts who assisted with the review and revision of the RSS sample sites (http://www.fao.org/forestry/fra/67944/en/). Funding for the FAO Remote Sensing Survey was provided by the European Commission.
CHAPTER 5

GLOBAL FOREST LAND-USE CHANGE FROM 1990 - 2005

This chapter was published as:

5.0 Abstract

A survey of global land use and land-use change was carried out to estimate the amount of the Earth’s surface in a forest land-use and the amount of gain and loss from 1990 to 2005. Based on a systematic sample of Landsat imagery from 1990, 2000 and 2005, the survey estimated the total area of the world’s forests in 2005 at 3.8 billion hectares, or 30% of the global land area. Overall, there was a net decrease in global forest area of 1.7% between 1990 and 2005, at an annual rate of change of 0.11%. This equates to an annual shift from forest land-use to other land uses of 3 million hectares per year between 1990 and 2000 and of 6 million hectares per year between 2000 and 2005.

Major regional differences were found in the net rates of forest area change – only Asia and North America experienced gains in forest area, all other regions saw net declines. South America had the highest net forest loss - some 3.3 million hectares annually between 1990 and 2005. Africa had the second highest net forest loss of 1.6 million hectares annually during the same period. Europe, including the Russian Federation, had net losses of 0.5 million hectares annually and Oceania lost just under 0.1 million hectares annually. North America experienced net gains in forest area of some 0.2 million hectares annually while Asia had a net gain of 1.4 million hectares annually between 1990 and 2005.

Forests were categorized according to four climatic domains: boreal, subtropical, temperate and tropical. There were significant gains in forest area in the boreal (0.9
million hectares annually) and subtropical (1.1 million hectares annually) between 1990 and 2005. There were also net gains in forest area in the temperate domain of 0.9 million hectares for this period. In contrast, the tropical domain had a net loss of forest area of 6.8 million hectares annually between 1990 and 2005. This net reduction in forest land-use was nearly 2.5 times the net forest area gained in the other three domains.

5.1 Introduction

FAO analyses and compiles data on the extent and state of the world’s forests through a process called the Global Forest Resources Assessment (FRA). Published every 5–10 years, the FRA report reflects the major issues of concern prevalent at the time of reporting. In response to post-Second World War needs, early FRAs focused on timber stocks, while more recent editions, including FRA 2010 (FAO, 2010), have addressed topics such as forest biodiversity, forest carbon stocks and the social benefits of forests.

The FRA is an important information source for global efforts to sustainably manage forests, reduce the concentration of atmospheric greenhouse gases and advance other international initiatives. According to guidelines for national greenhouse gas inventories published by the Intergovernmental Panel on Climate Change (IPCC) (Paustian, Ravindranath and van Amstel, 2006), FAO is the main source of activity data and emission factors for forest and other land-use categories in Tier 1 calculations. The IPCC guidelines suggest that, where more detailed country data are unavailable, aggregate information can be obtained from international data sources such as the FRA.
5.1.1 The FRA 2010 Remote Sensing Survey

The FRA 2010 Remote Sensing Survey was the result of a partnership between FAO, countries and the European Commission Joint Research Centre (JRC). Its goal was to obtain globally consistent information on the areal extent and changes in tree cover and forest land-use between 1990 and 2005 at the regional, climatic domain and global levels. This report presents the results of the global forest land-use component of the survey.

5.2 Methods and materials

5.2.1 Land cover and land use

This report includes global statistics on forest land-use derived from a land-cover classification and expert image interpretation. Land cover refers to the biophysical attributes of the Earth’s surface and can be detected directly from aerial imagery or satellite-borne sensors. Land use implies a human dimension or purpose for which the land is used (Lambin et al., 2001). Land use can be inferred from remotely sensed data but typically must be verified by local expert knowledge or data collected in the field. Accurate information on land use is critical for understanding the causes of forest-cover change and for developing effective policies and strategies to slow and reverse forest loss.
5.2.2 Systematic sample design

The survey used a systematic sample of 10 km x 10 km satellite image extracts at each 1-degree intersection of latitude and longitude (Mayaux et al., 2005; Ridder, 2007). Globally, this is equivalent to a 1% sample of the Earth’s land surface. Sampling intensity was reduced above 60 degrees latitude, north and south, to include only even degrees of longitude. This was done to avoid an increasing “weight” of samples in the high latitudes due to the curvature of the Earth. No sites were located higher than 75 degrees latitude, north or south. For Canada, the 1-degree grid was modified to use the Canadian National Forest Inventory’s 20-km grid of smaller 4-km² photo points (Gillis, Omule and Brierley, 2005). The final sample grid consisted of 15 779 samples worldwide (Figure 5-1).

Figure 5-1. The 15 779 1-degree grid sample site locations used in the survey, with reduced intensity above 60° latitude north and south. Canada samples were spaced on a 20-km grid to match the Canadian National Forest Inventory (inset, see Annex 1 and Annex 2). Sites processed by JRC are in grey and sites processed by FAO are in black.
In a number of national, regional and global studies (e.g. Hansen et al., 2008; Stehman, Sohl and Loveland, 2005; Potapov et al., 2008; Eva et al., 2010), sampling approaches have proved successful in producing results for forest area change with acceptable and known precision. In previous remote sensing surveys, an approach using a large sample of satellite imagery over broad geographic regions has been shown to suitably capture parameter estimates at the regional (i.e. > 100 000 hectares (ha)) and continental scales (Czaplewski, 2002).

A systematic sample was chosen for four main reasons (Ridder, 2007): land cover exhibits trends at the regional and continental scales and no a priori assumptions of forest area change intensity were considered; the layout of the latitude–longitude grid is not politically biased and is easy to understand; sample locations can easily be identified on maps; and FAO-supported national forest assessments are typically constructed based on the same grid.

5.2.3 Imagery data sources

Imagery from the United States Geological Survey’s Landsat Global Land Survey (GLS) provided the majority of data for classification and interpretation (Gutman et al., 2008). The Landsat sensor provides global coverage, a long time-series of acquisitions, and spatial and spectral characteristics suitable for the detection of changes in tree cover. Landsat acquisitions are referenced to the Earth’s surface by a grid of paths and rows, called the Worldwide Reference System (WRS). The GLS is a spatially consistent, multi-

For each sample site, Landsat optical bands 1–5 and 7 from the GLS1990, GLS2000 and GLS2005 datasets were compiled. These were clipped to a 20 km × 20 km box centred on each 1-degree latitude and longitude intersection to create imagery subsets. The central 10 km × 10 km of each image subset was used for area calculations and statistical analysis. In areas where the GLS acquisitions were cloudy or not seasonally matched, effort was made to obtain additional scenes from the Landsat data archive or directly from regional ground stations (for more detail see Beuchle et al., 2011; Potapov et al., 2010; Seebach et al., 2010).

For boreal, temperate and subtropical climatic domains, the GLS data were assumed to be the best available. If more than one GLS acquisition was available for a given site and date, the GLS acquisition with the lowest cloud cover was selected for classification (Lindquist et al., submitted).

5.2.4 Image preprocessing

Images were preprocessed to correct for radiometric differences caused by changes in atmospheric quality or sensor characteristics between scene acquisition dates for the same site. Image normalization has the effect of standardizing digital number values relative to dense tree cover on a per-site basis and enables the more efficient application of automated classification algorithms (Toivonen et al., 2006; Potapov et al.,
Potapov et al. (2010) describe the preprocessing methods used by the FAO team for areas outside the tropics. Bodart et al. (2011) describe the preprocessing methods used by the JRC team for the tropical and sub-Saharan Africa sites.

### 5.2.5 Automated land-cover classification

FAO and JRC both carried out automated land-cover classifications of preprocessed imagery. The JRC team processed sites within the tropics, sub-Saharan Africa (Beuchle et al., 2011) and Western Europe (Seebach et al., 2010) as part of its ongoing TREES-3, MONDE and FOREST projects (JRC 2010; see Raši et al., 2011 for details of the JRC land-cover classification processing chain). The FAO team processed all other sites (Figure 5-1). Although there were differences in the processing methods used by the two teams, the overall processing and importantly the output classifications are comparable. The processing methods consisted of the following common components:

- data acquisition;
- data preprocessing and image normalization;
- image segmentation;
- image classification.

The automated segmentation of land-cover polygons and preclassification of land-cover
types had two main goals: to create a spatially and temporally consistent dataset; and to avoid manual delineation, thus reducing the effort involved in the visual review and revision of land-cover and land-use labels.

The FAO–JRC land-cover classification methodology consisted of four main steps:

- image segmentation at level 1 (no minimum mapping unit or MMU) and level 2 (MMU approximately 5 ha in size);

- training data collection of representative sites for supervised classification;

- model construction and land-cover classification of level-1 objects;

- assignment of land-cover classification of level-2 objects.

All functions of segmentation and supervised classification were carried out using eCognition® image segmentation and processing software.¹

Image segmentation is the process of partitioning an image by grouping similar pixels into patches called objects (regularly referred to as segments or polygons) based on spectral similarity and spatial distinctiveness. The criteria for creating image objects from individual pixels in eCognition can be controlled by the operator by specifying values for a series of parameters such as size, shape and the degree of similarity to be achieved in the segmentation. These values affect clustering and control the overall shape and size of objects.

the objects created (Baatz and Schappe, 2000).

A multi-date segmentation routine used Landsat image bands from all three survey periods to create a single layer containing objects based on the spectral information in each period (Figure 5-2). Image segmentation was implemented in two parts. The FAO method was similar to the segmentation routines described by Raši et al. (2011), using parameters that allowed the creation of small, irregular-shaped objects based on the spectral reflectance values of Landsat bands 3, 4 and 5 (0.63–1.75 μm). These bands were chosen for their ability to discriminate differences in surface reflectance caused by changes in vegetation type (Desclée, Bogaert and Defourny, 2006; Duveiller et al., 2008). The first (i.e. level-1) segmentation created very small objects that ranged in size from a single Landsat pixel to greater than 100 ha and varied inversely with the spectral heterogeneity of the underlying Landsat image.

Figure 5-2. Example of three imagery dates combined to make a single composite image with segments that capture reflectance changes in each period.
The most recent image (i.e. 2005) was segmented first. The objects created during this process were used to constrain the segmentation of the image for 2000 and, in turn, those objects constrained the segmentation of the 1990 image. For the tropics, the segmentation was first applied to the pair of 1990 and 2000 images, then the dissolved objects for 2000 were used to constrain the segmentation of the image for 2005.

The target MMU of the level-2 segments was 5 ha (Ridder, 2007). The desired MMU was achieved by aggregating level-1 segments smaller than 5 ha with adjacent objects with the most similar average Landsat band 5 reflectance. Short-wave infrared reflectance was used due to its effectiveness in forest mapping applications (Horler and Ahern, 1986; Hoffhine and Sader, 2002). Land-cover classification was carried out on the spectrally homogenous level-1 segments. The level-2 segments were assigned class labels according to the underlying percent composition defined by the level-1 segments (Table 5-1).

