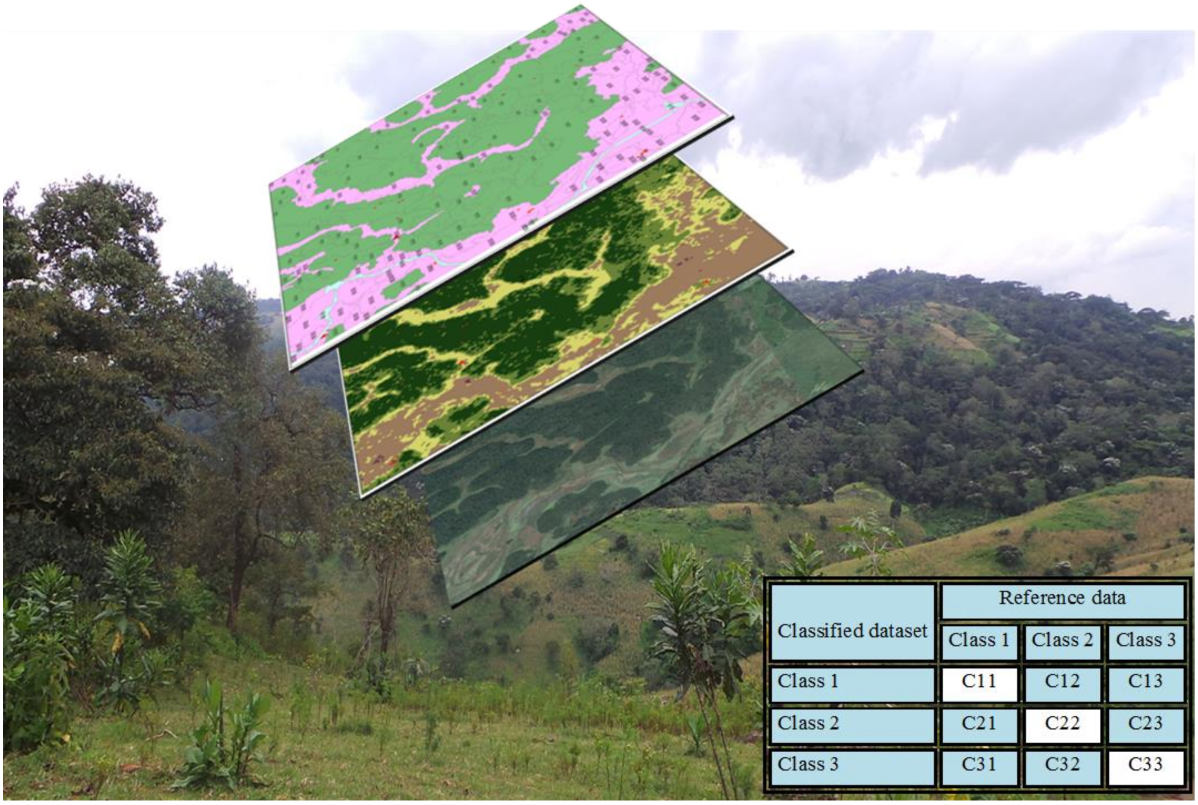
Centre for Geo-information

Minor thesis report GIRS-2014-29

Assessing the applicability of global forest cover change datasets to national REDD+ forest monitoring: The case of Ethiopia

Dereje Likissa Beyene

August 2014



WAGENINGEN UNIVERSITY
WAGENINGEN UR

Assessing the applicability of global forest cover change datasets to national REDD+ forest monitoring: The case of Ethiopia

Dereje Likissa Beyene

Registration number 77 04 21 064 010

Supervisors:

Benjamin DeVries

Prof. dr. Martin Herold

A thesis submitted in partial fulfilment of the degree of Master of Science
at Wageningen University and Research Centre,
The Netherlands.

August 2014

Wageningen, The Netherlands

Thesis code number: GRS-80424
Thesis Report: GIRS-2014 -29
Wageningen University and Research Centre
Laboratory of Geo-Information Science and Remote Sensing

Acknowledgements

First and for most, my praise goes to the almighty God that let me stay in life and gave me health to do such fruitful work. Next, I would like to express my deep gratitude sincere thanks to Benjamin DeVries and Prof.dr. Martin Herold. Their constructive comments, criticisms and wide open suggestions in line with my thesis work were very crucial. This thesis would not be effective without their close supervision and innovative ideas. I also want to extend my thanks to Erika Romijn for her responsive comment in line of REDD+ activities and forest definitions contexts and also to Niki De Sy for providing feedback during the proposal development and compiling FAO remote sensing survey data for the purpose of this study.

Lastly my appreciation for my wife Abebech Desisa is credible for her unfailing affection and responsibility sharing with my kids in the course of my study. Her strong initiation and encouragement makes my dream to become real. Nothing equals my love and admiration to her.

Abstract

In recent years, several global land cover change datasets have been derived from various spatial and temporal resolutions of remotely sensed data. Such datasets with accurate and up to date information are necessary for REDD+ land cover change monitoring activities in developing countries most importantly where data gaps are huge. Yet, to complement these datasets to national forest monitoring systems (NFMS) and measuring, reporting and verification (MRV) requirements of REDD+ an assessment of how well locations of mapped change would correspond to actual areas of change on the ground is crucial. Two existing global land cover/use maps derived from different initiatives such as FAO remote sensing survey (RSS) and Hansen et al. global forest cover change (GFCC), with the same remote sensing data sources and spatial resolution were compared to evaluate their accuracy and their adaptability to NFMS. Other external data sources were also explored to see their definition on forest and forest cover change.The study employed two sampling strategies, namely systematic sampling strategy to get access to available RSS tiles and simple random sampling to incorporate Hansen et al. GFCC datasets outside the selected tiles. Our results showed that forest definitions and methodologies for each dataset employed varies, underpinning the variation of forest cover extent and forest loss rate for the same period in Ethiopia. The accuracy of 5 land cover classes from change and no change category were compared and the analysis result indicate that 61% of Hansen et al. GFCC and 80% of FRA RSS datasets showed overall accuracy and agreement with reference data and reasonably acceptable. While, the overall spatial agreement between the two datasets was 53%, which is fairly less than the accuracy observed. With regard to classes stable forest has high agreement with reference data in both datasets with 71% and 83% user accuracy and 62% and 85% producer accuracy in GFCC and RSS, respectively. While other woodland class, is poorly classified in the former case with 25% producer and 51% user accuracy. The reason for more misclassification in Hansen et al. dataset was confusion of land cover classes mainly with agricultural areas and where tree cover is <50%. In conclusion, although both datasets are consistent in their methodology for adaptability, RSS dataset is more accurate and comparable, while Hansen et al. GFCC dataset is more complete and transparent to be used for MRV of REDD+ activities at national level in Ethiopia.

Keywords: Forest cover, Forest monitoring, datasets, REDD+, accuracy, applicability, Ethiopia

Index of Contents

Index of Figures

Index of Tables

Acronyms and Abbreviations

CRGE	Climate Resilient Green Economy
CBD	Convention on Biological Diversity
ETM+	Enhanced Thematic Mapper (plus)
FAO	Food and Agricultural Organisation
FDRE	Federal Democratic Republic of Ethiopia
FRA	Forest Resource Assessment
GEI	Google Earth Imagery
GFCC	Global Forest Cover Change
GHG	Greenhouse gas
GPG	Good Practice Guidance
IPCC	Intergovernmental Panel on Climate Change
LULCC	Land use land cover change
MA	Marrakech Accord
MRV	Measuring, Reporting and Verification
NASA	National Aeronautics and Space Administration
NFMS	National Forest Monitoring System
ODK	Open Data Kit
REDD	Reducing Emissions from Deforestation and forest Degradation
R-PP	Readiness Preparation Proposal
RS	Remote Sensing
RSS	Remote Sensing Survey
TCC	Tree Canopy Cover
TM	Thematic Mapper
UNFCCC	United Nations Framework Convention on Climate Change
USGS	United State Geological Survey
WBISPP	Woody Biomass Inventory and Strategic Planning Project

1. Introduction

1.1 Background

Forest cover change science has emerged as key element of global environment change giving emphasis to changes in land use and forest cover, not only in terms of their type and magnitude but also in location of changes as well (Rindfuss et al. 2004). In this respect the role of remote sensing domain is huge enabling spatially comprehensive and wall to wall coverage of forest cover change analysis. Estimates of extent of deforestation derived via remote sensing plays a prominent role in ongoing efforts to establish scientifically valid procedures for forest change monitoring, specifically accounting to policy approaches for reducing greenhouse gas emissions from forests (DeFries et al. 2007; Olofsson et al. 2014).

Ethiopia is one among developing countries that incorporated Reducing Emissions from Deforestation and forest Degradation (REDD+) forest monitoring approach as an important component of Climate Resilient Green Economy (CRGE) strategy. The strategy is designed to foster development protecting forest and reducing deforestation from forestry sector (FDRE 2011a, b). In order to report on this REDD+ activities measurement, reporting and verification (MRV) road map has been developed and thus requiring an appropriate national forest monitoring system (MoA 2013). Besides this fact, historical reference datasets on forest cover and land use is problematic in Ethiopia for MRV of REDD+ activities (Moges et al. 2010), and still Ethiopia is among the 49 tropical non-Annex I countries that have a very large capacity gap in forest monitoring (Romijn et al. 2012). In such case, there is a need to complement existing data sets to the national forest monitoring and MRV framework.

Several remote sensing applications have been developed for capturing forest dynamics to meet a range of information needs at global scale. Assessing the accuracy of estimated forest area and its change are therefore critical to ensure the integrity of datasets besides its importance for national input in biological diversity and for reducing emissions from deforestation and forest degradation. Moreover, detailed forest cover data can enhance better national forest management strategies with respect to holistic approach of REDD+ forest monitoring programme.

However, the results are rarely perfect and forest cover change accuracy can be affected by a range of factors including definitions, data used, change of mapping process, methodological and integration of diverse social and biophysical processes (Olofsson et al. 2014; Rindfuss et al. 2004). Such difficulties can amplify uncertainties of where and when the forest cover change occurs, specifically generating special problems for land-cover change analysis. As described by Olofsson et al. (2014), an accuracy assessment can help to identify errors emanating from misclassification and as such the sample data can be used for estimating both accuracy and area together with the agreement between datasets, indicating usability of these estimates for national, regional and global forest monitoring (Fritz et al. 2011; Giri et al. 2005).

Research in this thesis has employed an assessment of accuracy of the two global datasets to the national context of REDD+ forest monitoring to bridge the national data gaps, particularly emphasizing on the case of Ethiopia. The thesis assesses the overall accuracy of the two datasets for five selected land cover classes and their spatial agreement that can allow characterization of uncertainties achieved from the datasets for adaptability to national forest monitoring.

1.2 Problem statement

Many developing countries like Ethiopia lack baseline reference data to determine greenhouse gas emissions from deforestation and forest degradation and climate change mitigation strategies. The key problem is lack of institutional investment and scientific capacity to undertake forest resource inventory. In such countries putting and producing an accurate forest inventories and information is also challenging. For example, FAO (2009b) indicated that global forest land use change results from remote sensing estimates showed large difference with that of FRA (2010), mostly in African countries due to use of old datasets. This can reduce the accuracy of forest cover change estimates reported by different organization or agencies. Similarly, differences can occur between existing global land cover maps and data sets in employing these for national scale forest monitoring.

Due to such limitations detailed information on the accuracy of datasets are needed, particularly for developing countries to study the impacts of forest cover changes on REDD+ activities. So, assessing the accuracy of remote sensing products regarding forest cover and change in support of national forest and land use policy options is important to bolster information generation (Olofsson et al. 2014). Quantifying accuracy of forest extent and change and understanding whether such change in forest cover occurred can enhance the capacity of countries in their report to achieve the goal of Kyoto protocol and the evolving activities on REDD+ through National Forest Monitoring System (NFMS). Moreover, assessing the accuracy of datasets provides information for the user and helps to decide whether it is acceptable or not depending on the acceptable level of errors.

Recently two global forest cover change studies (Hansen et al. 2013a; Potapov et al. 2011), have been undertaken using medium resolution Landsat satellites images to provide information on the dynamics of global forest cover extent and change. Both datasets provide the extent of forest and its change over the past decade on global basis. However, due to the limited availability of reliable forest inventory information and limited capacity to carry out such inventories in developing countries, promoting the use of such Remote Sensing (RS) products for national and local forest monitoring is important. In addition, the accuracy and complementarity of these products and their consistency while using as an important input to national REDD+ forest monitoring is not largely studied.

Therefore, the hypothesis of this thesis research is that both datasets namely FAO Remote Sensing Survey (RSS) (Potapov et al. 2011) and global forest change 2000–2012 (Hansen et al., 2013) are expected to achieve and improve the understanding of national forest extent and change in agreement. The results from analysis of these two dataset will provide information to government officials, land managers and researchers to generate better information with regard to forest monitoring decisions and activities.

The aim of this paper is to pinpoint adaptability of global remote sensing forest cover change results to the national level considering the case of Ethiopia, through assessing their accuracy and agreement in forest/land cover classification and changes observed. Besides, the study investigated the conformity of these independent data products in order to provide a better understanding of forest monitoring for REDD+ at national level.

1.3 Objective and research questions

The objective of the study is to assess the applicability of global forest cover change products and methods for adaptability to national scale forest monitoring activities for REDD+ in Ethiopia. To answer this main objective the following three research questions were investigated:

1. How national forest monitoring systems and measuring, reporting and verification requirements of REDD+ activities comply with global forest cover change monitoring products?
2. What forest cover change accuracies arise from global forest cover change products at national level?
3. Can global forest cover change estimates be adaptable to national REDD+ forest monitoring activities?

2. Materials and methods

2.1 Study area and sample locations

The study was conducted in Ethiopia, which is located in East Africa situated approximately between 3°23’53” and 14°53’41” N latitude and between 33°00’00” and 48°00’00” E longitude (Figure 1). Sample site location was selected based on FAO Remote Sensing Survey (RSS) tiles data product. However, due to less number of reference data for FRA RSS tiles additional samples were taken outside the tiles in order to assess the accuracy of for Hansen et al. (2013a) Global Forest Cover Change (GFCC).