Table 5-1. Level-2, 5-ha MMU land-cover labelling scheme based on the percent composition of underlying level-1 segments, listed in descending order of priority.

<table>
<thead>
<tr>
<th>Level 1</th>
<th>% composition</th>
<th>Level-2 land-cover</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tree cover</td>
<td>( \geq 30 )</td>
<td>Tree cover</td>
</tr>
<tr>
<td>Other wooded land</td>
<td>( \geq 70 )</td>
<td>Other wooded land</td>
</tr>
<tr>
<td>Other land cover</td>
<td>( \geq 70 )</td>
<td>Other land cover</td>
</tr>
<tr>
<td>Water</td>
<td>( \geq 70 )</td>
<td>Water</td>
</tr>
</tbody>
</table>

Given the large number of samples and the complexity involved in classifying each site, a supervised automated classification approach was selected as the best processing option. The overall classification methodology (depicted as a generalized
flowchart in Figure 5-3) was as follows: For each site and date, a land-cover classification was produced with the following main classes – *tree cover, shrub cover, other land* (comprising herbaceous cover and bare ground/non-vegetated, which were grouped and not shown separately), *water* and *no data*. These classes were broadly in line with the IPCC land-use good-practice guidelines (Paustian, Ravindranath and van Amstel, 2006) when ultimately converted to land-use labels.

Imagery from 2000 was classified first. When there was a low likelihood of detecting change between surveys, the class label for objects in the image object layer for 2000 was transferred to the 1990 and 2005 image object layers. The objects determined to have a relatively high likelihood of change between 1990 and 2000 and between 2000 and 2005 were classified separately using training data automatically selected from non-changing objects in the same period. The 5-ha MMU objects were assigned class labels according to the proportion of labelled level-1 objects they contained.

5.2.6 Training the classification

The broad range of biophysical traits exhibited globally by tree cover presented a challenge for training data collection. For example, dense, dark, evergreen conifers have different characteristics to broad-leaved evergreens, which differ, in turn, from the characteristics of broad-leaved deciduous trees. The variations in biophysical features, changing seasonality and illumination conditions due to sun angle and slope position
combine to affect the spectral reflectance properties of tree cover and make it difficult to create reflectance-based models that can accurately classify tree cover in its myriad forms globally. The FAO classification methodology attempted to account for this variation by applying a single method for creating tree-cover classification models globally to each sample site and period. At each sample site, therefore, three separate models of land-cover classification were created and applied, one for each period.

For sites in the boreal, temperate and subtropical domains, training labels for each land-cover class were assigned to level-1 image objects using temporally coincident year 2000 Moderate Resolution Imaging Spectroradiometer (MODIS) Vegetation Continuous Fields (VCF) (Hansen et al., 2003) and 2005 GlobCover (Arino et al., 2008) land-cover products. Training class labels for water bodies were assigned based on the proportion of MODIS global water mask pixels (Carroll et al., 2009) falling within an individual image object. Data from GlobCover were used to assist with the classification of shrub-dominated land cover.

Artificial neural network classifiers were used to produce land-cover classifications for the FAO-processed sample sites. For each site, the network was trained and then applied to all year 2000 image objects. Objects with the same or similar spectral characteristics in 1990 and 2005 as in 2000 were automatically assigned the land-cover label from the 2000 image object. Where a large spectral change was detected between 1990 and 2000 or between 2000 and 2005, the 1990 and 2005 image objects were assigned labels based on individually created 1990 and 2005 classification models.
Figure 5-3. Generalized flowchart of the processing chain.
For the tropics, the object-based land-cover classification at level 1 was based on a supervised spectral library (Raši et al., 2011). Spectral signatures were collected from a common set of training areas representing the main land-cover classes within the tropics. For this purpose, the preprocessed Landsat ETM+ data for the year 2000 of all sample sites in a subregion were used. For each main land-cover class, several subclasses were identified, representing spectral variations due to site condition or land-cover subtype. For tree cover, for example, identified subclasses were dense evergreen forests, degraded evergreen forests, dry deciduous forests, mangroves and swamp forest. For each subclass, several training areas were selected. The number of pixels ultimately used for establishing the spectral signature of a subclass was generally higher than 1 000. Spectral signature statistics (means and standard deviations) were calculated at the level of subclasses. For South and Southeast Asia, for example, 73 spectral signatures were established as inputs to the digital classification of the four main land-cover categories. A generic supervised classification of the level-1 segmentation objects was performed uniformly for all sample sites, based on membership functions established from the spectral signature of each subclass for the Landsat spectral bands 3, 4 and 5. The membership functions were defined as an approximation of the class probability distribution. These membership functions were then applied to the imagery of the three years, i.e. extending the spectral signatures to 1990 and 2005. The subclasses resulting from supervised classification were not mapped as separate thematic land-cover categories but contributed to the mapping of the four main land-cover classes.

The supervised classification result obtained for the level-1 objects served as direct input to the thematic aggregation done at the level-2 segmentation (with a 5-ha
A sequential list of classification criteria (Table 5-1) was developed to label level-2 objects. For the purpose of forest monitoring, the main emphasis was on tree cover and tree-cover proportions within level-2 objects. For tropical sites, a tree cover mosaic class was introduced for objects containing partial tree cover at level 2: for example, a mapping unit containing 40 % tree cover (= total area of aggregated tree-cover objects at level 1) was still labelled tree cover mosaic. Level-2 objects were the only image object labels considered for the expert review-and-revision process described in later sections.

5.2.7 Land-use classes

Land-use classifications were based on FAO forest definitions (FAO, 2010), as follows:

- **Forest** – land spanning more than 0.5 ha with trees higher than 5 metres and canopy cover of more than 10 %, or trees able to reach these thresholds in situ. It does not include land that is predominantly under agricultural or urban land use.

- **Other wooded land** – land not classified as forest, spanning more than 0.5 ha; with trees higher than 5 metres and canopy cover of 5–10 %, or trees able to reach these thresholds in situ, or with a combined cover of shrubs, bushes and trees above 10 %. It does not include land that is predominantly under agricultural or urban land use.

- **Other land** – all land that is not classified as forest or other wooded land.

5.2.8 Conversion of land cover to land use
The conversion of land-cover class to land-use class was a two-step process. The first involved the automated conversion of land-cover classes to preliminary land-use labels (Figure 5-4). This conversion was presumed to account for the majority of polygons in the dataset. However, the accurate quantification of true land-use changes is complicated. The true land use of a given area must be examined in an ecological context that includes determining not only the vegetation present at the time of satellite image acquisition but also how the land will respond in the future (e.g. through regeneration, afforestation or deforestation) (Kurz, 2010).

![Figure 5-4. Land-cover and land-use classes and their associated numeric codes. In the conversion from land cover to land use, tree cover is converted to forest, shrub cover is converted to other wooded land, other land cover is converted to other land and water stays as water. Ideally, where there was a change in land use either to or from forest, the subclasses of other land use were to be used to identify the cause of the change.](image)

Operationally, FAO definitions required expert human interpretation to provide the context necessary for the accurate categorization of land use, especially where exceptions to the automated rules existed. The exceptions were as follows (see also
Figure 5-4): The *tree cover* and *tree-cover mosaic* land-cover classes were converted to the *forest land-use* class. Experts looked for exceptions where the land uses were either urban (e.g. trees in parks or gardens around houses) or agricultural (e.g. orchards). Urban areas with trees, orchards, oil-palm plantations, agricultural land with trees, and areas under agroforestry were identified and manually re-coded as *other land use with tree cover*. *Shrub cover* was converted to the *other wooded land* land-use class. Experts looked for exceptions, such as forest re-growth where trees were likely to grow taller than 5 metres, and re-coded those areas as *forest*. *Other land cover* was converted to *other land use*. Experts looked for exceptions such as temporarily un-stocked areas that may have had no trees at the time of the image but were likely to regenerate or be replanted, in which case they were re-coded as *forest*.

### 5.2.9 Expert interpretation, validation and correction of land cover and land use

The final assignment of land-cover and land-use labels was carried out by selected national forestry or remote sensing experts. The visual checks were conducted on all the imagery of three survey periods to review and revise the automatically assigned land-cover and land-use labels. The JRC developed a dedicated stand-alone computer application for this purpose (Simonetti, Beuchle and Eva, 2011). The aim of this tool was to provide a user-friendly interface, with an easy-to-use set of functions for navigating and assessing a given dataset of satellite imagery and land-cover/land-use maps, and to efficiently re-code areas where, according to expert judgement, changes were required (Figure 5-5).
Visual control and refinement of the digital classification results at object level 2 were implemented in three steps: (i) Obvious errors from the automatic classification were corrected, (ii) a revision of the mapping results was carried out by national experts, who contributed local forest knowledge to improve the interpretation. Nineteen regional workshops were held between September 2009 and July 2011, involving 204 national experts from 107 countries (Annex 3) and (iii) experienced image interpreters performed a final screening for errors overlooked or mistakenly re-introduced and controlled for interpretation consistency across the region, applying final corrections where necessary.
The review and revision of the classification was aided by very-high-resolution satellite imagery, Google Earth™, images from the Degree Confluence Project², Panoramio™, and existing vegetation maps, where available. Specific expert field knowledge was also important. The phase of visual control and refinement was designed as a crucial component for correcting classification errors and for implementing the change assessment.

5.3 Data analysis

All calculations are shown in Annex 4.

5.3.1 No data

Areas obscured by cloud or otherwise lacking data due to poor satellite coverage or low-quality images were coded as “no data” in both the land-cover and land-use polygons. Cloud-affected and shadow-affected imagery was most common in the tropics (Ju and Roy, 2008; Asner, 2001); about 9% of the 4016 tropical sample sites had no data for 2005. Where possible, areas obscured by cloud or shadow were re-coded manually based on an examination of the same location using images recorded at later or earlier dates, or by using national datasets, Google Earth® or local knowledge.

“No data” areas were considered an unbiased loss of information. If not resolved

² www.confluence.org.
using the methods above, a “no data” classification encountered in one time period was passed to the land-cover and land-use label in all other time periods during analysis to ensure that only areas with viable data concurrent to all survey periods were analysed. Survey sites missing a Landsat acquisition for any of the time periods were removed from the analysis. Ultimately, 13,066 sites were processed to generate the results after all ‘no data’ sites had been accounted for (Figure 5-6 and Annex 2).

![World Map](image)

*Figure 5-6. Final 13,066 sample sites used in the 2010 analysis. Note the absence of sites in the Eastern Russian Federation due to a lack of Landsat acquisitions from the 1990s time period.*

The proportion of forest and gross gains and losses were calculated relative to the total area of all viable image objects, or “good land”. Good land was considered to be any object not classified as water or “no data” (Annex 4, equation 1).