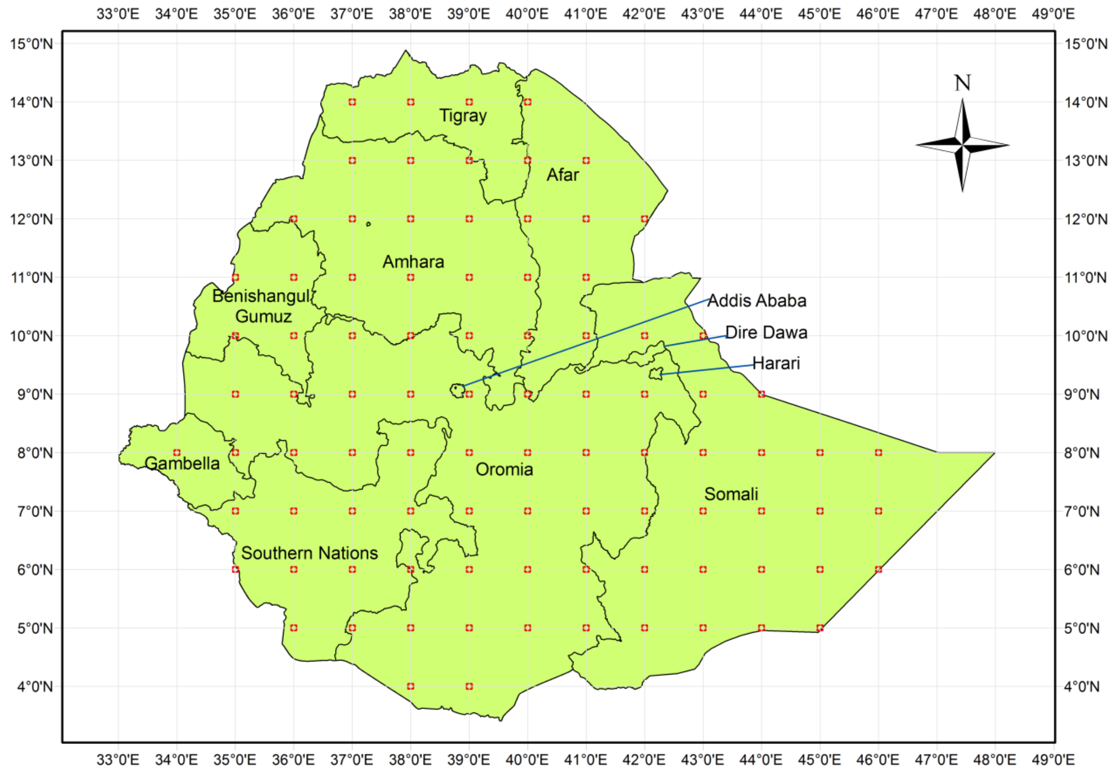


Figure 1. Study area and sample locations. The red polygons indicate that 10 x 10 km polygons at the intersection point of 1 x 1 degree. Samples for accuracy assessment of RSS and change agreement between Hansen et al. (2013a) and FRA RSS was selected from each tiles except for areas where historical google earth imagery was not available as a reference (Table 14 on Appendix I). Additional sample points were selected outside the 10 x10 km polygons only to assess the accuracy of Hansen et al. (2013) GFCC.

First all tiles were checked to determine if they have reference data based on FRA RSS (2010) grid based systematic sampling strategy. In total, only 51 sample tiles have reference data from google earth imagery according to FRA RSS tiles for the year between 2000 and 2006. Due to this problem accuracy assessment for FRA RSS and agreement between the two datasets was only based on samples located only in these 51 tiles.

2.2 Data sources and their methodological description

Two forest cover land use change datasets were used. The first one was a global RSS product by FAO (FAO 2009b; Potapov et al. 2011) and the second dataset was Hansen et al. (2013) high resolution global forest cover change (GFCC) map.

2.2.1 The FRA 2010 Remote sensing Survey

In order to complement national based Forest Resource Assessment (FRA) on forest cover extent and change, a global RSS has been conducted as part of United Nations Food and Agriculture Organization (FAO) and its partners namely the Joint Research Centre of the European Commission (JRC), USGS, NASA, South Dakota State University, Friedrich-Schiller University, and many national experts. The dataset encompasses global forest land use change from 1990 to 2005. It used multi date Landsat TM and ETM imagery to classify and analyse changes in forest and land use at 30 m x 30 m resolution for various ecoregions at global scale (Potapov et al. 2011). Employing automated supervised image classification and segmentation approach of each tile was classified in to land cover and then to land use polygons (FAO 2009b). The results are polygon layers containing forest cover and land use change information (i.e. detection of forest area, forest gains and forest losses) for the three time periods (1990, 2000 and 2005). The polygons are labelled by the following broad land use classes (forest, other wooded land, other land use, water) and no data to assess forest and land-use area change and the major drivers of these changes. The classification methodology of FRA RSS land cover/use is described in detail in (FAO 2009a, 2012).

At the national scale Ethiopia, FRA RSS product also has forest cover land use change polygons within 10 x10 km tiles at every intersection of 1^{o} by1^{o} (Figure 1 red polygons). Within the boundary of Ethiopia the dataset consists of 90 tiles with land use classes of forest, other woodland, other land use, water and no data. However, this study focuses only on two main classes namely (i) change classes-containing forest gain and forest loss and (ii) no change classes- consisting stable forest, other woodland and stable non forest) between the year 2000 and 2005.

2.2.2 High resolution global maps of 21st century forest cover change

The second dataset is from Hansen et al. (2013) high resolution GFCC map. Similarly, it was created by multi organizational team of researchers encompassing the first detailed map of GFCC from 2000 through 2012. This dataset provides a global quantification of forest extent, gain and trends in forest loss due to human and natural causes. According to the authors, the data product provide access to forest cover change information and change trend, particularly for countries which still have no the ability to view their trends in forest change. Therefore, this makes it possible to use the data provided for these countries in policy formulation and implementation i.e. to slow deforestation and forest degradation at national level.

In this study, the following three Hansen et al. GFCC products, processed from growing season Landsat 7 Enhanced Thematic Mapper (ETM) time series images (from 2000-2012) were used. These three datasets were downloaded from http://earthenginepartners.appspot.com/science-2013-global-forest/download.html on 11 May 2014 at 8:56 include: (i) tree canopy cover (TCC) for the year 2000. It is a canopy closure for all vegetation taller than 5m in height encompassing a range of 0–100% tree cover. This dataset was used as sampling and again as forest cover information of the country. (ii) Forest cover loss in time: annual forest cover loss product disaggregated to annual time scale of 2000 to 2005 in order to see the agreement between the two data sets. (iii) Global forest cover gain: indicates change from non-forest to forest for the period 2000–2012. This Hansen et al. GFCC data product has problem of disaggregation to annual time scale as well as to the study time (2000- 2005). For this purpose forest gain in this data product and the FRA RSS was checked for agreement based on visual assessment of reference data.

2.3 Data analysis methods

2.3.1 Sample point selection

Many researchers have presented options on the proper sampling scheme to be used based on the objectives of the research and include from simple random sampling to stratified sampling (Congalton 1991; Congalton and Green 2008; Olofsson et al. 2014). In order to assess the accuracy of each dataset and the agreement between them, two sampling strategy was adopted for the study in this thesis .The first one was systematic sampling strategy following the methodology of FRA RSS, where all tiles with available imagery were incorporated. The second one was simple random sampling to incorporate Hansen et al. GFCC dataset products outside the RSS tiles which help to assess accuracy of Hansen et al. dataset only.

Based on the first strategy, a total of 230 sample points were selected both for Hansen et al GFCC and FRA RSS, representing five major land cover classes namely forest loss, forest gain, stable forest, other woodland and stable non-forest cover areas. For the first two change classes, forest loss and forest gain 20 and 10 samples, respectively, were purposively selected if change was recorded in these datasets within the tiles. For the last three no change classes (stable forest, other woodland and stable non-forest areas), 200 sample points were sampled using simple random sampling strategy from Google Earth Imagery (GEI). This way the accuracy of RSS and the agreement between the two datasets were performed independently.

Second, due to complete coverage of Hansen et al. GFCC dataset additional simple random sampling method was preformed outside the tiles to see the accuracy of this dataset. In addition to 230 samples inside the tiles that can be used for both datasets, additional 170 samples were added outside the tiles for Hansen et al. GFCC dataset alone (i.e. 230 +170=400 samples for GFCC). From the total 400 sample points, 130 samples for forest loss, 60 for forest gain, 75 for stable forest, 35 for other woodland and 100 for stable non-forest classes were selected. In both cases, the determination of number of samples in each category was based on availability of information in both datasets and corresponding reference data.

2.3.2 Labelling selected sample points to forest or land cover classes

Different international and national organizations and researchers have produced different definitions of forest and other related land cover classes for different purposes including (i) for assessment of forest resources and its change, (ii) monitoring biological diversity in forests, or (iii) issues related to climate change. Moreover, plenty of land cover and land cover change definitions has been listed by Lund (2006). As a result there is difficulty to obtain comparable and consistent land cover classification results from different dataset. Similarly, the datasets used for this study have different land cover classes. The land cover categories generated by (FAO 2009b, 2012) was broadly in line with IPCC (2003) land cover land use change and forestry good practice guidelines. In Hansen et al. (2013) a generic land cover classification approach was employed, where percent tree cover, forest loss and forest gain are genereted. As a result the study undertake two procedures i.e harmonisation of forest or land cover change definitions employed by the two datasets and then assigning selected points to land cover classess.

First, following the methodology employed by FAO (2002a), we tried to harmonise the definitions and methodologies indicated in the two global datasets. This is in order to permit comparability and accuracy assessment of the datasets and thus determine the accuracy to fill the scarce availability of forest resource data at national level. The IPCC (2003) and UNFCCC (2001) definitions and guidelines were also used to harmonise them in line with MRV of REDD+ activities and NFMS. Given this situation, the author of this thesis has used definitions for key land cover classes that are widely recognised and internationally used, particularly those defined by (FAO 2006, 2009b). The complete terminology used is presented in Table 1.

Table 1. Harmonized definition of land cover classes used for assessment in this study

Broad category	Land cover /change classes	Definition
Change classes	Forest loss	Stand replacement or conversion of forest and other woodland class to non-forest covered class in the time period.
	Forest gain	change in an area from non-forest covered to forest in the study time
No change classes	Stable forest	Land covering more than 0.5 ha with trees higher than 5 m and a canopy cover of more than 10%. It does not include land that is predominantly under agricultural or urban land use.
	Other woodland	Land not classified as “Forest”, covering more than 0.5 ha; with tree height of more than 5 m and a canopy cover of 5-10%, excluding land largely under agricultural or urban land use. Or it is with a combined cover of shrubs, bushes and trees above 10 %.
	Stable non forest	All other land use classes with less than 5% tree cover including agricultural land, grass land, settlement, water, bare land and others which are not classified as “Forest” or “other wooded land”.

The definition for change classes (forest loss and forest gain) is common for both datasets. With regard to no change category, because of variation between the two data sets, standard definitions mostly employed by FAO are described in each class. Here, the threshold values indicated in Table 1 are taken from the forest definition of the of FAO global FRA (2006) and FAO remote sensing survey (2009b). This definition has been widely adopted for national forest inventories and for national forest monitoring systems (Magdon et al. 2014). Minimum tree cover which is the common parameter in both data sets was most commonly used during visual assessment supported by other defined elements.

Second, to assign each sample point to the land cover classes, historical data available on GEI was visually assessed spanning back and forward in time and then assigned to each class as a reference data for transparency. However, because of lack of access to GEI around the year 2000, land cover information available both in Hansen et al. GFCC percent tree cover dataset and on FRA RSS land use change polygons were used as forest cover information and assignment. The data from Hansen et.al 2000 tree cover data set was classified based on minimum tree cover definition for each no change classes (Figure 2A). Then it was overlaid with forest loss and forest gains as well with other datasets such as GEI and FRA RSS to check the accuracy and the agreement in labelling between the two data sets. Figure 2 and Table 15 indicate labelling procedures followed for each sample point selected, considering the land cover and land cover change definitions shown in Table 1.

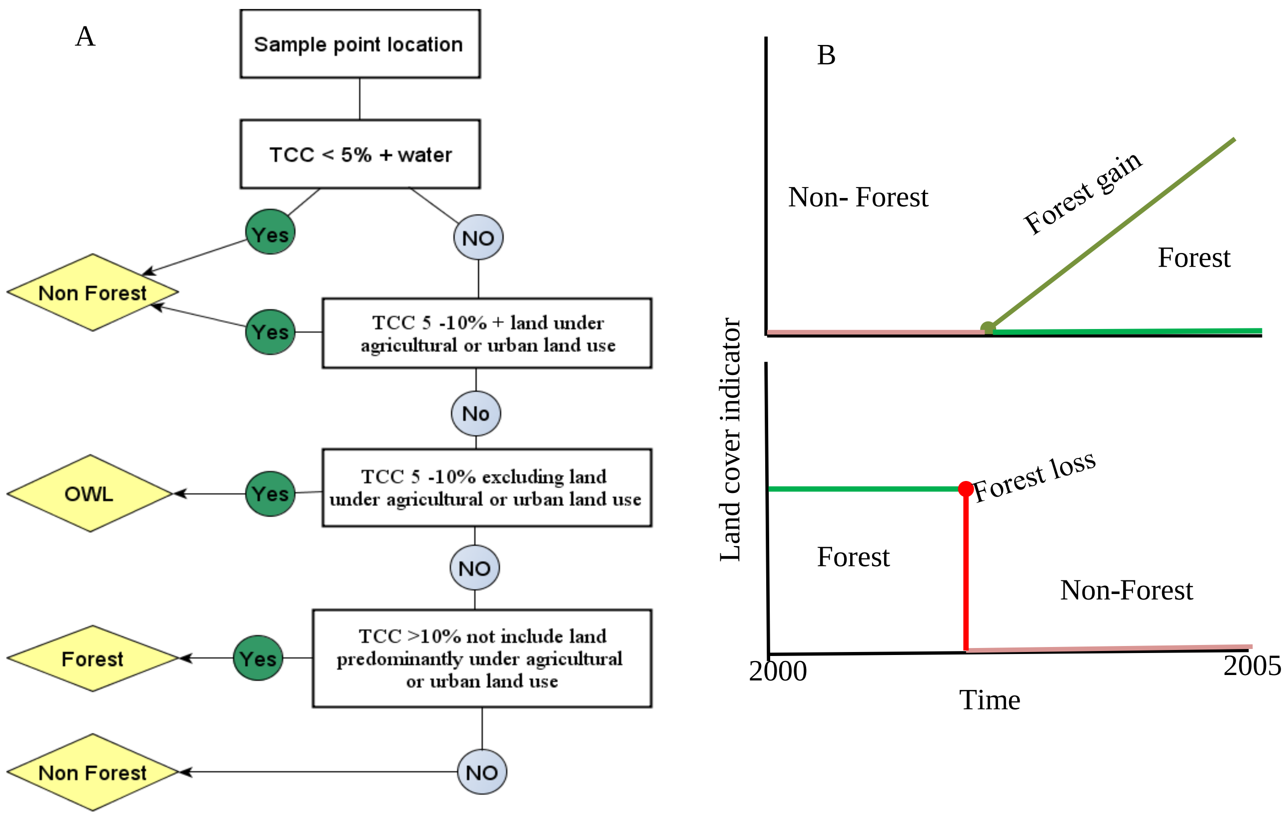


Figure 2. Conceptual frame work in assigning sample points to land cover and land cover change classes (Adapted from Magdon et al. (2014) and Watson et al. (2000))

Figure 2A indicate decision tree of no change classes in Hansen et.al GFCC. Figure 2B is land cover change assignment methods for change classes (forest loss and forest gain). Land cover indicators such as tree canopy cover and minimum area were assessed between the two points in time and labelled as one of the change classes if change in forest cover is observed. For example, the upper figure in 2B indicate change from non-forest cover to forest showing forest gain and the lower figure is inverse of the upper indicating forest loss. Since there are labelled land use land cover change polygons in RSS dataset, the information available in the dataset which overlay with the sample point is directly used for assessment.

The other important point to mention was the use of reference data for assessment of change or no change classes. Reference data closest to the target study period was used for assessment of both datasets agreement and FRA RSS accuracy analysis, particularly from 2006 and 2007 based on the methodology described by FAO (2012). This is to check the consistency of change or no change classes. The time coverage of Hansen et al. dataset is between 2000 and 2012. So, in order know the permanency of land cover classes observed in or around our study period, sample points for Hansen et al. GFCC accuracy assessment alone were checked until 2012.

Furthermore, if there was no change of forest to non-forest and vice versa between the two points in time, the sample site under assessment is labelled under one of the no change category as: (i) stable forest, if forest cover is permanent in GEI or other two datasets in agreement; (ii) stable non forest, if no forest covered is observed on GEI, recorded percent tree cover was less than 5% in 2000 and no change was observed until 2005 or later; or (iii) stable other woodland, if the recorded data indicate 5-10% tree cover in time (Figure 2). Areas not covered with more than 5% tree cover (e.g. land predominantly covered with agricultural fields, bare lands, roads and other infrastructures) without loss or gain of tree or forest cover was regarded as stable non-forest for this particular study. On the other hand, if change of forest to non-forest and Vice versa in time series is observed it was labelled under one of change category as (i) forest loss- if there was change of forest/other woodland covered area to non-forest class or open woodland in the time period., and (ii) forest gain, if change from non-forest class to forest or closed non-agricultural wood land covered is observed (Figure 2; Table 15).

2.3.3 Analysis of global forest cover change products in the context of national Forest monitoring system and MRV requirements of REDD+ activities

Maniatis and Mollicone (2010) indicated that developing countries willing to participate in the mitigation mechanism of REDD and to enhance the role of conservation and forest management activities needs to establish national forest monitoring system (NFMS). Guided with viewpoint, this section was tried to bring overall guidance on each datasets definition and how it perform related to Measuring, Reporting and Verification (MRV) criteria for NFMS provided by the Intergovernmental Panel on Climate Change (IPCC). The MRV for REDD+ specifically refers to measuring (which help to obtain information referred to as activity data), reporting (to collect relevant forest and forest change data for national estimate) and verification (to check the accuracy and reliability of reported information) of a country's forest, including its change over time (www.fcmcglobal.org). Overall, the result of this information helps to analyze (i) accuracy of the data sets and their agreement on forest cover or change (research question 2), (ii) the reliability of information generated and whether data comply with defined quality criteria such as transparency, comparability, consistency and completeness (research question 3). Thus, this aims to set bench mark and guidance on collating requirements which are relevant to NFMS and MRV of REDD+ activities from each datasets. It assessed the practical implementation of each data set to the Ethiopia's NFMS context by relying on secondary data sources.

2.3.4 Forest cover change accuracy and spatial agreement analysis

2.3.4.1 Accuracy assessment

In this specific research question quantitative accuracy assessment was undertaken based upon the collection of reference data which is used to compare the global forest cover change products in question. Although, there is some discrepancies in the availability of high resolution imagery the reference data used (i.e. GEI) is assumed to be correct. Most importantly, as indicated in Steven and Brown (2012) the accuracy assessment offers a validation of these products, providing an indication for reliability of forest and land cover classification results. This is important to help REDD+ managers and users of the products at country level like in Ethiopia and to be used throughout the decision making process.

First interpretations and class labelling for each sample point from reference data sources (GEI and its supplementary products like global forest watch) and the two datasets under question were carried out. Then comparisons of these reference data to the mapped datasets were organized and presented in an error matrix based on the methodology described by different studies (Boschetti et al. 2004; Congalton and Green 2008; Hollister et al. 2004). An error matrix is a square array of numbers organized in rows and columns which express the number of sample units (i.e. pixels, polygons or points) assigned to a particular land cover category relative to the actual category as indicated by reference data (Congalton and Green 2008) as shown in Table 2. The values in the columns represent the reference samples from high resolution GEI, while the rows represent values of the map for each global forest cover change dataset corresponding to the five classes.

Overall, the tabulated values across the diagonal of each matrix describe how much times the classified datasets and reference data samples have equal values of correct classifications. On the other hand, the off-diagonal values in the matrix indicate errors due to either of inclusion or exclusion to/from each particular class. Commission error is shown on the horizontal axis of classes, while omission errors are presented on the vertical axis. The accuracy of the outcome of each dataset in the error matrix were estimated using the following three techniques:

1. producer's accuracy - indicate the probability of a certain land cover of an area on GEI is being the same as classified forest or land cover change classes in global datasets (Hansen et. al. GFCC or FRA RSS). The error introduced is due to excluding an area from a class to which it does truly belong (false negative or underestimation).
2. User’s accuracy - indicates the probability of change classes assessed for accuracy in global datasets being the same as in GEI. The error introduced is due to including an area to the category which it does not truly belong (false positive or overestimation).
3. Overall accuracy - which summarizes the total agreement/disagreement between the two datasets (classified global datasets against GEI). It incorporates only the diagonal values excluding the omission and commission errors. All techniques were calculated following similar equations which were described and implemented previously in numerous studies (Boschetti et al. 2004; Foody 2002; Olofsson et al. 2014). The illustration of each dataset accuracy assessment was presented with confusion matrix (Table 8 and Table 9), by calculating the level of accuracy and agreement in both datasets.

Table 2. A typical example of confusion matrix used

Classified dataset	Reference data (GEI)		
	Class 1	Class 2	Class 3
Class 1	C11	C12	C13
Class 2	C21	C22	C23
Class 3	C31	C32	C33

$$producers\ accuracy\ (pa) = \frac{Cii}{\mathrm{C+i}} * 100 \text{ ....Equation (1)}$$

$$Users\ accuracy\ (Ua) = \frac{Cii}{Ci+} * 100 \text{ .... Equation (2)}$$

$$Overall\ accuracy\ (Oa) = \frac{\sum_{k=1}^{r} Cii}{N} * 100 \text{ .... Equation (3)}$$

Where C_{ii} represents total number of correct cells (samples) in a class, C+i indicate sum of cell values in a column, Ci+ represents sum of cell values in a row, and N is total number of samples in error matrix.

Besides, as indicated by Boschetti et al. (2004) omission error (Oe) and commission error (Ce) which are closely linked to producer and user accuracy by the relations of 1- Pa and 1- Ua, respectively were used to assess inclusion and exclusion of classes. Lastly, although both producer and user accuracy measures are important, we mostly examined the result using user's accuracy, because it intuitively represents how much the user should ''believe'' that a given sample from actual reference data is the class that the map from each data set claims it is (Langford et al. 2006).

2.3.4.1 Determination of spatial agreement between the two datasets

Determination of spatial agreement between the two datasets in forest change was also be carried out based on the concept described by Rindfuss et al. (2004) and by Fritz et al. (2011). Spatial agreement is concerned with similarity of land cover classification in to one category or similarity of occurrence of forest cover change (loss or gain) in space. For example, if change of forest cover in space between the datasets is similar, positive spatial agreement exist. Conversely, if changes in forest cover between the datasets in space is dissimilar in their change attributes, it indicates the two datasets have spatial disagreement (Rindfuss et al. 2004). This way, the agreement of forest cover and change estimates and experienced by an individual sample point (or polygon) were evaluated. The percent agreement were calculated using equation 4 as adopted from Giri et al. (2005).

$$Overall\ percent\ agreement = \frac{\sum_{k=1}^{r} Cii}{N} * 100 \text{ .... Equation (4)}$$

Where C_{ii} represents total number of samples that have spatial agreement between the two data sets and N is total number of samples in the matrix. Later on disagreement areas were cross-checked with GEI for mapping. Overall, the accuracy assessment between each mapped dataset and GEI (section 2.3.4.1) and agreement between the two datasets (this section) were done to address the questions related to how well locations of mapped changes in the two data product correspond to actual areas of change on the ground.

2.3.5 Analysis of adaptability of global forest cover change data sets to national REDD+ forest monitoring

The REDD+ forest monitoring approach involves the collection of data and information for the estimation of emissions and removals of GHG related to forest cover change. Two basic data inputs were used for GHG inventory namely activity data (AD) and emission factors (EF) that are usually come from national forest inventory (NFI) utilising RS technologies combined with ground measurements. This specific research question was more restricted to activity data, because it is more related to the two global land cover change datasets that are derived from RS data. Activity data provide area of deforestation forest degradation, forest conservation, sustainable management of forests and enhancement of forest carbon stocks which are main elements of REDD+ activities. This data is important for setting up forest reference emission level (REL) or reference level (RL).

However, Ethiopia lacks capacity to produce activity data for forest monitoring (Romijn et al. 2012). In such case, there is a need to complement existing data sets to the national forest monitoring and MRV framework considering the 5 reporting principles of IPCC good practice guidance (GPG) into consideration. This reporting requirements or principles also apply to REDD+ (Bernard and Minang 2011) to guide an independent assessment of the datasets understudy. Hence, the latest convention decision 4/CP.15 on REDD+ requests parties to "provide estimates that are transparent, consistent, as far as possible accurate, and that reduce uncertainties, taking into account national capabilities and capacities" as described in Achard et al. (2012).

Moreover, besides the accuracy assessment described, the two datasets were independently checked to verify the reliability of information provided or the procedures used to generate information for REDD+ national forest monitoring. Thus, adaptability level of each dataset for national REDD+ forest monitoring was assessed using the following indicators (parameters) as described in FAO (2013). The indicators described in this report were consistency, completeness, transparency, comparability and accuracy. These principles help to outline the extent of the datasets adaptability and comply with national REDD+ forest monitoring. The description of the five reporting principles used to assess the adaptability level of each data set to national REDD+ is presented in Table 3.

In order to analyse the data sets adaptability, first specific guiding parameters were derived for the 5 main reporting principles of IPCC described in FAO (2013). This is to develop an attribute value to each data set characteristics and frame the answers to be responded in each data set. Second, adopting the methodology described by Romijn et al. (2012) indicators and values were assigned to each specific parameter based on (i) description available in each data set (Section 2.2), and (ii) following the results obtained (Section 3.2 and 3.3). Different parameters receive different indicators and values depending

Table 3. Good practices guidance for REDD+ forest monitoring based on IPCC reporting principles

Guiding principles	Description
Consistency:	Capability of representing land cover categories consistently over time, without affection by discontinuities in time series. Consistent definitions and methodologies should be used along time and should aim to reflect the real annual differences in land cover change and carbon emissions, but not be subject to changes resulting from methodological differences. All relevant information is properly documented to the extent that all land cover changes and carbon pools can be tracked in the future.
Comparability	The data, methods and assumptions applied in the LCLUC accounting process must be those with widespread consensus and which allow producing meaningful and valid comparisons between areas, data sets and change classes. Other parties should follow the methodologies and standard formats to estimate and report forest inventories.
Transparency:	Data sources, definitions, methodologies and assumptions used in forest inventory are clearly described and appropriately documented. The integrity of the reported results able to be confirmed by a third party, so that anybody could verify its correctness.
Completeness:	All land within a country should be included with change in forest cover extent and land use change in order to report emissions and carbon removals of GHGs. Ability to solve lack of relevant data for measuring and monitoring forest area change.
Accuracy:	How land cover land cover change categories are closely estimated to the true value and the conversion between land cover categories match the existing historical data to estimate carbon stock changes and GHG emissions with minimum error. So far data set has to use appropriate methodologies to promote accuracy in forest monitoring mechanism and to improve future inventories as much as possible.

Sources: (IPCC 2003, 2006), Bernard and Minang (2011), Achard et al. (2012) and FAO (2013)

on the availability of information in each dataset corresponding to the specific parameter identified i.e. high=1, medium = 0.5, low = 0. Adaptability level of each data set was calculated using equation 5, that was developed by the author of this thesis following the methodology described by Romijn et al. (2012). This way the forest monitoring capacity of each data and its complementarity to NFMS and MRV principles was analysed in a transparent and consistent way.

$$AL = \frac{(\sum PV\ high*1)+(\sum PV\ medium*0.5)+(\sum PV\ low * 0)}{N} * 100 \text{ .... Equation (5)}$$

Where AL is the adaptability level, referring each datasets compatibility to national REDD+ forest monitoring, PV is parameter value and N is total number of specific parameters based on 5 reporting principles.

3. Results

This chapter presents the main findings of the study. The analysis of the three research questions formulated to address the issue of applicability of global forest cover and change datasets to national REDD+ forest monitoring systems are presented in consecutive sections. The first two sections presents the result of literature and documentary sources, where section 3.1 presents the analysis on national forest monitoring system and measuring, reporting and verification requirements and section 3.2 analysis complementarity of global forest cover change data sources and their definition to the stated requirements at national scale Ethiopia. In section 3.3 accuracy analysis and spatial agreement of the two datasets under study was presented. Under section 3.4, analysis of adaptability level of each dataset to the national REDD+ forest monitoring approach based on the five reporting principles of IPCC was presented.

3.1 Measuring, reporting and verification requirements and the national forest monitoring systems of REDD+ activities in Ethiopia

3.1.1 Measuring, reporting and verification for REDD+

The conference of parties on climate change held in Cancun, Mexico, in 2010 adopted a decision 1/CP.16, which marked a remarkable turning point in development of REDD+. The agreement states that parties have to collectively aim to halt, slow and reverse carbon emissions from deforestation and forest degradation, and thereby address the 5 main activities of REDD+ (reducing emissions from deforestation, reducing emissions from forest degradation, conservation of forest carbon stocks, sustainable management of forests, and enhancement of forest carbon stocks). The measuring, reporting and verification (MRV) is included in this agreements as one of the most critical elements necessary for the successful implementation of any REDD+ mechanism (Bernard and Minang 2011). Thus, enabling the assessment of national GHG emissions and removals in the forestry sector and reporting this to the UNFCCC in a transparent, accountable and verifiable manner is required. It is a monitoring action responding to what each country defines as necessary to ensure that its REDD+ actions are result based (Vietnam-redd.org 2011).