### 5.3.2 Adjustment for latitude and area weighting

Due to the curvature of the Earth, the actual area represented by a latitude/longitude grid sample decreases with latitude. Analyses of forest area and forest-
area change must take this into account by applying a correction to area measurements (Annex 4, equation 2).

Sites were also given a weight equivalent to the proportion of the total surveyed area represented by the site. Both latitude and area weights were incorporated in the survey analysis (Annex 4, equation 3).

5.3.3 Aggregation for regional and climatic domain analysis

Land-use classifications were summarized on a per plot basis and aggregated by FRA region and FAO climatic domain (Figure 5-7) (Iremonger and Gerrand, 2011). Each survey site was assigned to the FRA region and FAO climatic domain within which the majority of the site was located. Survey data were analysed using the statistical software packages R (2.12.2) and Systat (Ver. 13).
Figure 5-7. Regions (top) and climatic domains (bottom) used for aggregation and analysis.
5.3.4 Forest area: gains and losses

Total forest area was determined using the Horvitz-Thompson direct estimator following Eva et al. (2010) – that is, by calculating the mean proportion of forest (Annex 4, equation 4) over all sample sites within a region or climatic domain and multiplying this figure by the total land area of the region. Forest area for each site was calculated at the nominal date of image acquisition: i.e. without taking the real acquisition date into account. Global forest area totals were calculated by summing the total forest area per region. This was done because confidence intervals for regional totals were smaller than for climatic domains (Table 5-2). A similar approach was used to calculate gross and net forest area gains and losses. All calculations were made using the Mollweide equal area map projection.

5.3.5 Annualizing forest-area change

The satellite imagery used in the survey, while nominally representing 1990, 2000 and 2005, was acquired over a range of dates around the target year (Figure 5-8). Changes were calculated as mean annual changes, based on the date range represented by the imagery acquisition date at each site (Annex 4, equation 5).

5.3.6 Error

The statistical precision of all estimates are reported as the values from the 95 %
confidence interval expressed as percent of the mean (Annex 4, equations 6–8).

Reported errors are sampling errors only and do not account for classification errors or other sources of error.

Table 5-2 Mean forest area (‘000 ha ± standard error) by region (a) and climatic domain (b), 1990, 2000 and 2005. The sum of the forest areas of all regions was used as the global forest area total.

<table>
<thead>
<tr>
<th>Region</th>
<th>1990</th>
<th>2000</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>n 1990</td>
<td>2000</td>
<td>2005</td>
</tr>
<tr>
<td>Africa</td>
<td>2322</td>
<td>520 000 ± 7%</td>
<td>510 000 ± 7%</td>
</tr>
<tr>
<td>Asia</td>
<td>2863</td>
<td>500 000 ± 7%</td>
<td>510 000 ± 7%</td>
</tr>
<tr>
<td>Europe</td>
<td>907</td>
<td>1 080 000 ± 5%</td>
<td>1 070 000 ± 5%</td>
</tr>
<tr>
<td>North and Central America</td>
<td>4833</td>
<td>790 000 ± 3%</td>
<td>800 000 ± 3%</td>
</tr>
<tr>
<td>Oceania</td>
<td>769</td>
<td>120 000 ± 14%</td>
<td>120 000 ± 14%</td>
</tr>
<tr>
<td>South America</td>
<td>1372</td>
<td>860 000 ± 5%</td>
<td>820 000 ± 5%</td>
</tr>
<tr>
<td>World</td>
<td>13066</td>
<td>3 860 000 ± 2%</td>
<td>3 820 000 ± 2%</td>
</tr>
<tr>
<td>b.</td>
<td>n 1990</td>
<td>2000</td>
<td>2005</td>
</tr>
<tr>
<td>Boreal</td>
<td>3092</td>
<td>1 180 000 ± 3%</td>
<td>1 190 000 ± 3%</td>
</tr>
<tr>
<td>Subtropical</td>
<td>1958</td>
<td>320 000 ± 8%</td>
<td>330 000 ± 8%</td>
</tr>
<tr>
<td>Temperate</td>
<td>3831</td>
<td>560 000 ± 5%</td>
<td>570 000 ± 5%</td>
</tr>
<tr>
<td>Tropical</td>
<td>4185</td>
<td>1 730 000 ± 4%</td>
<td>1 670 000 ± 4%</td>
</tr>
</tbody>
</table>
Figure 5-8. The range of dates of satellite imagery used in the study by survey period. The table below the graph lists the earliest, latest, average and median dates for each survey period.

<table>
<thead>
<tr>
<th></th>
<th>Ref_1990</th>
<th>Ref_2000</th>
<th>Ref_2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>min</td>
<td>17/04/1984</td>
<td>22/06/1996</td>
<td>14/08/2003</td>
</tr>
<tr>
<td>max</td>
<td>22/08/1996</td>
<td>22/05/2003</td>
<td>11/05/2009</td>
</tr>
<tr>
<td>median</td>
<td>07/08/1989</td>
<td>12/09/2000</td>
<td>01/10/2005</td>
</tr>
</tbody>
</table>

5.4 Results and discussion

The statistical significance of weighted, annualized gains and losses in gross forest area and net change in forest area was tested for regions and climatic domains using several analyses including (i) Welch’s t-test to indicate whether the gains, losses and net change are different from 0 (Table 5-3), (ii) general linear models to calculate slopes and the significance of intercept and slope (Table 5-4), (iii) analysis of variance
(ANOVA) to detect interactions between climatic domain and year (Table 5-5), and (iv) restricted maximum likelihood (REML) analysis as a more robust tool for assessing differences and interactions assuming unequal variances of the sample populations (Table 5-6).

Table 5-3. Significance of net annual changes and gross annual gains and losses for regions and climatic domains, a * indicates a value significantly different from zero ($p < 0.05$) using Welch’s $t$-test.

<table>
<thead>
<tr>
<th>Region</th>
<th>Significant change, 1990-2000</th>
<th>Significant change, 2000 - 2005</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>net</td>
<td>gain</td>
</tr>
<tr>
<td>Boreal</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Subtropical</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Temperate</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Tropical</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Africa</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Asia</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Europe</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>North and Central America</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Oceania</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>South America</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>World</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

Table 5-4. $P$ values for the slope of the line formed by a general linear model relating annualized net change and gross gains and losses with survey period by regions and climatic domains. Significant differences ($p < 0.05$) between survey periods are in green. For net change, the direction of the arrow indicates whether there was a net forest area loss (↓) or gain (↑).
### Table 5.5. Results of ANOVA test for annual net forest area change, by climatic domain and year

<table>
<thead>
<tr>
<th>Domain</th>
<th>Net</th>
<th>Gain</th>
<th>Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boreal</td>
<td>0.167</td>
<td>0</td>
<td>0.001</td>
</tr>
<tr>
<td>Subtropical</td>
<td>0.895</td>
<td>0.178</td>
<td>0.009</td>
</tr>
<tr>
<td>Temperate</td>
<td>0.018</td>
<td>↑ 0.003</td>
<td>0.417</td>
</tr>
<tr>
<td>Tropical</td>
<td>0</td>
<td>↓ 0.664</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Region</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td>0</td>
<td>↓ 0.787</td>
<td>0</td>
</tr>
<tr>
<td>Asia</td>
<td>0.515</td>
<td>0.014</td>
<td>0.122</td>
</tr>
<tr>
<td>Europe</td>
<td>0.133</td>
<td>0.646</td>
<td>0.03</td>
</tr>
<tr>
<td>North and Central America</td>
<td>0.027</td>
<td>↑ 0</td>
<td>0.339</td>
</tr>
<tr>
<td>Oceania</td>
<td>0.595</td>
<td>0.438</td>
<td>0.78</td>
</tr>
<tr>
<td>South America</td>
<td>0.001</td>
<td>↓ 0.928</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>World</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>World</td>
<td>0.001</td>
<td>↓ 0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III SS</th>
<th>df</th>
<th>Mean squares</th>
<th>F-ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Climatic domain</td>
<td>1.096</td>
<td>3</td>
<td>0.365</td>
<td>237.686</td>
<td>0.000</td>
</tr>
<tr>
<td>Year</td>
<td>0.053</td>
<td>1</td>
<td>0.053</td>
<td>34.678</td>
<td>0.000</td>
</tr>
<tr>
<td>Climatic domain * Year</td>
<td>0.164</td>
<td>3</td>
<td>0.055</td>
<td>35.499</td>
<td>0.000</td>
</tr>
<tr>
<td>Error</td>
<td>40.162</td>
<td>26124</td>
<td>0.002</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 5-6. REML results for annual net change by climatic domain and survey period (1990-2000 and 2000-2005).

<table>
<thead>
<tr>
<th>Effect</th>
<th>Effect level</th>
<th>Estimate</th>
<th>Standard error</th>
<th>df</th>
<th>t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Climatic domain</td>
<td>Boreal</td>
<td>0.003</td>
<td>0.002</td>
<td>26123</td>
<td>1.083</td>
<td>0.279</td>
</tr>
<tr>
<td></td>
<td>Subtropical</td>
<td>0.002</td>
<td>0.002</td>
<td>26123</td>
<td>0.962</td>
<td>0.336</td>
</tr>
<tr>
<td></td>
<td>Temperate</td>
<td>0.002</td>
<td>0.002</td>
<td>26123</td>
<td>0.81</td>
<td>0.418</td>
</tr>
<tr>
<td></td>
<td>Tropical</td>
<td>-0.007</td>
<td>0.002</td>
<td>26123</td>
<td>-2.879</td>
<td>0.004</td>
</tr>
<tr>
<td>Year</td>
<td></td>
<td>0.000</td>
<td>0.000</td>
<td>26123</td>
<td>0.346</td>
<td>0.729</td>
</tr>
<tr>
<td>Climatic domain * Year</td>
<td>Year*Boreal</td>
<td>0.000</td>
<td>0.000</td>
<td>26123</td>
<td>7.217</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Year*Subtropical</td>
<td>0.000</td>
<td>0.000</td>
<td>26123</td>
<td>1.638</td>
<td>0.101</td>
</tr>
<tr>
<td></td>
<td>Year*Temperate</td>
<td>0.000</td>
<td>0.000</td>
<td>26123</td>
<td>1.667</td>
<td>0.095</td>
</tr>
<tr>
<td></td>
<td>Year*Tropical</td>
<td>0.000</td>
<td>0.000</td>
<td>26123</td>
<td>-3.069</td>
<td>0.002</td>
</tr>
</tbody>
</table>

5.4.1 The area in forest land-use declined between 1990 and 2005

Figure 5-9 shows the estimated forest area by region in 1990, 2000 and 2005, and Figure 5-10 shows the estimated forest area by climatic domain for the same years. Total forest area in 2005 was 3.8 billion ha, which is approximately 30% of the global land area. There was a net reduction in the global forest area between 1990 and 2005 of 66.4 million ha, or 1.7%.
Figure 5-9. Forest area by region for 1990, 2000 and 2005.
5.4.2 Global forest loss and gains