As quoted by Bernard and Minang (2011) on strengthening MRV for REDD+ and also defined by the United Nations REDD programme (UN-REDD 2009) which supports countries to develop robust and compatible national forest monitoring and MRV system, the key terms of MRV are defined as:

- Measurement - the process of data collection over time, providing basic datasets, including associated accuracy and precision, for the range of relevant variables. Possible data sources are field observations and measurements, detection through remote sensing and interviews.
- Reporting - implies the process of formal reporting of assessment results to the UNFCCC, according to predetermined formats and established standards, especially following the IPCC guidelines and good practice guidance (GPG).

- Verification - refers to the process of independently checking the accuracy and reliability of reported information (e.g. national forest inventory reports to the UNFCC).

Accordingly, Ethiopia has envisages participating in the implementation of REDD+ and has started setting up her MRV system for the determination of carbon benefits. The readiness proposal (R-PP) for REDD+ implementation, to measure forest area change and estimate emission factors through a combination of RS and NFI approaches has been prepared. However, this requires consistent forest monitoring approach, where Ethiopia lacks capacities to develop and implement a NFI and monitoring system for REDD+.

3.1.2 The national forest monitoring system

Several strategies were developed to assist developing countries in the establishment of REDD+ MRV systems. One important mechanism is through the establishment of national forest monitoring system (NFMS). The national forest monitoring activities intern provides technical services on MRV of forest related REDD+ activities. The 15th conference of the parties (COP15) to the UNFCC in 2009 (Copenhagen, Denmark) has adopted the methodological guidance decision 4/CP.15 for REDD+ activities. This decision requests developing country parties to establish robust, consistent and transparent national forest monitoring system that:

- Use a combination of RS and ground-based forest inventory approaches for estimating forest area change, forest related GHG emissions by sources and removals by sinks
- Provide estimates that are transparent, consistent, as far as possible accurate, and that reduce uncertainties, taking into account national capabilities and capacities
- Are transparent and their results are available and suitable for review

Accordingly, Ethiopia is one of the REDD+ participating country that has included the forest sector as a major focus area in its sustainable development strategy and plans to improve management of forest resources (www.norway.org.et 2012). Its national forest inventory (NFI) is designed in order to align with sustainable development strategy and REDD+ MRV needs. Three pillars of NFMS that are essential to undertake its REDD+ MRV functions (MoA 2013) are also designed. The 3 pillars include (i) the satellite land monitoring system (SLMS) which helps to collect and assess the activity data (AD) over time related to forest land, (ii) the national forest inventory (NFI) - to collect information on forest carbon stocks and changes, relevant for estimating emissions and removals and to provide emissions factors, and (ii) the national GHG inventory - as a tool for reporting on anthropogenic forest-related GHG emissions by sources and removals by sinks (FAO 2013; MoA 2013). The first two pillars were used to provide inputs into the third pillar, while they are operationalized in three phases. In this regard, RS can be used as an input both for yielding AD for MRV as well as for meeting broader forest monitoring needs.

However, information obtained from different sources including Herold (2009) and Romijn et al. (2012) both on assessment of national forest monitoring capacities in tropical non-Annex I countries based on

FAO FRA reports, Bernard and Minang (2011) on strengthening MRV for REDD and (www.norway.org.et 2012) on forest carbon monitoring capacity assessment indicated that, Ethiopia still lacks:

- complete and regular forest cover change monitoring capacity,
- standardized and consistent method of national and regional forest data collection and analysis,
- coordinated system for managing and integrating forest-related data, and
- inconsistency between regional, national and global forest monitoring strategy for REDD+.

The establishment of the UN-REDD+ program in Ethiopia is in response to real world challenges that developing country are facing. Developing consistent MRV system for REDD+ has the opportunity to identify Ethiopia's forest monitoring activities and actions associated with it. This can improve current issues related to (i) inadequate baseline forestry datasets needed to measure reference emission levels, (ii) inadequate government capacity and experiences of implementing REDD+ type measures at national, regional and local level forest monitoring systems without conflict.

3.2 Complementarity of global and national forest cover change monitoring systems with MRV requirements of REDD+ activities

3.2.1 Forest or land cover change definitions

The definition of forest and its change is fundamental in land cover classification systems, where classification is an abstract representation of what is actually exist in the field using well defined criteria (Di Gregorio 2005). There was an experts meeting organized by FAO (2002a) which discussed on the issue of harmonizing forest related definitions for use by various stakeholders. As described in the report from the meeting, land cover classification have to rely on the combination of a set of independent attributes allowing the user to define a wide variety of different land cover features within a standardized framework. The clear definition of forest with defined attribute is important, particularly related to REDD+ forest monitoring mechanisms.

From the change and no change categories indicated in Table 1, most importantly the options based on definition for the term forest and deforestation are presented and analyzed below. National scale and Hansen et.al definitional scenarios are also included in order to facilitate the choice for NFMS of Ethiopia.

3.2.1.1 Forest definitions

Various definitions of the term forest have been discussed during the first and second expert meeting, which was jointly organised by FAO and IPCC in Rome, in 2002. However, still there is difficulty to find a comprehensively applicable definition of forest, although definition influences REDD+ activities. Such variation in definition of forest and other related land cover change may be related to the terms used in different disciplines and the purpose of research activities. Among many, comparison of 4 global and 1 national level established forest definitions are presented in Table 4.

Table 4. Definitions of forest by source

FAO (2010a): Land spanning more than 0.5 ha with trees higher than 5 m and a 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.

UNFCCC (2001): 'Forest' is a minimum area of land of 0.05-1.0 ha with tree crown cover (or equivalent stocking level) of more than 10-30% with trees with the potential to reach a minimum height of 2-5 m at maturity in situ. A forest may consist either of closed forest formations where trees of various storeys and undergrowth cover a high proportion of the ground or open forest. Young natural stands and all plantations which have yet to reach a crown density of 10-30% or tree height of 2-5 m are included under forest, as are areas normally forming part of the forest area which are temporarily unstocked as a result of human intervention such as harvesting or natural causes but which are expected to revert to forest.

UNEP/CBD (2001): Forest is a land area of more than 0.5 ha, with a tree canopy cover of more than 10%, which is not primarily under agricultural or other specific non-forest land use. In the case of young forests or regions where tree growth is climatically suppressed, the trees should be capable of reaching a height of 5 m in situ, and of meeting the canopy cover requirement.

Hansen et al. (2013a): Defined in terms of tree cover where all vegetation taller than 5m in height at Landsat pixel scale is included.

WBISPP (2004): Land with relatively continuous cover of trees, which are evergreen or semi deciduous, only being leafless for a short period, and then not simultaneously for all species. The canopy should preferably have more than one story.

Moreover, the IPCC special report by (IPCC 2000) showed 3 ways that some authors define forest. The 1st one is in terms of legal or administrative requirements which have no any relationship to the vegetation characteristics and associated carbon on the land. Others define it in terms of land use implying particular purpose for which the forest land is used (FAO 2012). For instance, wood production and not used to a significant extent for other purposes. The above two cases according to IPCC may have a confounding information on forest cover and estimation of related carbon stock. The 3rd category defines forest in terms of vegetative land cover (biophysical attributes) characterizing it by relative tree cover (FAO 2012; Hansen et al. 2013a; IPCC 2000). The 3rd category specifies variation in forest cover using area cover, minimum canopy cover, minimum height, and minimum biomass. Some attributes (parameters) that are used to quantify forest in different data sources are presented in the Table 5.

Table 5. Attributes and thresholds of 'forest' used in different data sources.

Parameters		Data source				
		FAO	UNFCCC	(UNEP/CBD)	Hansen et al.	(WBISPP)
Threshold	Min. area (ha)	0.5	0.05 - 1	0.5	0.09 ha**	n/a
	Min. height (m)	5	2-5	5	5	n/a
	Min. crown cover (%)	10	10 - 30	10	0-100	>20
	Strip width (m)	20	n/a	n/a	n/a	n/a
Binary	Temporally unstocked areas	yes	yes	no	no	no
	Forestry land use	yes	no	yes	no	yes
	Young stands	yes	yes	yes	no	yes

Adapted from Schoene et al. (2007) and FAO (2002b)

The key parameters and features of various forest definitions quoted in Table 4 are presented schematically in Table 5, where in the binary section "yes" indicate the presence of the parameter in the definition of forest corresponding to the data source, while "no" represents the absence. In the threshold section "n/a" indicates uncertain because of data not available while reviewing. ** is single pixel value.

The major difference observed from definitional scenarios of forest (Table 5) is variation in usage of quantitative thresholds to define forest (i.e. minimum height, area and crown cover). Although, almost all of them included these threshold parameters for classification and identification of changes in forest cover and its extent, due to variation of the purpose, the parameters corresponding to the data source are defined differently. The FRA definition has complete set of parameters. The Marrakech Accord (MA) of the UNFCCC (2001) mainly concerned with carbon accounting and defines all areas containing trees within the defined parameter as forest. The convention on biological diversity (CBD) which has a concern more on biodiversity issues seems to follow the FAO FRA forest definition with the exception that there is variation in strip width and temporarily unstocked areas consideration. Similarly, except in Hansen et al. GFCC young trees that do not yet meet the minimum height or crown cover criteria, but provided that they are expected to reach these parameter thresholds as maturity is included in all forest definitions.

The national definition, which is the base for FRA forest resource assessment do not explicitly comply in most parameters. This can make difficulty to transpose national data to global scale, and in turn might raise a question on the accuracy of forest extent and its change that has been reported to the FAO. For example, contrary to global FRA and CBD, the national dataset WBISPP (2004) classified all other woodlands and shrub land that has less than 5 m in height with tree canopy cover (TCC) >20% as low woodland cover. Therefore, this national scale definition needs complete set of parameters and values in order to complement with global data sets.

Moreover, a minimum tree cover (crown cover) is the most commonly used parameter in all data sources, to quantitatively assess forest and non-forest areas in land cover land use based definitions (Magdon et al. 2014) Table 5. In definition implemented by Hansen et al. (2013a) land cover with percent tree cover from 0-100 was implemented for all vegetation taller than 5 m in height. However, because of Hansen et al. GFCC data set is freely available online to download, it is easy to disaggregate percent TCC and adaptable to other thresholds for forest extent and change analysis.

3.2.1.1 Forest loss (deforestation) definition

Forest loss (deforestation) is the conversion of forest to another non-forest land use or the long-term reduction of the tree canopy cover below the minimum 10% threshold of FAO (2010a), or the direct human-induced conversion of forested land to non-forested land (UNFCCC, 2001). Besides, Hansen et al. defined forest loss as change from a forest to non-forest state through complete removal of tree canopy cover at the Landsat pixel scale. It includes areas of forest converted to agriculture, pasture, water reservoirs and urban areas. On the other hand, forest gain is the other way definition of forest loss in many data sources, referring to the establishment of trees on non-treed (forest) land due to afforestation, reforestation or natural regeneration. Hansen et al. GFCC definition indicate that forest gain emphasize a change in land-cover from non-forest to forest state (i.e. from ~ 0% TCC to > 50% TCC densities). Attribute and threshold differences between major existing definitions of deforestation from the FAO FRA**,** UNFCCC (2001), Hansen et al. (2013a) and the national WBISPP are shown in Table 6.

Table 6. Attributes and thresholds of forest loss or deforestation used by different data sources

Parameters	**Data source**			
	FAO (2010a)	UNFCCC	Hansen et al.	WBISPP (2004)
Crown cover change	< 10%	< 10-30 %	~ 0%	n/a
Transition from forest to non-forest	yes	yes	yes	yes
Land-use change	yes	no	no	yes
Only directly human-induced	no	yes	no	no
Does not consider temporary loss (non-stocked condition)	yes	yes	no	yes

Adapted from Schoene et al. (2007)

Table 6 is schematically presentation of deforestation definition parameters considered in four different sources. All except the first row are binary parameters where, “yes” represents presence of the parameter in the definition of deforestation, while the “no” shows the absence and “n/a” represents data not accessed while reviewing.

The major deforestation definitions applied globally and nationally imply both similarity and differences. With regard to their similarity, (i) all deforestation definition consist transition from forest state to non-forest, and (ii) except Hansen et al. GFCC temporarily unstocked conditions as a result of harvesting is not considered as forest loss due to it may not lead to a non-forest state. Two major differences were also observed in deforestation definition. These include:

- Differences in the threshold of minimum crown cover change. For example, according to Hansen et.al data set there have to be a change from more than 50% TCC to ~ 0% TCC, while it can be a reduction to the state of less than 10 to 30% TCC change according to MA.
- The term conversion of forest to other land use is not considered in Hansen et.al and MA. Causes of deforestation is also ambiguous where the MA considered only human induced conversion of forested land, while the other three data sources in the Table 6 (FAO, Hansen et al and WBISPP) considered natural perturbations in addition.

Overall, the definition of forest is a key factor for setting reference scenario and MRV system of REDD+. However, from Table 4,Table 5 andTable 6 it can be generalized that variation in definition of forest and forest loss, as well their associated land cover classes can lead to ambiguous interpretations from global to national scale. Without clear and standard definitions it is difficult to measure the progress or pay for performance, which is central to the REDD+ discussions. The scope of forest cover and forest cover change interpretation is also critical issue for REDD+ implementation, since it determine what is being measured, reported and verified. So, lack of common and clear understanding of what to be constituted as forest and other land cover classes can affect MRV system of REDD+ activities. This is specifically immense in Ethiopia, where no standard forest definition has been publicized to ensure consistency in the estimation of forest related emissions. According to IPCC (2000) the use of such different definitions can account to different estimates of forest extent and change in carbon stocks and MRV of REDD+ forest monitoring mechanisms as shown under following section.

3.2.2 Forest and forest loss extent estimation in Ethiopia

The information on the extent of forest and forest loss is important and necessary for forest resource assessment and change monitoring. It is also important to establish links between REDD+ forest monitoring actions at national and regional levels.

3.2.2.1 Variation in forest extent estimation (global, national and regional)

The comparison of forest extent and forest cover loss indicated in Table 7 showed variation between datasets for the same period. As an illustration, based on national definitions of the Woody Biomass Inventory and Strategic Planning Project (WBISPP 2004) by ministry of agriculture and rural development of Ethiopia the estimated total area under forest for the country was 3.65 million ha (about 3.3% of total land area) for the year 2000. Five years later the area under forest was 3.33 million ha (about 3% of total land area). On the other hand, based on FAO global definitional scenario the estimate indicate 13.7 million ha by 2000 and 13 million by 2005 which amounts 12.4 and 11.8% of the total land area (FAO 2006, 2010a, b), respectively. According to these figures there was a great pick of forest cover based on FAO FRA estimate than national based WBISPP estimate. However, it is very important to mention that forest cover assessment was based on WBISPP 2004 national data using reclassification, calibration and extrapolation method.