Worldwide, the gross reduction in forest land-use was 9.5 million ha per year between 1990 and 2000 and 13.5 million ha per year between 2000 and 2005. This reduction was partially offset by gains in forest area through afforestation and natural forest expansion of 6.8 million ha per year between 1990 and 2000 and 7.3 million ha per year between 2000 and 2005. Thus, the rate of annual net forest loss increased significantly ($p < 0.05$) from 2.7 million ha between 1990 and 2000 to 6.3 million ha between 2000 and 2005 (Table 5-7). Figures 5-11 and 5-12 show these changes by geographic region and climatic domain.
Table 5-7. Mean annual net forest area change and 95% confidence intervals between survey periods for FAO regions and FAO climatic domains. Global net change was calculated by summing estimates for FAO regions.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td>-1091</td>
<td>-2712</td>
<td>306</td>
<td>560</td>
</tr>
<tr>
<td>Asia</td>
<td>1419</td>
<td>1367</td>
<td>564</td>
<td>703</td>
</tr>
<tr>
<td>Europe</td>
<td>-437</td>
<td>-638</td>
<td>303</td>
<td>578</td>
</tr>
<tr>
<td>North and Central America</td>
<td>323</td>
<td>55</td>
<td>190</td>
<td>287</td>
</tr>
<tr>
<td>Oceania</td>
<td>-101</td>
<td>-61</td>
<td>87</td>
<td>136</td>
</tr>
<tr>
<td>South America</td>
<td>-2779</td>
<td>-4275</td>
<td>516</td>
<td>863</td>
</tr>
<tr>
<td>Total</td>
<td>-2666</td>
<td>-6264</td>
<td>902</td>
<td>1410</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Climatic Zone</th>
<th></th>
<th>CI</th>
<th></th>
<th>CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boreal</td>
<td>776</td>
<td>1153</td>
<td>565</td>
<td>1088</td>
</tr>
<tr>
<td>Subtropical</td>
<td>1212</td>
<td>902</td>
<td>295</td>
<td>380</td>
</tr>
<tr>
<td>Temperate</td>
<td>787</td>
<td>1152</td>
<td>288</td>
<td>364</td>
</tr>
<tr>
<td>Tropical</td>
<td>-5648</td>
<td>-9111</td>
<td>775</td>
<td>1238</td>
</tr>
<tr>
<td>Total</td>
<td>-2873</td>
<td>-5904</td>
<td>1044</td>
<td>1730</td>
</tr>
</tbody>
</table>

5.4.3 Regional differences in forest loss and gain

In South America, significant forest conversion to other land uses occurred in both survey periods: 2.8 million ha per year between 1990 and 2000 and 4.3 million ha per year between 2000 and 2005. In Africa, there were statistically significant net annual forest area losses of 1.1 million ha between 1990 and 2000 and 2.7 million ha between 2000 and 2005.

Europe, including the Russian Federation, had a statistically significant net annual loss of forest area of 0.4 million ha between 1990 and 2000 and 0.6 million ha between 2000 and 2005. Oceania had significant net annual forest losses of 0.1 million ha between
1990 and 2000 and no significant change in forest area between 2000 and 2005.

There was a significant mean annual net gain in forest area in North America between 1990 and 2000 of 0.3 million ha, but there was no significant net change between 2000 and 2005. In Asia, there were significant mean annual net gains in forest area of 1.4 million ha between 1990 and 2000 and 1.4 million ha between 2000 and 2005.

Figure 5-11. Gross gains and losses and net changes in forest area, by FRA region, 1990-2000 and 2000-2005.
Figure 5-12. Gross gains and losses and net changes in forest area, by climatic domain, 1990-2000 and 2000-2005.

Net forest loss was highest in the tropical climatic domain in both time periods: 5.6 million ha per year between 1990 and 2000 and 9.1 million ha per year between 2000 and 2005.

There were significant net annual gains in forest area in the temperate climatic domain of 0.8 million ha between 1990 and 2000 and 1.2 million ha between 2000 and 2005.

In the boreal climatic domain there were significant net annual gains in forest area of 0.8 million ha between 1990 and 2000 and 1.2 million ha between 2000 and 2005. The
high coefficient of variation in these estimates, however, indicates a large range in estimates of forest area change, which could be due to problems in the classification of land use and land cover in this zone.

The subtropical climatic domain showed significant net annual gains in forest area of 1.2 million ha between 1990 and 2000 and 0.9 million ha between 2000 and 2005.

5.4.4 Differences in the annual rate of change by region and climatic domain

There was a significant interaction between climatic domain and year (Table 5-5), meaning that the differences between survey periods were not the same across climatic domain types. These differences in the rate of net forest change between time periods were significant in the boreal and tropical climatic domains and insignificant in the subtropical and temperate domains (Table 5-6). The only climatic domain that showed a net decrease was the tropics, where the annual net change increased from a loss of 5.6 million ha in 1990–2000 to a loss of 9.1 million ha in 2000–2005.

The REML analysis in Table 5-6 allows for spatial and temporal correlation and unequal variance between populations and may be more robust than ANOVA for the analysis of survey data. REML analysis is used to decrease the chances of committing a Type 1 error when determining the statistical significance of some results (Picquelle and Mier, 2011).

In recent decades the tropics have been considered the largest source of net forest loss. This study confirms that trend and the fact the most of the loss occurred in Africa and South America (Table 5-7).
5.5 Comparison with other FAO studies

The following section compares estimates of forest area and forest area change made in this project with those derived from previous FAO pantropical remote sensing surveys and those presented in the FRA 2010 tabular reports (using country-supplied data).

5.5.1 Comparison with FRA 2000 pantropical remote sensing data

FAO (2001) conducted a remote sensing-based survey of forest area in the tropics for the years 1990 and 2000; hereafter, that survey is referred to as RSS 2000. RSS 2010 data were aggregated using the same geographic boundaries as those used in RSS 2000 (Figure 5-13), and the estimates of forest area, gross forest area loss and net forest area change for the years 1990 and 2000 were compared (Figure 5-14).
Figure 5-13. Distribution of RSS 2000 sample sites in the pantropics. The 117 sampling units of the survey were selected over the entire pantropical zone following a two-stage random sampling method based on geographical divisions (subregions) and forest cover or forest dominance.
Figure 5-14. Pantropical forest area for year 1990 (top) and 2000 (bottom) as estimated from RSS 2000 (blue bars) and this study (red bars).
Estimates of total forest area and gross forest area loss for the period 1990–2000 were not significantly different ($p < 0.05$) between the two surveys. The difference in estimates of net forest area change was not significantly different in Asia and South and Central America between the two surveys, but it was significantly different ($p < 0.05$) in Africa (Figure 5-15). RSS 2000 targeted areas of forest cover and did not include samples from non-forest, which could explain why estimates of net forest loss were generally higher in RSS 2000 than in RSS 2010.

![Figure 5-15. Comparison of pantropical net forest area change and gross forest area loss between 1990-2000 as estimated by RSS 2000 (blue bars) and this study (red bars).](image)

RSS 2000 consisted of 117 full Landsat scenes (representing a total sample area
of 250 million ha) and, in the area coincident to both surveys, RSS 2010 consisted of 3,631 sample sites (representing a total sample area of 36 million ha). The larger number of samples in RSS 2010 increased the precision of its estimates compared with those made in RSS 2000.

Figure 5-16 shows a complete timeline of tropical forest area estimates, by region, for 1980, 1990, 2000 and 2005 derived from FRA remote sensing surveys. The estimates for 1980 were derived from RSS 2000 and the estimates for 1990, 2000 and 2005 were derived from RSS 2010.

![Figure 5-16. Pantropical forest area for the years 1980, 1990, 2000 and 2005 as estimated from RSS 2000 (1980) and this study (all other years).](image)

5.5.2 Comparison with FRA 2010 tabular reports
The estimates of forest area and rates of change in RSS 2010 differ from those presented in the tables contained in FRA 2010 for both forest area and annual forest area change. Differences between the “state” (e.g. forest area) and “trend” (e.g. forest area change) of forest land-use are complex. In the following section, differences between RSS 2010 and FRA 2010 tabular reports (hereafter referred to as FRA 2010) are examined with respect to several key criteria, including the definition of forest, the reporting methods of both surveys, and the overall quality of the reported information.

5.5.3 Differences in forest area

The estimate of forest area in Africa in 2000 was almost 200 million ha (29 %) greater in FRA 2010 than in RSS 2010 (Figure 5-17). On a percentage basis, the greatest difference was in Oceania, where the estimated forest area in 2000 was 41 % (81 million ha) greater in FRA 2010. Similar differences in forest area were observed for 1990 and 2005 estimates.
Figure 5-17. A comparison of forest area, by region, as reported in FRA 2010 tabular reports (blue dots) and the remote sensing-based estimates presented in this paper represented by black bars indicating the 95 % confidence intervals about the mean.

Differences in forest area estimates between this study and FRA 2010 are likely due to differences in survey and reporting methods and to an issue in remote sensing arising from the definition of forest. The methods used to derive estimates in FRA 2010 vary by country and include the use of national forest inventories, remote sensing-based studies and expert opinion. FRA 2010 country questionnaires had a standard template to improve consistency between countries, but differences between countries in reporting standards still led to inconsistencies in the analysis of both the state and trend of forest area. For example, some countries did not submit completed FRA questionnaires for FRA
2010. For such countries, forest area state and trend were derived from ancillary data sources or previously reported figures (FAO, 2001). Depending on the frequency and standard of reporting, there is a risk that estimates are out of date and of unknown accuracy (Matthews, 2001).

Africa currently has the oldest data, on an area-weighted basis, of all the FRA regions (O. Jonsson, personal communication, 2012). The use of outdated information, which required extrapolation, sometimes over decades, to produce estimates for FRA 2010 contributes to the variation observed between forest area estimates in the two studies.

The definition of forest used in both FRA 2010 and RSS 2010 is characterized by a low threshold for tree canopy cover (i.e. > 10 %), which is difficult to detect using medium spatial resolution satellite imagery and to delineate accurately in the field at anything other than the plot level. Forest area with canopy cover less than 20 % may not be reliably detected from medium spatial resolution satellite imagery such as Landsat. Work is ongoing to determine canopy-cover percentage thresholds classified as forest in RSS 2010 through the incorporation of high spatial resolution imagery at selected locations. More consistent characterization of low-canopy-cover sites could reduce some of the difference between the two methodologies.

To test the theory that difficulty in delineating low-canopy-cover forest (usually in drier forest areas) contributes to differences in forest area estimates between FRA 2010 and RSS 2010, the proportion of dry ecological zone per region was related to the absolute difference in forest area estimates. Figure 5-18 shows a high degree of correlation between the area of dryland and differences in forest area estimates between
FRA 2010 and RSS 2010; uncertainty in estimating dryland forest area, therefore, may contribute to differences in forest area estimates.