Table 7. Forest cover extent for the year 2000 and forest loss between 2000-2005 at national and regional states of Ethiopia. The *** is forest loss calculated based on forest cover difference observed between 2000 and 2005 from WBISPP report as described by FAO (2010b), while ** is country level forest cover extent in percent from country total land area, and * is country level forest loss. Forest cover calculated for Hansen et.al. GFCC was based on more than 20% tree cover extent and forest loss was disaggregated to 2000-2005 from 2000-2012. Regional share of percent forest cover and percent forest loss is from country level total.

Region	National estimate for year 2000 in (FAO 2010b)				Hansen et al. GFCC			
	Forest cover		Forest loss***		Forest cover		Forest loss	
	Area (ha)	%	Area (ha)	%	Area (ha)	%	Area (ha)	%
Oromia	2205619	60.40	173607	55.3	6962651	42	45114	51.6
SNNPR	740271	20.27	101844	32.44	4648175	28	26832	30.7
Gambella	491805	13.47	30219	9.63	1408279	8.05	2862	3.3
Harari	216	0.01	0	0	3.33	0	4	0
Amhara	92744	2.54	8278	2.64	427790	2.58	1432	1.6
Tigray	9332	0.26	0	0	16630	0.1	170	0.2
Beneshangul	68495	1.88	0	0	3092471	18.67	10457	12
Afar	39197	1.07	0	0	5425	0.03	20	0
Somali	4257	0.12	0	0	2784	0.02	608	0.7
Regional total	**3651936**	**100**	**313948**	**100**	**16563209**	**100**	**87497**	**100**
Country total	**3651936**	**3.2****	**313948**	**8.6***	**16563209**	**14.6****	**87497**	**0.53***

Besides, FAO on FRA 2000 (FAO 2010a, b) estimated 13,705,000 ha forest cove (12.4% of total land area) in the year 2000 and 13,000,000 ha by the year 2005. The estimated percent tree cover with more than 20% threshold (to harmonies with national definition) from Hansen et al. GFCC for the whole country indicate that there was 16,563,209 ha tree cover which accounts 14.6% of total land area. The size of forest cover is relatively higher than FRA estimate for the country and has no comparative figure with the national estimate of 3.2% of total land area. These taken together indicate that there is difficulty to get a reliable estimate on forest cover and forest cover change of Ethiopia.

The most forest covered regions of Ethiopia lies in south and south-western part of the country, predominantly in Oromia and Southern Nations and Nationalities. Similarly, variation in forest extent is reported between data sources for regional states. As an example, Oromia regional state has about 7 million ha forest cover (42% of national forest covered area considering Hansen et al. GFCC estimation) and 2.2 million ha (60.4% of the national total based on WBISPP national forest cover extent estimation). The second largest forest covered region was Sothern nations, covering 20.3% according to national report and 28% according to Hansen et al. GFCC. On the other hand, in both data sources it was reported that Harari, Afar, Somali and Tigray regions have very low forest cover compared to other regions (Table 7).

3.2.2.2 Forest loss estimation in Ethiopia (global, national and regional)

There was similar variation with regard to forest loss estimation of the country between different datasets, making difficulty to provide consistent and reliable deforestation rate. The FAO (2010a) estimates indicate that 141,000 ha of forest have been lost annually between 2000 and 2005 and the average annual deforestation rate for this specific period was about 1.05% of the total forest cover in 2000. Deforestation rate between the time periods was more than 5.1% of year 2000 forest area. In the same period (2000-2005) from Hansen et.al GFCC, the total area of forest loss estimated was 87,731 ha (0.53% out of >20% total tree cover in 2000), which is much lower than FAO FRA estimate that accounted for >10% canopy cover in forest cover estimation. Annual forest cover loss of Hansen et.al GFCC (0.12%) is still much lower compared to FAO (2006) and WBISPP (2004) national estimation (1.7%).

Given the diversity of regional contexts within Ethiopia, forest change occurred with different rates in different parts of the country. Based on the forest loss assessment undertaken for the country between 2000-2005, both data sources have no common change rate. However, in both cases forest loss is concentrated in the southern and south-western forested areas of the country. The most forest covered regions, Oromia and Southern nations experienced the highest forest loss, accounting 52% (45,114 ha) and 31% (26,832 ha) share of national forest loss respectively, as computed from Hansen et.al yearly forest loss dataset. Similarly, WBISPP 2005 in Moges et al. (2010) computed 1.6% forest loss for the three forested regional states (Oromia, SNNP, Gambella), although Gambella is the 4th regional state according to Hansen et.al GFCC forest loss dataset. The other regional states with low tree cover also experienced low deforestation rates.

In conclusion of section 3.2 Ethiopia has inconsistent national forest inventories, and still has incomplete and out of date forest monitoring system, as of other few tropical countries (FAO 2009). All global data assessed under this section have varied estimates and definitions on forest resources of Ethiopia. Most of these data sources, particularly the global once (Hansen et. al GFCC and FAO FRA) has intended to complement the country based forest monitoring system and to improve the measurement and reporting of national forest area and change. However, still variation in forest definition exists, leaving variation on national estimates of forest cover and change.

3.3 Forest cover change accuracy analysis

During the last decades, several global forest cover and land use change datasets have been derived from varied spatial and temporal resolution of remotely sensed data. However, these remotely sensed datasets were not free from errors. To complement these datasets to NFMS an assessment of how well locations of mapped change would correspond to actual areas of change on the ground is important. In this specific research question, basic visual assessment was conducted to identify obvious errors and other concerns corresponding to accuracy of two global forest cover change products and their agreement in forest cover change. The two global forest cover change datasets understudy were first aggregated to the common definition to assess the accuracy assessment of each dataset (Table 1; Table 2). Then each dataset was overlaid spatially to one another and the reference dataset (GEI) to allow an assessment for each datasets accuracy and agreement/disagreement between them. The result observed from accuracy assessment and the agreement between the two datasets was presented under section 3.3.1 and 3.3.2, respectively.

3.3.1 Accuracy estimates for global forest or land cover datasets

3.3.1.1 Accuracy of global forest cover change (Hansen et al. GFCC)

Hansen et al. GFCC dataset is the first high resolution global map of the 21st century. According to authors it provide detailed information on forest extent, forest loss and gain from global to local scale. The published map also intended to provide information in policy formulation and implementation to slow deforestation and then climate regulation, which can be an important input for REDD+ activities. To determine the applicability of the dataset to the stated objectives at national level Ethiopia, we carried out an accuracy assessment of this dataset based on sample point location. The outcome of 400 sample points relative to reference data set was presented in the error matrix below.

Table 8. Confusion matrix of estimated Hansen et al. GFCC corresponding to the reference data from GEI. The diagonal matrix from the table indicates sample point classes which are correctly classified; while off diagonal elements represent either of commission or omission errors (Foody 2002).

Hansen et. al. GFCC	Google earth imagery (reference data)					Row total	Users acc.
	Forest loss	Forest gain	Stable forest	other woodland	Stable non forest		
Forest loss	86	6	11	9	18	**130**	**66**
Forest gain	4	33	6	2	15	**60**	**55**
Stable forest	4	0	53	10	8	**75**	**71**
other woodland	2	1	9	18	5	**35**	**51**
Stable non forest	6	1	6	33	54	**100**	**54**
Column total	**102**	**41**	**85**	**72**	**100**	**400**	
Producer acc.(%)	**84**	**80**	**62**	**25**	**54**		

From confusion matrix in Table 8 the overall accuracy estimated from the total 400 sample point selected is 61%. Based on this sample point study estimation, 61% of Hansen et.al GFCC derived map for the national level of Ethiopia was correctly classified as the same in the high resolution GEI. Per class cross referenced user accuracy ranges from 51% (for other woodland class) to 71% (for stable forest class). This user's accuracy indicates the probability that a sample from forest cover map actually matches the real world observation from reference dataset. As an example, a user's accuracy of 66 % for a category of forest loss in Hansen et al. GFCC would indicate that 66% probability of correctly detecting and classifying forest loss areas from the total category of sample points assigned to this class. The remaining percentage (34%) indicate inclusion of non-deforested areas to the forest loss category which it does not truly belong. For illustration, out of 130 samples classified to forest loss in Hansen et al. GFCC dataset, 8% and 14% are indeed turned out to be stable forest and stable non-forest classes, respectively, in GEI. This might cause overestimation of the area of deforestation for the whole country.

On the other hand, the producer's accuracy estimated for this dataset ranges from 25% representing other woodland class to 84% representing forest loss class. There was omission of 16%, 20%, 38%, 75% and 46% from forest loss, forest gain, stable forest, other wood land and stable non forest classes respectively. This cause underestimation of the categories due to they are actually exist in reference of the GEI. From this estimated producer's accuracy, relatively forest loss and forest gain followed by stable forest are correctly classified and mapped. The low producer accuracy of woodland class was mostly prominent in dry shrub land covered part of the country.

The overall accuracy observed for Hansen et al. GFCC dataset is less than. But, if we are most specifically interested to see the change classes for REDD+, forest loss and forest gain claim 84% and 80% of producer's accuracy which is quite good classification. A more careful look at the error matrix showed that there was significant confusion in discriminating forest loss and forest gain, most importantly from stable non-forest and stable forest classes (Table 8). Therefore, although the producer of this dataset claim that 84% of forest loss and 80% forest gain classes were correctly classified as such, the users' of this dataset will find that only 66% and 55% of the dataset is actually forest loss and forest gain, respectively, indicating less classification.

3.3.1.2 Accuracy of FRA Remote Sensing Survey

The Remote sensing survey was sample based forest resource assessment to obtain globally consistent information on forest area and change at global, regional and climate domains. It was designed to improve measurement and reporting of forest area and change over time and provide an input to national and international reporting processes like for REDD+ activities. In order to assess the applicability of this dataset to the national scale Ethiopia 230 sample points were selected and assessed for accuracy with available GEI between 2000 and 2005. The outcome of these sample points are presented in the error matrix (Table 9).

Table 9. Confusion matrix of FRA RSS forest cover change estimate corresponding to the reference data from Google earth imagery. The numbers corresponding to each class are sample points.

FRA RSS	Google earth imagery (reference data)					Row total	User's Acc. (%)
	Forest loss	Forest gain	Stable Forest	Other woodland	Stable non forest		
Forest loss	**15**	1	0	2	2	**20**	**75**
Forest gain	0	6	1	1	2	**10**	**60**
Stable Forest	2	0	60	9	1	**72**	**83**
Other woodland	1	0	8	48	9	**66**	**73**
Stable non forest	0	1	2	3	56	**62**	**90**
Column total	**18**	**8**	**71**	**63**	**70**	**230**	
Producer acc. (%)	**83**	**75**	**85**	**76**	**80**		

The overall accuracy estimated from the error matrix is 80%. The user's accuracy computed for each category showed that there is more probability of land cover classification from RSS dataset matching what is actually exist on the ground when referenced with GEI. Stable non-forest and stable forest classes have high probability of assignment to the same class followed by forest loss and other woodland classes. Comparison of visual assessment of land-cover classifications from high spatial resolution GEI for selected tiles and sample points in Ethiopia indicated that FRA RSS yield users accuracy of nearly 90% for stable forest (no change area) and 83% for forest and 75% for forest loss. The producer accuracy also indicates more probability of referenced sample areas being the same as the classified RSS map relatively in all classes. Concerning to the forest gain there was less omission (2 sample errors) than commission (4 sample errors) causing the area of forest gain overestimated (Table 9). The more commission in forest gain (40%) indicate the inclusion of other classes to the forest gain category with reference to the actual high resolution GEI.

3.3.1.3 Comparison of users and producers accuracy

A comparison of the automatically classified land cover change for each dataset with high resolution GEI showed a difference. Overall, the average cross-referenced accuracy of global forest cover land use change datasets against the reference data of GEI in Ethiopia vary from 61% in Hansen et al. GFCC and 80% in FRA RSS (Table 8, Table 9). Some classes, for example, stable forest and stable non forest in RSS have high producer and user accuracy, indicating more accuracy of the classes mapped. Such experiences were more clearly observed in RSS dataset than Hansen et.al GFCC. On the other hand, there was more confusion to assign forest gain and forest loss. Inspection to the error matrix showed that, there was high producer than user accuracy which is due to the inclusion of other stable land cover classes as forest loss although they are stable. For example, around 60% user accuracy of forest gain in both dataset might be due to the commission of other classes (40%) as forest gain.

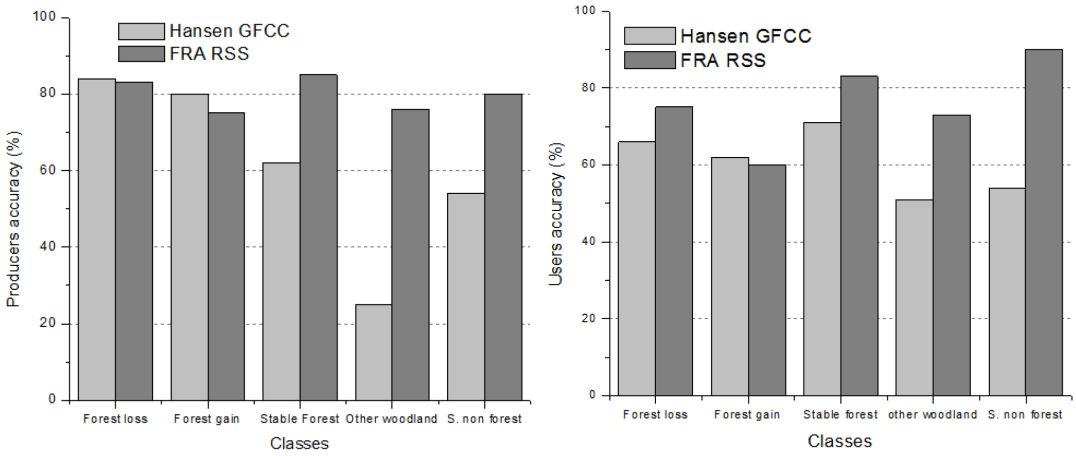


Figure 3. Comparison of producers and users accuracy for Hansen et al. GFCC and FRA RSS, based on assessment made for each against GEI (from Table 8 and Table 9). Adopted from Herold et al. (2008).

However, as shown in Figure 4 broadening from sample point to spatial extent indicate under estimation of forest loss and forest gain classes. The classification result of other woodland class is also match lower in Hansen et.al GFCC compared to FRA RSS with 75% exclusion and 49% inclusion, in total showing the under estimation and misclassification of this specific class to stable non forest, stable forest and forest loss classes. Omission and commission error to the forest gain and forest loss classes from Hansen et al. GFCC dataset cross referenced with GEI is indicated in Figure 4.