Figure 5-18. Relationship between proportion of dry climatic domains per region and the proportional difference between FRA 2010 and RSS 2010 forest area estimates for that region.

5.5.4 Differences in net forest area change

The estimates of net change in forest area in RSS 2010 also differ from those reported in FRA 2010. Overall net change was much lower in this study (66.4 million ha) than in FRA 2010 (107.4 million ha). The magnitude of the annual rate of change was also different. RSS 2010 results indicate that the annual rate of net forest area loss
increased from about 3 million ha in the period 1990–2000 to 6 million ha in the period 2000–2005. FRA 2010, on the other hand, indicated a decrease in the rate of annual net forest loss from 8.3 million ha in 1990–2000 to 4.8 million ha in 2000–2005.

Differences in net change estimates between the two surveys are due largely to uncertainties in forest area and change in Africa, Asia and South America (Figure 5-19). In the period 1990–2005, RSS 2010 estimated a lower net decrease in forest area in Africa and South America and a higher net increase in forest area in Asia compared with FRA 2010. RSS 2010 indicated a net increase in forest area in Asia in both periods, while FRA 2010 estimated a net decrease in forest area between 1990 and 2000 and a net increase between 2000 and 2005.

Figure 5-19. A comparison of net change between the remote sensing-based estimates in this paper (bars with 95% confidence interval indicated) and FRA 2010 tabular reports (diamonds) for 1990-2000 (left) and 2000-2005 (right).
It should be noted that FRA 2010 did not report specifically on forest loss as a distinct and separate variable; rather, forest change estimates were derived from the difference between forest area estimates over time. Thus, errors in forest area reporting may be compounded, or they may confound estimates of forest area change.

5.6 Causes of land-use change

The type or cause of land-use change was not assessed in this study as originally planned. The attribution by national experts of land-use types to more detailed classes proved difficult in the time allotted during the review-and-revision workshops. Thus, while the conversion of forest land-use to other land uses and vice versa can be analysed readily, RSS 2010 results do not indicate whether forest losses are attributable to specific uses (e.g. pastureland or cropland). Likewise, gains in forest area could be due to natural expansion or the establishment of planted forests.

Existing scientific literature can be used to gain insight into the causes of forest land-use conversion. Survey results re-affirmed that tropical zones account for the largest portion of global net forest loss. Gibbs et al. (2010) re-analysed RSS 2000 data and estimated that the total net increase in agricultural area between 1980 and 2000 in the tropics was greater than 100 million ha, nearly 80% of which came from previously intact or disturbed forest land-use. Given the sustained and increasing demand for agricultural products for food and energy, it is likely that the causes of forest conversion to other land uses in the period 2000–2005 in the tropics are also predominantly due to the expansion of agriculture (Lambin and Meyfroidt, 2011).

RSS 2010 results indicate that forest area increased in the temperate climatic
domain, likely due to increases in planted forests in temperate Asia. Liu and Tian (2010) document a large increase (51.8 million ha) in forest area in China due to the establishment of planted forests, a process that began in the 1950s and continues today. FRA 2010 confirmed in part this finding for China, reporting an increase in forest area of about 2.5 million ha annually – of a total of 49.7 million ha – between 1990 and 2010.

RSS 2010 results also show an increase in forest area in the boreal climatic domain, although this increase is a surprise and more difficult to explain. The increase may be due to the forest regrowth that has occurred on large areas of abandoned farmland since the collapse of the former Soviet Union. Kuemmerle et al. (2010) estimate the natural expansion rate on abandoned farmland in Ukraine since 2000 at 8,600 ha per year. Similar rates of natural expansion of forest may be occurring on the nearly 26 million ha of abandoned farmland in the Russian Federation, Belarus and Kazakhstan (Lambin and Meyfroidt, 2011).

Another possible explanation for the detected increase in forest area in the boreal climatic domain could be misidentification of burned areas as non-forest land-use in earlier time periods. In Canada, a largely automated review and revision of land-use classifications was undertaken using the large Canadian National Fire Database (Stocks et al., 2003) to identify burned areas and reclassify other land cover to forest land-use where a fire was considered to be the cause of forest loss. The Canadian National Fire Database includes fires greater than 200 ha in size and represents about 97% of the total area burned annually in Canada (Stocks et al., 20023). The mislabelling of small fires as non-forest land-use or any discrepancies between the RSS 2010 land-cover detection and the Canadian National Fire Database may have contributed to an artificial increase in
forest land-use area as burnt areas regenerate.

5.7 Accuracy assessment

A formal accuracy assessment of the land-use classification was not performed as part of this study. It is difficult to find data sources of higher spatial resolution, appropriate temporal resolution or greater reliability, especially globally, against which to check the automatically classified and expert-revised land-use labels. A comparison of the automatically classified land-cover labels before and after expert review and revision indicated overall agreement of 77–81 % (Lindquist et al., submitted). Comparisons of expert-revised land-cover classifications with high spatial resolution satellite imagery for selected sites in the Russian Federation indicated that expert revision could yield accuracies of nearly 100 % for a forest/other land dichotomous classification scheme (Bartolev, 2012 unpublished data).

It is expected that land cover will reflect the underlying land use in most instances; therefore, the accuracies achieved by the methods used should provide an indication of the overall accuracy of estimates. However, the exceptions to the land-cover/land-use equivalence generalization are important and significant. In the future, further effort will be directed at devising a method for assessing more thoroughly the accuracy of the land-use classification.

5.8 Conclusion
This is the first survey of its kind to measure, in a systematic way, losses and gains in forest land-use between 1990 and 2005 at the global, regional, climatic domain and ecological zone levels of aggregation. The results presented in this report indicate that forest conversion to other land uses is most prevalent in the tropical climatic domain and, within this domain, in South America. Other climatic domains were remarkably stable in terms of net forest land-use change over the period 1990–2005.

The systematic survey design permitted estimates of gross forest area gains and losses and net changes in forest area, each with an estimate of precision. The exhaustive review-and-revision process by national-level forestry and remote sensing experts made possible the correction of classification errors and the identification of land uses not discernible from remotely sensed data sources alone, and provided an improved ecological context for the monitoring of forest cover and forest land-use change globally.

5.8.1 Integration of coarse resolution satellite imagery to help classification

The survey benefited from the use of global coarse spatial resolution datasets to both normalize and classify the relatively finer spatial resolution Landsat samples. Although coarse spatial resolution satellite imagery is often unsuitable as a stand-alone data source for detecting change, several studies have shown the effectiveness of using such data for the purpose of selecting training data for land-cover classifications at finer spatial resolutions. For example, Hansen et al. (2008) showed the utility of using coarse spatial resolution data from the MODIS VCF product to delineate potential training sites for a forest/non-forest classification in Central Africa. Similar methods have also been applied successfully in the Brazilian Legal Amazon (Broich et al., 2009), Indonesia
(Broich et al., 2011), and the boreal region of the Russian Federation (Potapov et al., 2008; Potapov, Turubanova and Hansen, 2011).

5.8.2 Visual review and revision of classification important

Visual control and correction was an important part of the land-cover and land-use classification processes and had a large impact on the final results. A comparison of the initial results from the automated land-cover classification and final reviewed-and-revised results for the tropics indicated that about 20 % of the polygon labels were revised by national experts (Raši et al., 2011). Similar results were obtained for sites in the boreal, temperate and subtropical domains (Lindquist et al., submitted). The visual refinement process also had a notable effect on estimates of forest area and forest area change: for Southeast Asia, for example, the net rate of change in tree cover (loss) from 1990–2000 was assessed at 0.9 % before and 1.6 % after visual control (Raši et al., 2011).

5.8.3 The utility of Landsat for global monitoring

Land-cover classification and change detection methods that leverage available data from the current generation of Landsat sensors is critical for maintaining a record of land-cover changes until the new generation of sensors comes online. The Landsat programme has the longest continuous time-series of similar remotely sensed Earth observations and is a critical component in the analysis of change in land cover and land use since the 1970s. Landsat 7, the latest sensor, was launched in 1999 but suffered a mechanical failure in May 2003 that created no-data gaps in the across-track scan line covering 23 % of each image (Williams, Goward and Arvidson, 2006). Sampling
methods, such as those described in this report, are a suitable use of the currently available Landsat image acquisitions and should be used to leverage the large amounts of information freely available in the Landsat archive (Woodcock et al., 2008).

5.8.4 Establishment of global networks

The project established two very important global networks. One was the global survey grid, which will be updated with data from 2010 as part of the next FRA (to be released in 2015). The second and perhaps more important network comprises the many national experts who participated in the survey and who remain important points of contact and sources of forest remote sensing and land-use expertise in individual countries.

Annex 5-1 Country-specific review-and-revision methodologies

Every effort has been made to produce consistent results at a global scale. Some countries, however, used unique methods to review and revise land-cover and land-use classifications. Those methods are described here.

Canada
Data for Canada were derived using the classification methodology described in the main body of this report but applied across the Canadian National Forest Inventory (NFI) photo-plot grid system (Gillis, Omule and Brierley, 2005). The NFI uses 2 km × 2 km plots with 20 km horizontal and vertical spacing (i.e. a 20 km systematic grid), producing more than 18 000 individual plots. For the purposes of RSS 2010, a 25 % sample of the plots (i.e. every fourth plot) was selected for initial analysis (Figure 5-1). In total, 4 052 2 km × 2 km plots were analysed across Canada.

At each plot location, level-1 segments from imagery captured in 2000 were directly assigned land-cover labels based on the Canadian Earth Observation for Sustainable Development of Forests (EOSD) dataset (Wulder et al., 2006). The EOSD dataset is a 25 m spatial resolution, Landsat-based, 23-class land-cover classification for the forested areas of Canada. The 23 EOSD classes were aggregated into the simple 5-class legend, and level-1 segments for 2000 were assigned a value based on the majority land cover of the underlying EOSD data. The full methodology, as described in the main body of this report, was used where no EOSD data existed (i.e. in largely non-forested portions of Canada) and to classify 1990 and 2005 segments.

The initial conversion of land cover to land use was completed following the survey conversion rules, as described in the main body of this report. Next, a series of automated re-coding procedures was implemented in the review-and-revision phase of land-use validation. These procedures involved re-coding polygons to forest land-use in cases where commercial timber harvest activity was indicated from NFI photo-plot data, where a forest fire occurred during the period of analysis (as indicated in the Canadian National
Fire Database; Stocks et al., 2002), or where no known deforestation (on the basis of NFI land-use and deforestation information) had occurred. Remaining sites were examined by image interpreters to ensure the accuracy of the final land-use classification.

Parameter estimates were calculated separately for Canada and integrated into analyses of FRA regions and FAO climatic domains.

**Russian Federation**

The Russian Federation used a stratified sample of 300 RSS sample sites to estimate forest area and forest area change for the three survey periods. A total of 1,961 complete RSS sample sites were contained within the Russian Federation. Landsat data were available for 1,219 of these for all three time periods; this incomplete coverage is due to the lack of satellite data acquisitions for the eastern part of the Russian Federation in 1990. Although all 1,961 sample sites were processed to the extent possible using the methods described in the main body of this report, expert review and revision of all sample sites in the Russian Federation was not possible in the timeframe of the study.

Cloud-free, seasonal 250 m spatial resolution data from MODIS were used, along with vegetation change indices, to create 23 strata according to percentage forest cover and amount of indicated change in forest cover. A probability-based selection process was implemented to select the final plots for review and revision based on a minimum separating distance (i.e. plots were preferred to be further apart within any single stratum) and minimum number (ten) per stratum. A total of 282 RSS sites were expertly interpreted for land-cover and land-use classification.
The parameter estimates and statistical variance of the stratified sample were incorporated with those of the systematic sample for Europe and used in analyses of the boreal climatic domain.

**United States of America**

RSS results for the United States of America were derived from the National Land Cover Dataset (NLCD) (Vogelmann *et al.*, 2001; Homer *et al.*, 2004). The NLCD is a 21-class land-cover product for the conterminous United States based on Landsat satellite data. The 21 classes were reduced to the five simple land-cover classes required for RSS 2010. Level-2 segments for 1990, 2000 and 2005 were assigned land-cover labels directly from the NLCD dataset for each survey period. Land-cover labels were adjusted to land use using the automated conversion rules described in the main body of this report. A probability-based sample of sites, by FAO climatic domain, was selected for review and revision for continental United States and Alaska. At each review-and-revision site, the accuracy of the land-use call was evaluated against the NLCD and high-resolution aerial photography. The results of the accuracy assessment were used to adjust the overall area of land-use category for the United States in its entirety and for each FAO climatic domain.
Annex 5-2  Survey sites processed vs analysed

The table below lists, by region or country-specific grouping, the number of sample sites processed (grand total), analysed and not analysed. The main reason that survey sites were not analysed was missing data in one or more time periods due to cloud cover, a lack of satellite image acquisitions, or other data anomalies.

<table>
<thead>
<tr>
<th>Region/country</th>
<th>Analysed</th>
<th>Not analysed</th>
<th>Grand total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td>2 322</td>
<td>196</td>
<td>2 518</td>
</tr>
<tr>
<td>Asia</td>
<td>2 863</td>
<td>184</td>
<td>3 047</td>
</tr>
<tr>
<td>Canada</td>
<td>3 737</td>
<td>315</td>
<td>4 052</td>
</tr>
<tr>
<td>Europe</td>
<td>625</td>
<td>55</td>
<td>680</td>
</tr>
<tr>
<td>Oceania</td>
<td>769</td>
<td>29</td>
<td>798</td>
</tr>
<tr>
<td>Russian Federation</td>
<td>282</td>
<td>1 679</td>
<td>1 961</td>
</tr>
<tr>
<td>South America</td>
<td>1 372</td>
<td>129</td>
<td>1 501</td>
</tr>
<tr>
<td>North &amp; Central America</td>
<td>1 096</td>
<td>126</td>
<td>1 222</td>
</tr>
<tr>
<td><strong>Grand Total</strong></td>
<td><strong>13 066</strong></td>
<td><strong>2 713</strong></td>
<td><strong>15 779</strong></td>
</tr>
</tbody>
</table>
Annex 5-3  Review-and-revision contributors


**North America:** M. Gillis, S. Healey, C. Meneses-Tovar


**West Asia:** S. Chukumbaev, H. Samadi, M. Shojalilov


**Oceania:** C. Howell, P. Lane, M. Mutendeudzi
Summary of national and regional review-and-revision workshops

<table>
<thead>
<tr>
<th>Area of interest</th>
<th>Location</th>
<th>Date</th>
<th># countries</th>
<th># experts</th>
<th>Female</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brazil</td>
<td>S J dos Campos</td>
<td>Sep 2009</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Central Africa</td>
<td>Brazzaville</td>
<td>February 2010</td>
<td>8</td>
<td>16</td>
<td>1</td>
<td>15</td>
</tr>
<tr>
<td>North America</td>
<td>Salt Lake City</td>
<td>March 2010</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>South Africa</td>
<td>Cape Town</td>
<td>March 2010</td>
<td>8</td>
<td>16</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>West Europe</td>
<td>Rome</td>
<td>March 2010</td>
<td>14</td>
<td>14</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>Central America</td>
<td>Panama City</td>
<td>July 2010</td>
<td>7</td>
<td>12</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>Southeast Asia</td>
<td>Bangkok</td>
<td>August 2010</td>
<td>14</td>
<td>23</td>
<td>5</td>
<td>18</td>
</tr>
<tr>
<td>East Asia</td>
<td>Beijing</td>
<td>Sep 2010</td>
<td>3</td>
<td>16</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>South America</td>
<td>Valdivia</td>
<td>Nov 2010</td>
<td>7</td>
<td>14</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>France</td>
<td>Nogent</td>
<td>Nov 2010</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>East Africa</td>
<td>Nairobi</td>
<td>Dec 2010</td>
<td>6</td>
<td>11</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>West Africa</td>
<td>Dakar</td>
<td>March 2011</td>
<td>13</td>
<td>18</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>New Zealand</td>
<td>Rome</td>
<td>March 2011</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Australia</td>
<td>Canberra</td>
<td>April 2011</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>East Europe</td>
<td>Budapest</td>
<td>May 2011</td>
<td>9</td>
<td>14</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>Ireland/Latvia</td>
<td>Teleconference</td>
<td>May 2011</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Sudan</td>
<td>Khartoum</td>
<td>May 2011</td>
<td>1</td>
<td>18</td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td>Brazil</td>
<td>Campinas</td>
<td>June 2011</td>
<td>1</td>
<td>9</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Italy</td>
<td>Rome</td>
<td>June 2011</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>West Asia, North Africa</td>
<td>Rome</td>
<td>July 2011</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Russian Federation</td>
<td>Moscow</td>
<td>September 2011</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td><strong>107</strong></td>
<td><strong>204</strong></td>
<td><strong>36</strong></td>
<td><strong>168</strong></td>
</tr>
</tbody>
</table>
1. For every sample site, the following variables were extracted from the PostGreSQL database:

- tile unique ID (rss_id)
- latitude (lat) and longitude (lon) of the centre of the tile
- climatic domain (domain)
- region (continent)
- total tile area (total)
- water area (water)
- no data area (nodata)
- forest area in 1990, 2000 and 2005 (forest90, forest00, forest05)
- area of gains and losses of forest in 1990-2000 and 2000-2005 (gain9000, loss9000, gain0005, loss0005)
- Julian date of image acquisition for 1990, 2000, 2005 (jdate90, jdate00, jdate05)

2. Then, the following variables were calculated:

- Area of land within the tile (gla)
  \[ gla = total - water - nodata \]
- Latitude correction factor (corrlat)
\[
\begin{align*}
\text{if } \text{lat} \leq 60^\circ \text{ then } \text{corrlat} &= \cos(\text{lat}) \\
\text{if } \text{lat} > 60^\circ \text{ then } \text{corrlat} &= 2 \times \cos(\text{lat})
\end{align*}
\]

NB: The number of samples was reduced to include only even degrees of longitude above 60 degrees latitude (Figure 5-1 shows the thinning of samples at high northern latitudes).

- Weight of the sample \(i\) \((w_i)\)

\[
w_i = \frac{\text{gla}_i \times \text{corrlat}_i}{\sum_j \text{gla}_j \times \text{corrlat}_j}
\]

- Proportion of forest in 1990 \((p_{for90})\)

\[
p_{for90} = \frac{\text{forest90}}{\text{gla}}
\]

- Proportion of forest in 2000 \((p_{for00})\)

\[
p_{for00} = \frac{\text{forest00}}{\text{gla}}
\]

- Proportion of forest in 2005 \((p_{for05})\)

\[
p_{for05} = \frac{\text{forest05}}{\text{gla}}
\]

- Annualized proportion of gains, losses and net change for 1990–2000

\((p_{again9000}, p_{aloss9000}, p_{anet9000})\)
\[
\begin{align*}
\text{pagain9000} &= \frac{\text{gain9000}}{\text{gla} \times (\text{jdate00} - \text{jdate90})} \\
\text{paloss9000} &= \frac{\text{loss9000}}{\text{gla} \times (\text{jdate00} - \text{jdate90})} \\
\text{panet9000} &= \text{pagain9000} - \text{paloss9000}
\end{align*}
\]

NB: \text{pagain0005}, \text{paloss0005} and \text{panet0005} are calculated in the same way.

3. For any subset \( S \) of samples (e.g. one climatic domain), average value \( (\bar{x}) \) and standard deviation \( (\text{std}) \) of \( \text{pgf90}, \text{pgf00}, \text{pgf05}, \text{pagain9000}, \text{paloss9000}, \text{panet9000}, \text{pagain0005}, \text{paloss0005} \) and \( \text{panet0005} \) were calculated with the survey package of R (Lumley, 2004) using the following formula:

\[
\bar{x} = \frac{\sum_{i \in S} w_i \times x_i}{\sum_{i \in S} w_i}
\]

\[
\text{std} = \sqrt{\frac{\sum_{i \in S} w_i \times (x_i - \bar{x})^2}{\sum_{i \in S} w_i}}
\]

4. Final values (e.g. of annual loss in forest area between 1990 and 2000 in a given climatic domain) was obtained by multiplying the average and the standard deviation by the area of the region \( (A) \):

\[
\text{loss} = \text{paloss9000} \times A \pm 1.96 \times \frac{\text{std(paloss9000)}}{\sqrt{N}} \times A
\]
CHAPTER 6

RESEARCH SUMMARY AND RECOMMENDATIONS
6.0 Summary of Chapters

The research conducted and presented in this dissertation addressed the methods, limitations and results of assessing forest land-use change globally from a sample of medium spatial resolution satellite imagery.

Chapter 3 addressed research question 1, how do cloud, cloud shadow, haze and missing data specifically affect the amount of data required to create spatially exhaustive composites suitable to monitor land cover and land-use changes over time? The results of this research found that wall-to-wall mapping of forest area and change with medium spatial resolution Landsat imagery required a considerable amount of input imagery (> 5 acquisitions per path/row) to achieve cloud-free coverage within the humid tropics. Such data requirements may preclude the production of annual or bi-annual updates of spatially exhaustive forest area and change estimates and may overly complicate the calculation of change rates and accuracy assessments due to missing data and the effects of pixel-based compositing with multi-date acquisitions. Additionally, the increased data volume may be overwhelming when producing forest area and change estimates over large areas without the benefit of high power computing facilities. These findings point to the possible benefits of sample-based estimation methods that do not strictly require complete areal coverage within sample areas.