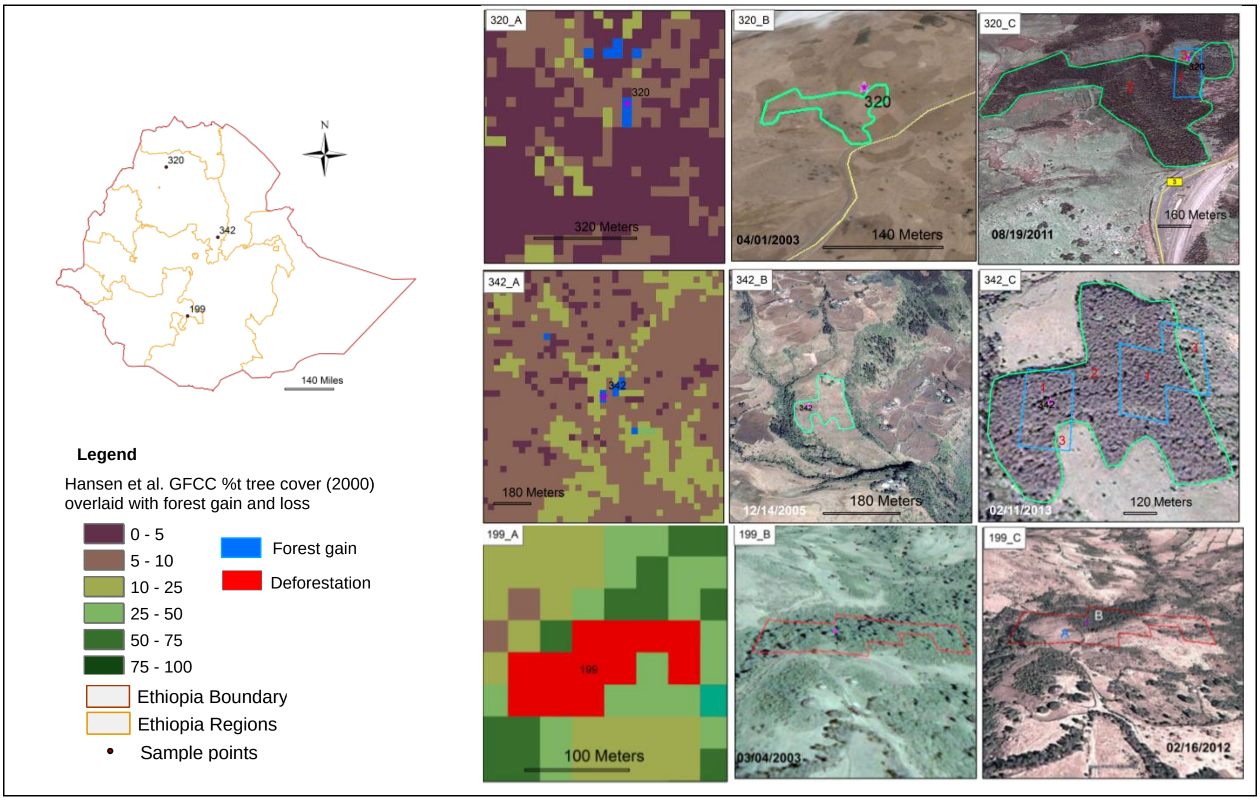


Figure 4. An example of forest gain and loss in Hansen et al. GFCC data set and error of omission and commission observed during mapping.

In Figure 4 samples 320 and 342 demonstrate omission and commision error of forest gain, while sample 199 demonstrates omission and commision error of forest loss for the area centred in the point and indicated in the national map. Both maps labelled with A are classified percent TCC overlaid with forest gain (blue colour) for the upper two and forest loss observed in 2005 (red colour) in the lower image (199_A). The middle figures represent GEI (i) without tree cover (320_B and 342_B) and (ii) with tree cover (199_B) during the imagery date indicated on each image. The final map labelled with C is forest gain observed (320_C and 342_C) and forest loss observed (199_C). The extent of forest gain and forest loss observed from Hansen et al. GFCC with label A and that of corresponding C has agreement and disagreement (commission and omission errors). For example, in both samples (320_C and 342_C) the polygon with sign 1 shows that there is agreement in extent of forest gain, sign 2 indicate omission error and sign indicated by 3 indicate commission error while mapping. While 199_C indicate there was partial agreement in deforestation where blue colour A is deforestation agreement and white coloured B area indicate commission error.

3.3.2 Agreement between FAO RSS and Hansen et al. GFCC dataset classes

Areas of agreement and disagreement between the two datasets were also compared based on 230 sample sites that are commonly selected within the tiles. Table 10 shows spatial agreement along the diagonal line between the datasets and disagreement on the off-diagonal values.

Table 10. Spatial agreement between Hansen et. al. GFCC and FRA RSS

	FRA RSS					
Hansen GFCC	Forest loss	Forest gain	Stable forest	other woodland	Stable non forest	Row total
Forest loss	6	0	5	7	2	20
Forest gain	0	2	0	0	1	3
Stable forest	3	3	46	12	10	74
other woodland	2	1	10	21	2	36
Stable non-forest	9	4	11	26	47	97
Column total	20	10	72	66	62	230

From the agreement matrix in Table 10 numbers in the diagonal shows each class agreement from the total number of observed sample points. For example, out of 20 sample points assessed in both datasets, only 6 sample sites has spatial agreement indicating forest loss in both datasets. The overall agreement between the two dataset is only 53%. The other off-diagonal values indicate disagreement between the two datasets corresponding to the land cover classes considered in the sample site.

Moreover, in Hansen et al. GFCC dataset woodland has insignificant relationship or agreement relative to reference data and the FRA RSS, suggesting no acceptable levels of accuracy in this class. This was particularly vivid in dry areas of the country where there was more misclassification of woodland areas as non-forest classes although most area is covered with woodland as observed from reference datasets. But the inherent difference between the two data datasets might be due to different definitions of forest, different tools and methods used and forest types considered. This makes it difficult to determine if there was due to misclassification or other source of errors.

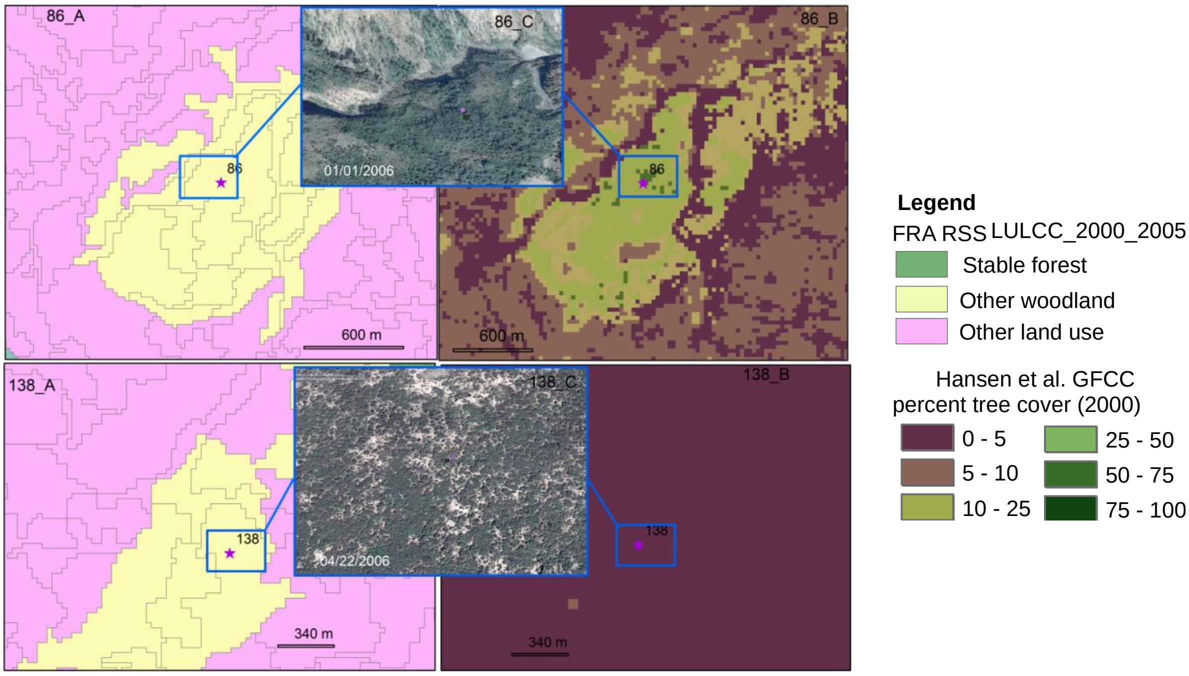


Figure 5. Examples of disagreement in land cover classes mapped in the data sets and actual observation from high resolution Google earth imagery

AS shown in in Figure 5 sample 86 and 138 there was disagreement in mapping between the two data sets (i.e. RSS labelled with “A” left figures and Hansen et al. GFCC labelled with “B” right figures) .On sample 86 RSS is miss labelled as woodland also the area is forest (86_C) which complement to Hansen et al. GFCC dataset (86_B). On sample 138 both are misclassified as wood land (138_A) and non-forest (138_B) respectively, while the area was forest cross referencing to high resolution Google earth imagery (138_C).

Overall, visual inspection of disagreement areas based on sample site showed that there was more discrepancy in other woodland, non-forest covered and forest gain classes. Highlighting only the major differences, three major areas of disagreement between the two datasets were found in (i) Somali region along south eastern border of Oromia and south Oromia. These areas were mostly covered with other woodland class in FAO RSS, and non-forest in Hansen et. al. GFCC dataset. (ii) south-western part of the country, particularly in Glabella regional state where the area was completely covered with non-forest in FRA RSS and was more than 10-50% tree cover in Hansen et al. GFCC. But, inspections of reference data suggest that the area was non-forest with agricultural land and woodland mixed with grasses. (iii) The third disagreement area was northern most part of the country in Amhara and Tigray regional states. In these areas more forest gain was observed from reference data-set, particularly after 2005, but where not detected by Hansen et al. GFCC. This might cause the underestimation of forest gain in this area, which is in complement with the accuracy assessment of Hansen et al. GFCC regarding the global forest gain.

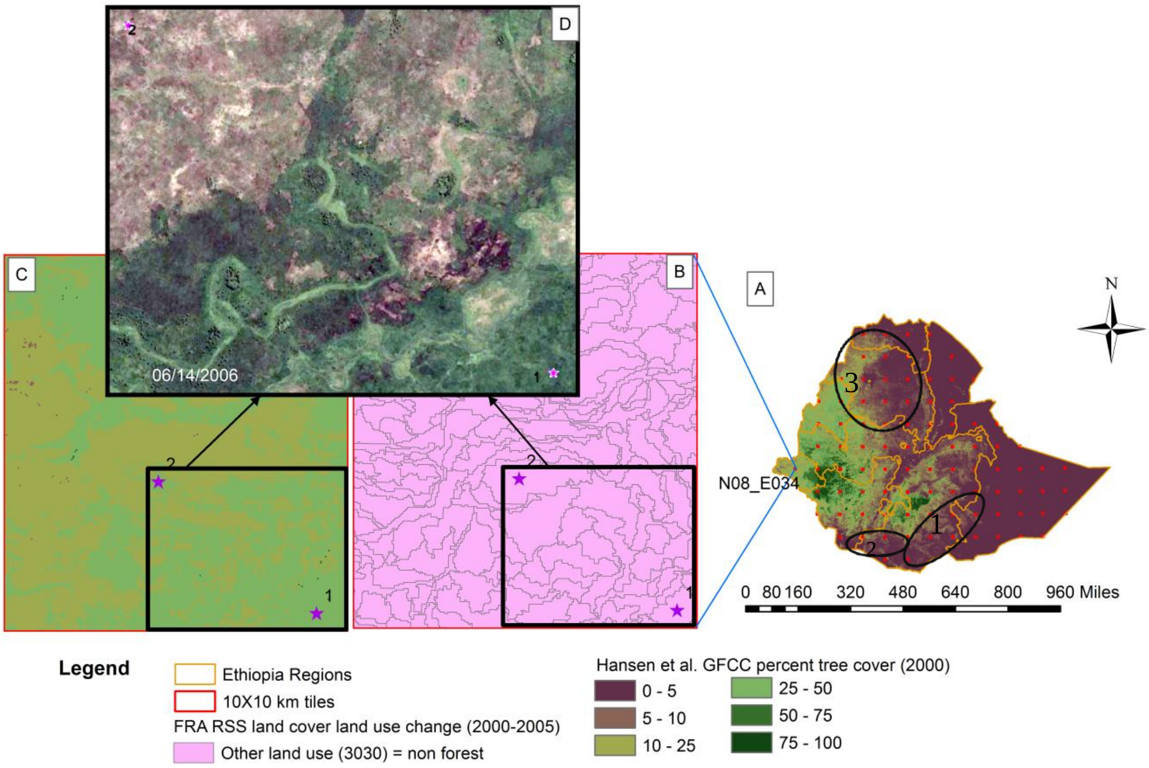


Figure 6. Land cover disagreement between FRA RSS (B) and Hansen et al. GFCC dataset (C) for the area located in south western Ethiopia, specifically Gambella regional state for global tiles 08°N and 34°E (A).

Overlying both datasets on available GEI the area within the black box was non-forest covered with agricultural and sparse shrubs on the north part and woodland or bamboo mixed with grass along the river (Figure 6 D). Relatively there was more disagreement with Hansen et al. GFCC dataset than with RSS result. However, no forest gain or loss was recorded in both data sets and the reference data in the study period for this site. Figure 6A shows that the Hansen et al. GFCC percent tree cover map and the red boxes represent 10 X 10 km tiles within 1° X 1° intersection from where sample sites were selected to check the agreement between the two data sets. The black ellipse coded with number 1-3 are also another examples of disagreement areas discussed above, specifically related to woodland and forest gain.

3.4 Adaptability of global forest cover change estimates to national REDD+ forest monitoring approaches

Remote sensing based forest monitoring provide activity data that help countries to assess forest cover changes and provide basis for country to reports on GHG emissions or removals under UNFCCC, along with emission factors. As recommended and indicated by IPCC (2003) and FAO (2013) RS data to be used in NFMS and for MRV of REDD+ activities should be (i) in compliance with the principles of consistency, completeness, comparativeness, accuracy and transparency, (ii) used to measure annual changes in land cover/use through a consistent methodological approach over time, including the assessment of historical rates of deforestation and degradation, (iii) help countries to assess and also access activity data, and (iv) reflect historical trends in land use management that ensure the estimates reported are transparent and comparable. Accordingly, the MRV system on submission of information on reference levels should also be consistent with the suggested global observation of forest cover and land dynamics (GOFC-GOLD) methods and the emerging standards. So to say identification of existing data in line with IPCC good practice guiding principles and standard is important for REDD+ national forest monitoring. Giving emphasis to AD, Table 14 indicate the 5 IPCC reporting principles and indicators related to the reporting parameters and the data sources used to see their complementarity into the REDD+ national forest monitoring context.

Table 11. Level of adaptability of global forest cover change datasets to the NFMS for REDD+ based on the five IPCC reporting guidelines. Parameter values are assigned as: H= high, M= medium and L= low based on the level of information that each corresponding parameters exist in the dataset.

IPCC principles	Specific parameters	Hansen et. al GFCC			FAO RSS		
		Parameter value			Parameter value		
		H	M	L	H	M	L
1. Consistency	1.1 Capability of representing land use categories			√	√		
	1.2 Capability of assessing FCC over time	√				√	
	1.3 Consistent methodologies used over time	√			√		
	1.4 Land cover conversions can be tracked on a spatially explicit basis	√			√		
2 Completeness	2.1 All land within a country is included	√					√
	2.2 provide country specific forest resource data	√				√	
	2.3 Land use change is included			√	√		
	2.4 provide activity data	√				√	
	2.5 Enhance report on emissions and carbon removals from GHGs		√			√	
	2.6 Change in forest cover extent included	√				√	
3 Transparent	3.1 Assumptions, definitions and methodologies used are clearly explained	√			√		
	3.2 Assumptions, definitions and methodologies used can facilitate replication by users	√			√		
	3.3 The data sources used for assessment should be available or open	√				√	
4 Comparability	4.1 Forest cover and change estimates are comparable with other external reports and estimates		√		√		
	4.2 Methods and assumptions applied accounts those with REDD+ consensus		√		√		
	4.3 Have common methodology and guidance with national R_PP		√		√		
5 Accuracy	5.1 Can achieve acceptable level of accuracy with reference data used		√		√		
	5.2 The data set is adjusted to national circumstances to improve accuracy		√		√		
	5.3 Capable of representing land use categories			√	√		
	5.4 Can represent change between land cover categories to estimate GHG emission and sinks		√		√		
Total score		**10**	**7**	**3**	**13**	**6**	**1**

Based on equation 5 under section 2.3.5 the adaptability level of each dataset to national REDD+ forest monitoring approach was calculated. Accordingly, the overall adaptability of Hansen et al. GFCC data set is 67.5 %. On the other hand FAO RSS has 80% adaptability level, indicating more conformity to national forest monitoring approach than Hansen et al. GFCC dataset. This does not completely mean that FRA RSS has no drawbacks and the reverse is true to Hansen et al. GFCC. It is true that, although overall assessment of the datasets using the IPCC guiding principles indicate the more acceptable level of adaptability for RSS data set, both have their own strength and limitations.

According to the parameters we set for assessment, almost both datasets qualify the consistency principles, except the Hansen et al. data not represent land use change classification system at all and FRA RSS data set followed bi-temporal than time series approach of change detection. With regard to other principles, FRA RSS has more accuracy and comparability. This might be due to the more accuracy and agreement in forest cover extent and loss it has with reference to high resolution Google earth imagery (section 3.3). On the other hand Hansen et al. GFCC dataset more complies with the principle of completeness and transparency due to the major advantage of wall-to-wall mapping and openness of the dataset for users. Overall, the major advantage and limitation of the two datasets (one over the other) are concluded in the Table 12.

Table 12. Major strengths and limitation of Hansen data set and FRA RSS

Data set	Strength	Limitation
FRA RSS	- provide more acceptable level of accuracy, - involve national forestry and RS experts to complement classification results to national circumstances, - provide ecological context forest cover and forest land use change and - It is broadly in line with UNFCC LULCC guideline that may suffice REDD+ activities more.	- incomplete due to sample based and not provide forest cover changes at national level, - Some inconsistent classification and labelling of agricultural and settlement with tree covered areas as forest and other forest classes as woodland and vice versa.
Hansen et.al GFCC	- It is well to well coverage mapping, - more open and freely downloadable, - guarantee access of timely information about forests and its change and - can empower local people/external body everywhere to better manage and analyse forest cover changes	- Less accurate, specifically underestimating dry woodland covered parts of the country. - Relatively less comparable with other external data sources and with reference data, due to over estimation of forest cover extent and underestimation of forest loss. - It less likely follow IPCC land cover land use GPG

4. Discussions

4.1 Variation in forest or land cover change class mapping

4.1.1 Disagreement on forest resource of Ethiopia and the challenge for REDD+ MRV

Ethiopia has been in the REDD+ process since 2008 and is now a REDD+ participant country to reduce deforestation and forest degradation and to enhance carbon stocks (FDRE 2011b; Moges et al. 2010). The country also created and submitted R_PP for implementation of REDD+ activities and then to transpire in to the green economy vision (FDRE 2011b). One important mechanism designed to implement the objectives of climate resilient green economy (CRGE) in Ethiopia is through forestry sector (FDRE 2011a), where determination of forest resources of the country has immense role in MRV for REDD+ activities. According to national forest definition described in WBISPP (2004) and on overview of REDD+ process in Ethiopia by Moges et al. (2010), forest resource of Ethiopia broadly consist natural high forest, woodland, and shrub land.

On the other hand, Moges et al. (2010) indicated that Ethiopia is endowed with woodlands lands accounting 29 million ha (25% of the country area) and store more carbon. The regions with the largest area of woodlands reported are Somali (45%), Oromia (34%), and Benshangul Gumuz (8%). Similarly, based on FAO forest definition and considering forest proclamation number (No. 542/2007) which included high forests, plantations, woodlands and bamboo forests in to forest land cover class forest resource of Ethiopia was estimated to be more than 35 million ha. In complement to the idea successful implementation of REDD+ activities also require consideration of other land cover classes (e.g. woodland), besides natural forest, because of they have wide role in carbon stocks in Ethiopia (Moges et al. 2010). However, our result from Hansen et al GFCC dataset indicted that there was more exclusion of woodland classes in these regions (specifically in Somali and Oromia) as shown on Figure 5 (sample point 138_B) and Figure 8 (sample point 22_B). Reasons for variation and its influence are discussed in the next section.

4.1.2 Variation in forest definition and the challenges of land cover class mapping

The substantial variation in definition of forest has important ramifications in land cover/land use mapping and leaves many ambiguities (Watson 2009) on measuring and reporting for REDD+ activities. In line with this study, our study also revealed that different forest and forest loss extent were quantified while assessing different global and national forest cover change data sets for Ethiopia (Table 7). For example, utilising a 10% minimum canopy cover threshold as a key characteristic of the forest classification in the year 2000, the study observed 31, 861,605 ha forest (about 28% total land area) from Hansen et al. GFCC dataset and 13,705,000 ha forest cover (12.4% of total land area) from FAO FRA dataset (FAO 2010a, b). Similarly, considering 20% national minimum canopy cover threshold more unacceptable variation was identified between national and global data sets on forest cover and forest cover change estimates. In this case, the national WBISPP (2004) estimated 3,651,936 ha (3.2% of total land area) forest extent, while the Hansen et al. GFCC estimate showed 16,563,209 ha (14.6 % total land area). Although, similar data sources were used, the main reason for the difference is the classification scheme used. The national WBISPP used 25 ha MMU, where community forest or patches of woodland or shrub land less than 25 ha in agricultural area are classified as a cultivated land or non-forest(WBISPP 2004).

These two global land cover/use data sets used for accuracy assessment were prepared using different classification systems and methodologies, although they have similarity mainly (i) in use of the same data sources (Landsat derived) and (ii) minimum mapping height (i.e. > 5 m) and inclusion of plantations and natural forest in their definition of forest. For instance, the FAO RSS used the FAO land cover land use definition (Table 4), first to automatically classify land cover classes, and then to aggregate classified land cover classes in to 5 ha minimum mapping unit (MMU) based on percent tree cover and tree cover proportions. On the other hand, Hansen et al. GFCC classified forest as land with more than 5m tree cover at Landsat pixel size. In this case all trees above 5 m are identified as forest (Hansen et.al. 2013).

The current study imply that the choice of forest definition which result in the variation in determination of threshold value is one main reason for variation in forest extent and forest loss observed from different data sources (Table 5; Table 6). In agreement with, Zomer et al. (2008) showed the impact of variation in determination of threshold value to categories land cover classes on afforestation and reforestation activates. Similarly, a report by Watson (2009) on forest carbon accounting also indicated that the use of minimum or maximum threshold values in forest definition can induce inclusion or exclusion of one land cover class to the other and vice versa. Both authors reported that, the usage of low minimum canopy cover threshold can induce inclusion of other land cover with sparse forest covered and agricultural areas as forest. Conversely, use of high maximum canopy cover threshold can enhance exclusion of areas of sparse forest without efforts to account them for REDD+ activities.

Besides, the choice of minimum area and height thresholds have a considerable impact in land cover calcification and then in REDD+ activities. Consideration of low minimum forest area may permits the inclusion of patches around farms and settlements, often serving as woodlots as forest (Watson 2009). This was in complement with the case of Hansen et al. GFCC. On the other hand, high forest area thresholds encourage inclusion of contiguous forest areas as forest (the case of FRA RSS). For example, the application of 0.5 ha FAO minimum mapping criterion removes forest patches <0.5 ha from forest class and classify it as other woodland or other land use (non-forest), but which can be considered as forest if it was detected within the single pixel minimum mapping unit of Hansen et al. Low tree height thresholds may also permit short woody, forest vegetation to be included in forest classes (Neeff et al.

2006; Watson 2009). Over all, the choice of forest definition can affect land cover categories and land that have to be available for different REDD+ activities. Such a loos definition might be reason for variation in forest extent and also prevents their accurate comparisons on forest carbon stocks.

4.2 Accuracy analysis

Many accuracy assessment studies were undertaken for many different purposes, where these purposes can influence the choice of appropriate accuracy assessment level and results (Wulder et al. 2006). In these studies different figures were suggested in determining acceptable level of accuracy for land cover/use classes. For example, a review of Anderson (1976) on land use land cover classification system indicates that 85% minimum level of accuracy in the identification of land use and land cover categories from remote sensing data, even though it was considered to be traditional to measure success or failure of a land cover mapping. Although no consensus are reached, several recently completed regional-scale land cover mapping projects (Herold et al. 2008; Wulder et al. 2006) also indicated the acceptable level of producer's and user's accuracies in the range of 50–70%.

Likewise, the present study employed an accuracy assessment of the two datasets to assess how the maps from each data set were accurately mapped and how it complies for REDD+ activities. Eventually, the levels of overall accuracy of both datasets and producer and user accuracy estimates observed from the 5 individual land cover classes vary. The overall accuracy for FAO RSS was 80%, which has more acceptable level of classification accuracy relative to Hansen et al. GFCC dataset that has 61% fairly classification accuracy. Relative to Hansen et al. GFCC all classes in FRA RSS have well classification ranging from 60% (forest gain) to 90% (stable non-forest) users accuracy and 75% (forest gain) to 85% (stable forest) producers accuracy (Table 9). No formal accuracy assessment was undertaken for this dataset at global scale (FAO 2012). However, the revision of the dataset showed 77-81% overall agreement of land cover classification labels before and after the expert review, which is in agreement with our results.

In Hansen et al. GFCC the classification accuracy for individual classes varies greatly. Forest class relatively has good user accuracy (71%) followed by forest loss and forest gain with fair or moderately low classification (66 and 55% respectively) than the other two remaining classes. Contrarily, the producer of the Hansen et al. dataset complains that the accuracy of forest loss (84%), forest gain (80%) and stable forest (62%), respectively, appears to be well or fairly classified. The forest loss accuracy estimated was in complement with global accuracy assessed and reported by Hansen et al. (2013b) . However, the accuracy of estimated forest gain in our case was converse of the data producer. Hansen et al. (2013b) estimated user's accuracy of 82% and producers accuracy of 48% globally, while in the current study only restricting to sample points we observed 80% producer and 55% user accuracy. However, switching visual assessment at sample site to measuring area change for two sample sites showed as underestimation of forest gain detected, revealing the omission error estimated by the producer (Figure 4 sample 320 and 342) . The remaining two classes' namely stable non-forest and other woodland class had low and insignificant producer accuracy in respect of GEI respectively, suggesting unacceptable level of accuracy in woodland class due to more omission (Table 8).

Here it is important to mention what factors may have contributed for the variation between the two datasets and low accuracy results in some classes. Variation in accuracy between the two datasets might be arise due to variation in intended application as reported by Wulder et al. (2006) or the variation in definition of forest and classification methodology implemented in these products as revealed by Zhu et al. (2000). On the other hand, pointing to the land cover classes two main factors might be the reason for variation in classification accuracy. The first, one is the methodology used to define and classify forest. The FAO RSS used 5 ha as a minimum mapping unit for classification to land cover/use category, while GFCC used Landsat pixel size to identify percent tree cover and forest cover changes. This may enhance inclusion and exclusion of land cover classes (Watson 2009; Zomer et al. 2008). For example, many stable non-forested agricultural land and wetland areas were detected as forest with more than 10% tree cover in Hansen et al. GFCC might be due to low minimum mapping unit and growing season data used for mapping (Figure 8 sample 31 and 79). In line with this point, Zhu et al. (2000) also indicated the influence of timing of data acquisition and forest definition results in misclassification and then leads to lower class accuracy.

The second and most important point that causes variation in agreement and less accuracy of Hansen et al. GFCC than RSS, particularly in dry shrub or woodland covered part of the country might be due to additional FAO forest definition for woodland class mapping. In this case crown cover of more than 10% of trees not able to reach a height of 5 m at maturity in situ (e.g. "dwarf or stunted trees") are considered as other woodland cover (http://www.fao.org/docrep/006/ac840e/AC840E02.htm, accessed on 27 July, 2014). However, in (Hansen et al. 2013a, b), such areas were classified as non-forest, because they were considered to be less than 5 m. However, inspecting the reference data with visual assessment showed that more of such area is covered with woodlands or shrubs, but with visual inspection alone it is difficult to understand weather the area is more than 5 m or below.

4.3 Agreement between FRA RSS and Hansen et al GFCC datasets

The spatial agreement matrix presented in Table 10 showed that the overall agreement between FAO RSS and Hansen et al. GFCC datasets was only 53%. The disagreement observed between the two datasets is in line with the report of Giri et al. (2005), where the author of this paper also observed disagreement in forest distribution between the two global land cover products namely moderate resolution imaging spectrometer (MODIS) land cover dataset and global land cover 2000 (GLC-2000). Moreover, as shown in the study of Fritz et al. (2011) on highlighting uncertainty in global land cover maps for user community, the spatial disagreement between non-forested cropland classes and forest classes from three global land cover products, that is GLC-2000, MODIS and Glob Cover_2005 were large. In agreement with our result, Fritz et al. (2011) also demonstrated major disagreement between these datasets in northern part of Ethiopia, where large cultivated area in the location was labelled as other woodland in MODIS and forest in GlobCover_2005. In line with Herold et al. (2008), the current study also showed that variation in class accuracy as the major driver of spatial disagreement between datasets. Relatively, less accurate land cover land use classes (Table 8; Table 9) showed low agreement among datasets. Similarly, study conducted by Colson et al. (2009) revealed existence of variation in forest class due to variation in definition, where more forest definitions in use has been shown by Lund (2006).