Chapters 4 and 5 of this dissertation explored a sample-based method to generate areal estimates of forest land-use and change for the years 1990 – 2000 – 2005 at regional, continental and global scales. The survey used single-best date imagery and object-based image analysis (OBIA) to produce results for each epoch of analysis. Sampling with
single-best date imagery overcame problems associated with spatially exhaustive mapping including missing data due to lack of acquisitions, cloud, cloud-shadow or other quality issues. OBIA methods facilitated the review and revision of land-use classifications at each sample site by over 300 international experts. The objects, or polygons, provided meaningful delineations of the Earth’s surface that could be more easily identified in functional terms, e.g. an agricultural field or tree plantation, than individual pixels. Experts were able to use the functional clues provided by the objects in conjunction with their local knowledge to classify land-use for each sample site.

Chapter 4 assessed the strengths and weaknesses of this processing methodology with the results presented in Chapter 5. The object-based image analysis, automated classification procedure and expert review and revision were revisited to assess (i) whether the image segmentation and aggregation to a 5 hectare MMU introduced a systematic bias in the automated results, (ii) the effects of the subsequent expert review and revision on the results of the automated classification, (iii) the efficiency gained by supplying automatically derived land cover classification results for review and revision and (iv) the effect of image segmentation and aggregation to a 5 hectare MMU on fine-scale change detection.

The methods provided an efficient means of processing over 11,000 sample sites, 33,000 Landsat 20x20 km sample tiles and more than 6.5 million individual polygons over three epochs. Objects were first classified to land cover using a supervised and largely automated process. The goal of the automated classification was to ease the burden of human interpretation, the ultimate step in the classification process. Expert human interpreters from around the world were then enlisted to review and revise, in a
two-step process, first the land cover classification and then again, the subsequent land-use classification after the corrected land cover had been converted to land-use using simple decision rules. A comparison of automated classification results with expert-corrected results for land cover yielded agreements of approximately 80%, e.g. only 20% of the object labels had to be revised by expert human interpretation. The automated land cover classification methods thus provided suitable results and dramatically decreased the amount of time necessary for expert review and revision.

Error was introduced into the automatically derived segment-based results, however, when compared to pixel-level estimates. Larger segment sizes are typically more heterogeneous spectrally and thus, depending on the rules governing class aggregation, may over or underestimate class area. Image segmentation applied to medium spatial resolution imagery for change detection was susceptible to errors of omission when change was very fine-scale or was characterized by a complex shape (e.g. high perimeter to area ratio).

Implementing a relatively large MMU (5 ha) facilitated expert human interpretation, review and revision of the results and overcame issues related to fine-scale spectral heterogeneity. However, if the local change dynamic is very fine-scale, a large MMU applied to medium spatial resolution imagery may be incapable of adequately capturing this. In these areas, a strict adherence to a MMU may not be helpful and employing a much smaller or no MMU is advisable.

The methods were novel because they focused on characterizing global forest land-use and change, not only tree cover. Land use, as previously mentioned, references the way humans relate to and use a particular area of land in terms of economic function.
As such, it does not readily lend itself to automated monitoring from remote sensing instruments as it is not strictly classifiable at the instantaneous time of sensor overpass.

Expert human examination of the sample sites was critical to convert land cover to land use labels; this interactive effort recommends sampling as a method for land use estimation as wall-to-wall interpretation of land use is impractical. Land use classification is extremely important and is particularly useful information when attempting to account for emissions or sequestration of carbon dioxide. For example, a forest fire may temporarily remove vegetation overstory and subsequently kill the trees in a forest stand. Immediately post-burn, the land cover designation of this area may be classified as non-tree cover. However, the long-term land use of this area has not changed and remains forest. To account for this area as deforestation and a net carbon emission source, would be an error and likely overestimate total atmospheric carbon dioxide emissions.

Chapter 5 presented the results of the method described in Chapter 4 and addressed research question 3, what is the global extent of forest land-use, how does the area of forest land-use differ by geographic region and major climatic zone, and how have these areas changed over time? Chapter 5 presented global, climatic domain and regional estimates of forest land-use area and change between 1990 and 2005 based on a systematic sample of Landsat medium spatial resolution satellite imagery. The survey estimated the total area of the world’s forests in 2005 at 3.8 billion hectares, or 30 % of the global land area. Overall, there was a net decrease in global forest area of 1.7 % between 1990 and 2005, at an annual rate of change of 0.11 %. This equates to an annual shift from forest land-use to other land uses of 3 million hectares per year between 1990
and 2000 and of 6 million hectares per year between 2000 and 2005.

Major regional differences were found in the net rates of forest area change – only Asia and North America experienced gains in forest area. All other regions saw net declines. South America had the highest net forest loss - some 3.3 million hectares annually between 1990 and 2005. Africa had the second highest net forest loss of 1.6 million hectares annually during the same period. Europe, including the Russian Federation, had net losses of 0.5 million hectares annually and Oceania lost just under 0.1 million hectares annually. North America experienced net gains in forest area of some 0.2 million hectares annually while Asia had a net gain of 1.4 million hectares annually between 1990 and 2005.

Forests were categorized according to four climatic domains: boreal, subtropical, temperate and tropical. There were significant gains in forest land-use in the boreal (0.9 million hectares annually) and subtropical (1.1 million hectares annually) climatic domains between 1990 and 2005. In the case of the Boreal domain, gains in forest land-use were largely attributable to forest encroachment on abandoned agricultural lands (Lambin and Meyfroidt, 2011; Keummerle et al., 2010). Estimates for the boreal, however, could be affected by a large area of missing data in eastern Russia where no Landsat acquisitions exist for the 1990 epoch and where results were extrapolated from sites where data existed. Gains in forest land-use area in the subtropical domain were largely due to forest gazetting and planting in China (Liu and Tian, 2010). There were also net gains in forest area in the temperate domain of 0.9 million hectares for this period. In contrast, the tropical domain had a net loss of forest area of 6.8 million hectares annually between 1990 and 2005. This net reduction in forest land-use was
nearly 2.5 times the net forest area gained in the other three domains.

The results of this study improved on previous estimates of global forest land-use in three main ways: (i) by producing consistent and comparable figures of forest land-use over time including estimates of uncertainty, (ii) enabling aggregation and estimation of results at a variety of spatial scales and (iii) fostering national ownership of the process and results by engaging national experts for review and revision of sample sites.

6.1 Recommendations for future research or work

6.1.1 Decrease uncertainties in estimates of forest cover, use and rates of change.

Estimates of the global areal extent of forest land-use, land cover and their changes over time are still highly uncertain (Table 6-1) (Sexton et al., 2015; Keenan et al., 2015). In a comparison of year 2000 forest area estimates, forest defined as a land cover varies by almost 20 %, from 32.7 to 41.5 million square kilometers. Forest defined as a land use varies by 6 %, from 38.2 to 40.8 million square kilometers. Likewise, estimates of global gross forest cover loss between year 2000 and 2005 vary from 0.88 to 1.0 million square kilometers, a difference of 12%. Estimates of global forest land-use loss for the same period taken from FRA reports are much lower than those for forest land cover and range from 0.23 to 0.24 million square kilometers or, when results from the global remote sensing survey are included, 0.32 million square kilometers, a difference of 4% and 28%, respectively.

More worrying, however, is that even estimates of the trend in forest land-use area change between the 1990s and the 2000s are uncertain, with some claiming the rate of
forest loss is increasing (Lindquist et al., 2012) and others claiming the rate is decreasing (FAO, 2010; FAO 2015). The increasing trend in forest land-use, however is mirrored by several studies which indicate increasing rates of tree cover loss for similar periods (Kim et al., 2015). The three main factors contributing to the uncertainty in global forest area and change estimates are likely to be (i) the definition of forest used in each analysis, (ii) the survey methodology and (iii) the limitations of medium spatial resolution data to detect relatively sparse tree canopy cover (Lindquist et al., 2012; Coulston et al., 2013). Regarding (i) above, there is a definite need for alignment between forest definitions used for international reporting requirements (for REDD+, for instance) and the methods used to monitor and report on that forest area (Sexton et al., 2015). Lund (2014) documents over 1500 different, operational definitions of the term ‘forest’ and over 200 different definitions of the term ‘tree’. Some of these are land cover (e.g. biophysical) definitions and others are based on land use. Ideally, consensus on forest definitions could be reached among researchers and between applications such that, if results differ, a readily available explanation exists as to why.

Regarding point (ii), methods for generating forest land-use and tree cover estimates vary widely in methodology and scope and range from spatially explicit, wall-to-wall mapping of tree cover only, sample-based assessments of tree cover and forest land-use and survey responses provided to formal questionnaires for forest land-use. Hansen et al. (2013) used a pixel-based, wall-to-wall mapping approach with Landsat data to produce maps and estimates of tree cover and annual tree cover losses globally. Tree cover was estimated on a continuous percentage basis representing the per-pixel tree canopy cover. Gong et al. (2013) also used a pixel-based, wall-to-wall mapping
approach with Landsat data but used distinct land cover class labels to distinguish
tree-dominated classes from other land cover classes. Hansen et al. (2010) used a
stratified random sampling approach with Landsat data and strata derived from a MODIS
land cover change product to estimate tree cover losses at national and biome levels.
FAO and JRC (2012 [chapter 5 of this dissertation], 2014) used a degree latitude-
longitude survey of Landsat data and OBIA to estimate tree cover, forest land-use area
and change over time globally, by continental/country groupings and by large climatic
domains. FAO (2010, 2015) used responses submitted by individual countries in
response to a questionnaire regarding the area of forest land-use for the nominal reporting
cannot be ignored and, though no formal assessment has been carried out determining the
affect on estimates, contribute to the differences in the results obtained.

With respect to point (iii), many countries choose forest definitions with low
thresholds for tree cover (e.g. 10 % canopy cover in the case of the FAO definition, for
instance) or land use characteristics that simply cannot accurately be monitored with
existing satellite sensors, especially using highly automated approaches. As chapter 5 of
this dissertation shows, these difficulties in monitoring capabilities occur frequently in
zones of dryland forest. Inconsistencies in the area estimates of dryland forests
contribute a great deal to the uncertainty of estimating global forest area (Sexton et al.,
2015; Lindquist et al., 2012). In the Global Forest Resources Assessment from 1980,
(FAO, 1980), the area of open woodlands was 38 % of the total area of forest in the
tropical climatic domain. The physical characteristics of dryland forest including sparse
canopy covers, seasonal phenology and an understory generally characterized with a high
albedo, make them difficult to accurately map with medium resolution satellite imagery (Giri et al., 2005; Lindquist et al., 2012; Bodart et al., 2013, Gong et al., 2013). It is likely that the increased spatial resolution of optical remote sensing instruments, such as the just-launched Sentinel-2 mission (Drusch et al., 2012), will enable a more accurate and consistent characterization of dryland forests.

**Table 6-1. A comparison of recent global forest cover and land use area and change estimates. All areas in million sq. km. Net figures indicated with (*).**

<table>
<thead>
<tr>
<th>Forest Land Cover</th>
<th>Global Forest Area 2000</th>
<th>Global Gross Forest Loss 2000 - 2005</th>
<th>Forest Definition</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forest Area</td>
<td>32.7</td>
<td>1.01</td>
<td>&gt;25% Canopy</td>
<td>Hansen et al., 2010</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>&gt; 5 m Height</td>
<td></td>
</tr>
<tr>
<td></td>
<td>37.3</td>
<td>N/A</td>
<td>Clear stem observable</td>
<td>Gong et al., 2013</td>
</tr>
<tr>
<td></td>
<td>41.5</td>
<td>0.88</td>
<td>&gt;25% Canopy</td>
<td>Hansen et al., 2013</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>&gt;5 m Height</td>
<td></td>
</tr>
</tbody>
</table>
### Forest land-use

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>&gt;10% Canopy</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>40.8</td>
<td>0.24*</td>
<td>&gt;10% Canopy</td>
<td>FAO, 2010</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;5 m Height</td>
<td></td>
</tr>
<tr>
<td>38.2</td>
<td>0.68, 0.32*</td>
<td>&gt;10% Canopy</td>
<td>Lindquist et al., 2012</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;5 m Height</td>
<td></td>
</tr>
<tr>
<td>39.5</td>
<td>N/A</td>
<td>&gt;10% Canopy</td>
<td>FAO and JRC, 2014</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;5 m Height</td>
<td></td>
</tr>
<tr>
<td>40.6</td>
<td>0.23*</td>
<td>&gt;10% Canopy</td>
<td>FAO, 2015</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;5 m Height</td>
<td></td>
</tr>
</tbody>
</table>

### 6.1.2 Improve operational methods of integrating high spatial resolution optical data

Chapters 4 and 5 of this dissertation highlighted the utility of using OBIA techniques and statistical sampling designs with medium spatial resolution optical imagery to produce consistent estimates of global forest land-use. The advantage of OBIA being that expert interpreters were presented with image objects more likely than individual pixels to represent identifiable, ground-based phenomena linked to a particular process occurring on that land. A statistical sampling scheme meant that these same interpreters could spend more time distinguishing land use from mapped land cover and make a more careful interpretation at each sample site than would be possible for spatially exhaustive mapping.
High spatial resolution optical imagery is increasingly used for large-area forestry applications, notably for estimating finer-scale forest characteristics (Boyle et al., 2014) and monitoring more subtle forest area changes (Bholaneth & Cort, 2015). Increased resolution should improve interpretations of land-use as anthropogenic activity, crop type and re-growth or planting, for example, should be more easily observable. Thus estimates of forest land-use should also improve.

OBIA and statistical sampling techniques will be extremely important for the operational characterization of high spatial resolution data. OBIA will be more important for classifying the higher resolution data because pixels are more likely to be much smaller than the actual target of interest, e.g. a tree canopy or small clearing (Blashke and Strobl, 2001). Thus, targets for classification will be composed of multiple pixels with similar characteristics. OBIA will be used for intelligent aggregation of these separate pixels into objects more accurately classified into the target of interest than would otherwise be possible with traditional pixel-based approaches. Statistical sampling of these data will be an efficient means of producing well-interpreted, reliable results with known precision. Estimates generated from sampling high spatial resolution imagery will be used for improved estimates of many forest land cover and land use characteristics not currently reliably detectable.

6.1.3 Ensure continuous and frequent acquisition of Landsat-like data

The need for spatially explicit, continuous monitoring of the Earth’s land surface will not diminish in the coming years. This will necessitate at least two items, neither of which are particularly novel, but which bare repeating and elucidating here
including (i) sustained image acquisitions at a more frequent time interval and (ii) increased computer processing power and storage capacity for data users.

Continuous and more frequent earth observations are critical to maintain the long-term satellite earth surface data record and to improve on limitations in the current image archive due to the 16-day revisit period of the Landsat sensor and effects of cloud cover (Lindquist et al., 2008 [Chapter 3 of this dissertation]; Ju and Roy, 2008; Roy et al., 2014; Kovalskyy and Roy, 2014). The earth science community has argued for the continuity of Landsat and Landsat-like observations for some time (Wulder et al., 2008). The recent (2013) launch of Landsat-8 as the successful culmination of the Landsat Data Continuity Mission (Irons et al., 2012; Roy et al., 2014) will hopefully provide high-quality acquisitions for ‘no less than 5 years’. The European Space Agency (ESA) will also contribute to the continuation of earth observation data with the newly operational Sentinel-2 satellites (Drusch et al., 2012; Malenovsky et al., 2012). These satellites will provide medium to fine-scale spatial resolution observations compatible with Landsat (Roy et al., 2014) with a 5-day revisit period at the equator for a planned 15-year time period.

The United States recently initiated the Sustained Land Imaging Program to ensure acquisition of Landsat or Landsat-like Earth observations into the foreseeable future (NASA, 2015; Loveland and Dwyer, 2012). The program evolved from a 2007 interagency working group’s exploration into formalizing the requirements for a well-structured and governed national land-imaging program (FLIIWG, 2007; Freeborn, Green and Lauer, 2006). The original working group’s results were then further elaborated and specific recommendations were made on how to achieve sustained land
imaging by the Committee on Implementation for a Sustained Land Imaging Program (National Research Council, 2013). Landsat 9 is now in development with a planned launch date in year 2023 (USGS, 2015).

### 6.1.4 Increased access to high performance computing facilities

Increased access to high-powered computing and large storage infrastructures will have to be made available to more users in order to store and process the increasing volumes of data required to produce useful information products (Giri et al., 2013; Roy et al., 2014). The Sentinel-2 mission will produce almost 2 Terabytes per day of imagery data (in compressed form) (Drusch et al., 2012) and Landsat-8 compressed scenes will contribute another 400 to 600 gigabytes per day.

Large computing infrastructures are, however, becoming available and rapidly changing the way geospatial science is carried out. Hansen et al. (2013) processed over 600,000 Landsat scenes to produce the first, global, wall-to-wall depiction of forest cover and annual forest cover gains and losses between 2000 and 2012 at the Landsat scale. To produce this result, the authors utilized over a million CPU hours and were able to divide their processing in a highly parallelized manner between 10,000 different computers. Roy et al. (2010) created monthly, seasonal and annual wall-to-wall composites of Landsat acquisitions corrected to surface reflectance over the conterminous United States. The authors processed thousands of individual Landsat acquisitions to complete the spatially continuous coverage and now distribute these data free of charge to the general public via a system called Web-Enabled Landsat Data or WELD. The WELD datasets
will become available globally as well. Data such as these will allow applied research to advance more quickly because scientists can spend a proportionately larger amount of time analysing results instead of preparing remotely sensed imagery to be analysed (Hansen and Loveland, 2012).

Finally, perhaps most easily and most importantly, the ability to access all of the new data and the high-powered processing tools ought be made available as well in developing countries. However, the challenges to data access and processing remain much the same as they were as documented by Roy et al. (2010) with lack of sufficient internet bandwidth being chief, currently, among the limitations.

6.1.5 Exploit the power of the global Landsat image archive via time series analyses to automatically produce land use classifications

The Intergovernmental Panel on Climate Change (IPCC) Good Practice Guidelines for Land-use change and Forestry (Penman et al., 2003) specifies six major land use types that must be accounted (e.g. area estimates and changes) for reporting on carbon stock changes as part of the UNFCCC REDD+ reporting process. These land use classes are forestland, cropland, grassland, wetland, settlements and other land. Chapter 5 of this dissertation and published as Lindquist et al., 2012, represents the first consistent and systematic assessment of global forest land-use. There is an urgent need to produce assessments of the other 5 land use categories as well, including transitions between them (Houghton, 2010).

At the moment, the problem of delineating land use, as opposed to land cover, is best and perhaps uniquely solved by the intervention of human expert interpretation
(Kurz 2010, Blashke et al. 2014). However, advancing time-series techniques are making the classification of human land uses and change much more reliable. There are numerous examples now of how tracking individual pixel reflectance values over time can lead to a more detailed classification of an area not just in terms of uni-directional change events (e.g. complete overstory removal) (Kennedy, Cohen and Schroeder, 2007; Zhu and Woodcock, 2014) but also of more subtle changes including tree cover regrowth (Kennedy, Yang and Cohen 2010) and degradation (Vogelmann et al., 2013). Yan and Roy (2014) use a dense time-series of fully corrected and converted to surface reflectance Landsat data in combination with image segmentation algorithms to accurately map and describe agriculture field shape and size for several locations within the United States. The authors take advantage of the regular seasonal signature of agricultural cropping and uniformity of field shapes to produce maps classified as a land use. It is likely that, as mining of the Landsat archive and computing power increase, so too will the ability to differentiate land use from land cover information in satellite imagery without the necessary intervention of human analysts.
6.2 References


Ball, G. H., & Hall, D. J. (1965). ISODATA, a novel method of data analysis and pattern classification. STANFORD RESEARCH INST MENLO PARK CA.

Bartolev, S. 2012. Unpublished data on accuracy of FRA RSS land cover classification compared to high spatial resolution satellite imagery.


doi:10.1016/j.isprsjprs.2009.06.004.


Czaplewski, R. L. (2003). Can a sample of Landsat sensor scenes reliably estimate the


Foody, G. M. (2000). Mapping Land Cover from Remotely Sensed Data with a Softened Feedforward Neural Network Classification. *Journal of Intelligent and Robotic*


Horler, D.N.H., & Ahern, F.J. 1986. Forestry information content of Thematic Mapper


Landsat-7 ETM+ Automated Cloud-Cover Assessment (ACCA) Algorithm.


Kennedy, R.E. and Cohen, W.B., (2003). Automated designation of tie-points for image-


Lambin, E., Turner, B., Geist, H., Agbola, S., Angelsen, a, Bruce, J., Coomes, O., et al. (2001). The causes of land-use and land-cover change: moving beyond the myths. *Global Environmental Change*, 11(4), 261-269. doi:10.1016/S0959-3780(01)00007-3.


NAFORMA (in prep). National Forest Resources Monitoring and Assessment Final Report to the Tanzania Ministry of Natural Resources and Tourism - Forestry and Beekeeping Division. United Nations Food and Agriculture Organization.


means from two-stage sampling. *Fisheries Research* 107(1–3): 1–13. DOI:


Pimm, S. L., Jenkins, C. N., Abell, R., Brooks, T. M., Gittleman, J. L., Joppa, L. N., …


van der Werf, G. R., Morton, D. C., DeFries, R. S., Olivier, J. G. J., Kasibhatla, P. S.,


Walpole, M., & Al., E. (2009). Tracking Progress Toward the 2010 Biodiversity Target


Wulder, M.A., Dechka, J.A., Gillis, M.D., Luther, J.E., Hall, R.J., Beaudoin, A. &