Certainly, differences in the accuracy estimated and disagreement between the two datasets and land cover classes can indicate uncertainties in forest cover and its change in Ethiopia, particularly woodland covered areas of the country. The reasons for difference could be many. However, divergence in definitions and use of different methodologies applied for mapped data are the major once (Fritz et al. 2011; McCallum et al. 2006). On the other hand, the use of too broad or too small threshold in subsequent of variation in forest definition in land cover classification are main reasons for over or underestimation of the land cover and land cover change classes identified. These all including differences in time of imagery acquisition used and use of national expert review for validation of classified data (FAO 2012) can account for differences in agreement between the datasets or land cover classes examined

4.4 Adaptability of global forest cover change estimates to national REDD+ forest monitoring

The analysis shows that FAO RSS has 80% adaptability level, while it was 68% for Hansen et al. GFCC (Table 11). Relatively both results are consistent with the accuracy assessment result observed. Moreover, although there is no huge variation, this overall adaptability level indicates the more conformity of FRA RSS to the national forest monitoring approach than Hansen et al. GFCC dataset. Most importantly, the variation observed might be due to the former is more in line with IPCC land use good practice guide line in forest definition and land cover/use classification system (FAO 2012).

Restricted to individual datasets, the following major strengths and limitations were identified in addition to these described in Table 12 and have to be considered while adapting and complementing the data set to the REDD+ MRV and national forest monitoring system. The major strength of FRA RSS is the participation of national forestry and remote sensing experts from more than 100 countries. This causes examining of critical circumstances of respective countries while they participated on correction, optimisation and validation of land cover land use labels classified by FAO and other participants. This is in line with Giri et al. (2005), where the authors identified the advantage of involving experts from national, regional, and international organizations for optimized global land cover change mapping (GLC_2000). In addition, the classification scheme, data source and methodology implemented by FAO RSS accounted for variation in biophysical features and changing seasonality (FAO 2012). This might be another reason for more accuracy and adaptability of FRA RSS relative to Hansen et al. GFCC which in turn accounted only growing season data sources. On the other hand, FRA RSS dataset lacks wall to wall observations and mapping, underpinning the incompleteness of the dataset. This might raise uncertainties, particularly in uncovered areas of the country. However, the methodology developed is transparent, accurate and repeatable, that will help to reproduce wall to wall land cover mapping for the whole area of the country if further work will be carried out.

Related to Hansen et al. GFCC, the major advantage is that it adopts global wall to wall forest extent and change mapping that can be updated on an annual basis examining annual change on forest cover (WHRC 2013). Further, it has transparency and consistency where the full dataset is open to everyone, the dataset can be updated based on comments received from users and data reviewers together with inconsistencies of the earlier work, and the data can be freely downloadable to consistently monitor forest area changes with good spatial and temporal detail that can be viewable at http://earthenginepartners.appspot.com/science-2013-global-forest. But, the major limitation of this dataset is lack of accuracy and comparability related to reference dataset and national estimates, where forest extent was overestimated and forest loss and gain were under estimated. In line with this, Tropek et al. (2014) commented on underestimation of forest loss and gain and overestimation of forest extent as main limitation of this product. Similarly, the present study was looked back for two sites of which the data were collected from previous field work using open data kit (ODK collect) in September and October 2013 (Figure 7).

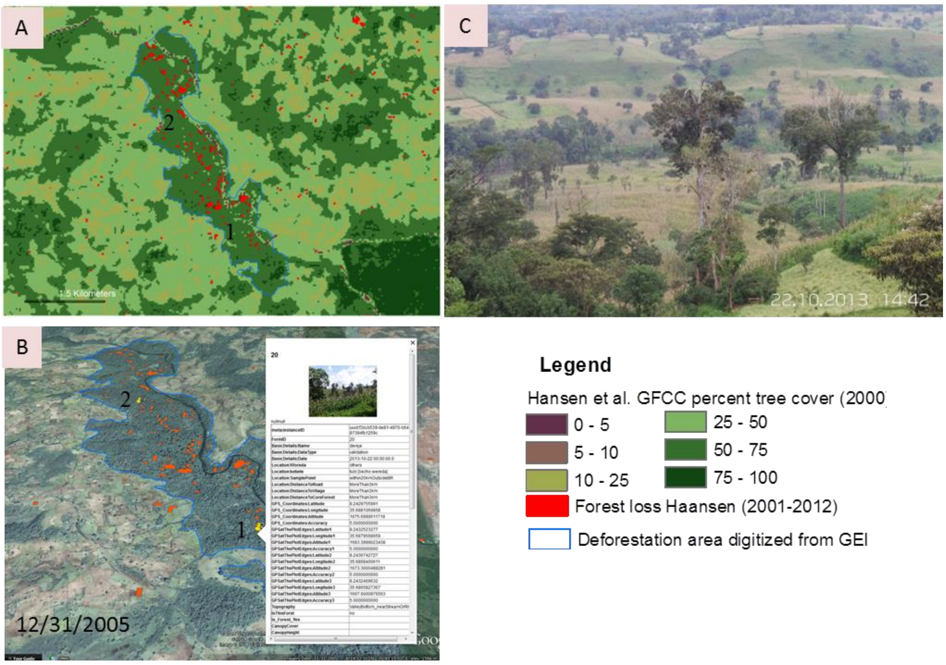


Figure 7. Underestimation of forest loss by Hansen et al. GFCC dataset.The upper left figure (A) is Hansen et al. GFCC processed tree canopy cover overlaid with forest loss detected from 2001 to 2012 with red polygons. The image in figure 7B is GEI in 2005 showing the existence of forest inside the area of blue polygon overlaid with the same forest loss detected by Hansen et al. GFCC. The metadata on the image B was field attribute information (GPS location, deforestation date from local farmers, current status of the area) collected for sample site 1 verifying the area as non-forest with tagged image on the

ODK data. Figure 7C is a photograph taken during field work for the site 2, having similar information as site 1. During the visit to the study area, the area was completely deforested before 2012 and covered with agricultural crops and with remnants of big trees along riverine, but this was underestimated in Hansen et. al. GFCC dataset.

Overall, the state of the art of forest cover/change monitoring approach was different between the data sources. Both assessed data sources vary in method of generation, mapping scale (e.g. sampling Vs wall to wall), definitions and classification systems employed, attributes collected, accuracy, time scale covered in the assessment, and data format (tabular, pixel, polygon area). The results estimated from the two main datasets were also found to be relatively variable. Even within a country, high variation in labelling of land cover categories were observed that might be due to application of different procedures and definitions of forest in different eco-regions. The state of knowledge of local experts and the way they perceive forest can also vary according to local condition. As a result of these variations in accuracy, spatial agreements and forest definition between datasets, it is worthy that the users of datasets particularly, in Ethiopia have to cautiously examine the compassion of these land cover products before use for national forest monitoring application of REDD+ activities.

In agreement with the work by Fritz et al. (2011), the present study also recommend that user of these products for national, regional and local REDD+ forest monitoring activities have to examine the strengths and limitations of each data product identified before use. Moreover, examining the disagreement areas of these two products in combination with other highest resolution imageries and field campaign validations can enhance their combined usage for national forest monitoring system. As such the two global datasets have more significance to improve and produce repeatable and comparable national land cover datasets for REDD+ implementation. More importantly it can be helpful to use them in agreement through compromising the drawbacks of one dataset over the other.

4.5 Problems in accuracy assessment

Although the basic approaches to accuracy assessment seem relatively straightforward, many problems are often encountered while assessing the datasets. These problems are mostly associated with limited amount of information to effectively assess change on the datasets with respect to actual ground data. Some of these interrelated problems that limit the quantification of classification accuracy of the datasets are described below.

One important aspect was related to sampling designs and sample selection. The sampling design used to select sample points upon which the accuracy assessment based is of major importance in land cover classification and accuracy estimation (Foody 2002). In addition, Foody (2002) indicated that availability of high quality fine spatial resolution imagery at an appropriate time is a proxy for ground observation. In this line our study was not without limitation, where actual data collection within time series from GEI was frequently constrained with cloud or lack of imagery (Table 14), restricting access to more reliable land cover information in time and space. As a result the current study employed two sampling strategy i.e. systematic sampling strategy to incorporate all RSS tiles with available imagery and simple random sampling to incorporate Hansen et al. GFCC products outside the RSS tiles. So, the sampling strategy and number of samples designed Wulder et al. (2006) for accuracy assessment in each class might have an influence on the evaluated classification accuracy (Foody 2002; Friedl et al. 2000).

Unfortunately, other source of errors such as mislocation of the samples on classified datasets with the ground truth data (Giri et al. 2005) and interpretation approaches (Ismail and Jusoff 2008) might contribute to the pattern of misclassification observed from the error matrix. In response of this Hollister et al. (2004), suggested that accuracy of landscape structure cannot be determined solely from site-specific assessments. Moreover, many studies (Fritz et al. 2011; Giri et al. 2005) documented, the importance and reliability of area based land cover and change accuracy and agreement between land cover products for the user community.

4.6 The need of harmonized and standard forest definition in Ethiopia

Here it is important to point that, different figures on the extent of forest and deforestation rate was observed for the same period in Ethiopia. This might be due to the fact that the authors probably used different definitions of forest and different tools and methods as described under description of datasets (section 2.2). These figures are confusing for users and decision makers at different levels including for REDD+ forest monitoring actions. In agreement with this several studies (Lund 2006; Romijn et al. 2013) pointed the impact of variation and choice of forest definition from local to global scale based on land cover, forest cover, land use and other properties on estimates of emissions from deforestation and forest degradation. Moreover, Romijn et al. (2013) concluded that the divergence in definition and classification lead to inaccurate and inconsistent forest cover and forest cover change estimates for REDD+ activities, particularly in developing countries like Ethiopia because of they have difficulties in capacity to meet them.

Despite, the importance of the diverse information and the methodologies applied, users and producers of data require consistent as well as interoperable land cover information. Recently, the development of earth observation and its analysis necessitated the harmonization of land cover information from different datasets (Colson et al. 2009). Much effort has been done to improve the comparability and compatibility of land cover information from different datasets (Herold et al. 2008) and harmonization of land cover classification systems and forest related definitions at global scale (FAO 2002a; Herold et al. 2008). In line with this, study by Magdon et al. (2014) on the issue of translating criteria of international forest definitions into remote sensing image analysis, pointed the need of standard forest definition to produce credible and comparable forest cover estimates. This attempt of land cover classification harmonization could lead to ease national as well as regional variation in forest cover and its change estimation, which is important in MRV of REDD+ activities. This is critically important for the success of national as well as regional REDD+ forest monitoring mechanisms in Ethiopia, where quality of information is improved and ambiguities and misunderstandings can be avoided in MRV of REDD+ activities.

Therefore, there is a need to establish a robust, consistent and harmonized forest definition and its monitoring mechanism for REDD+ in Ethiopia. Clear understanding on the determination of national forest emission level (REL) as part of national forest conservation, management and enhancement of carbon stock is important. This is needed mainly to, (i) have one agreed rate of deforestation for the country that can be used by all stakeholders and forestry experts in agreement, (ii) clearly measure role of REDD+ forest monitoring actions reducing ambiguities and misunderstandings between UNFCCC parties, and (iii) reduce measuring and reporting burden by improving the quality of information. In this regard, first Ethiopia has to build consistent forest definitions and monitoring approach at Federal and regional levels. The Federal level should provide coordination and consistent defining framework which includes clear forest definition, land cover classification categories and their mapping standards and monitoring approaches. Regional and local level forest monitoring standards have to complement with Federal guiding principles. Second, designed coordination between REDD+ approaches and national forest inventory of Ethiopia is required, to provide capacities for MRV of REDD+ activities.

Therefore, future research efforts should focus towards finding harmonised and consistent forest and other land cover land cover change related definitions for national, regional and ecological based circumstances in Ethiopia. This is important to reduce ambiguity in forest definitions and enhance effectiveness of national forest monitoring and reporting system for REDD+ activities.

5. Conclusions

This study has attempted to assess the complementarity, accuracy and adaptability of existing global land cover products to the national forest monitoring system of REDD+ in Ethiopia. The findings identified are concluded as follows.

1. Our analysis related to complementarity of global forest cover change datasets to national forest monitoring and MRV of REDD+ showed that forest definitions and methodologies that each dataset employed varies, underpinning the variation of forest cover extent and forest loss rate for the same study period. For example, the national Woody Biomass Inventory and Strategic Planning Project (WBISPP) estimated 313,948 ha (8.6% forest area of 2000) forest loss between 2000 and 2005. While FAO FRA estimated more than 5.1% and Hansen et al. Global Forest Cover Change (GFCC) estimation was very low i.e. 87,497 ha (0.53% forest area of >20% canopy cover). It implies ambiguity in forest definition and then the difficulty to get a reliable and consistent estimation on forest resource of the country. These suggest the need to establish robust, consistent and harmonized forest definition and its monitoring mechanism for MRV of REDD+ in Ethiopia.

2. Five land cover classes in Hansen et al GFCC and FRA Remote Sensing survey (RSS) datasets were assessed for accuracy and agreement. The result showed 61% of Hansen et al. dataset and 80% of FRA RSS has overall accuracy and agreement with reference data. These results are fairly good (with variation in acceptable level), since both dataset employed different classification schemes and different definitions, although they have similarity in data sources and spatial resolution. With regard to classes stable forest has high agreement, consistently exhibiting almost one to one relationship with the reference data and between the datasets. Thus, relatively forest appears to be well classified. Other land cover classes have more consistency of producer and user accuracy in RSS dataset than Hansen et al. GFCC where woodland class is poorly considered. This implies that users of the data at national level to estimate forest area and carbon emission level should be cautious when using area estimates, particularly from the later. If necessary better to examine the similarities and differences in their area of interest and also explore other supplementary sources.

3. The overall adaptability level of each global dataset to national REDD+ forest monitoring approach shows about 68 % for Hansen et al. GFCC and 80% for FAO RSS. Relatively RSS dataset has more conformity to national forest monitoring approach. Other than their accuracy, the main reason for difference might be the RSS is more designed in line with IPCC reporting principles and definitions that may suffice REDD+ forest monitoring mechanisms more. However, although overall assessment indicates more adaptability of RSS dataset, both have their own strengths and limitations to totally comply with REDD+ NFMS. Remote sensing survey dataset is more accurate and comparable, while Hansen et al. GFCC is more complete and transparent in use for measuring, reporting and verification of REDD+ activities at national level, Ethiopia. This makes difficulty to judge which individual dataset is more useful for REDD+ applications. More importantly, it can be helpful to use them in agreement though compensating the drawbacks of one over the other.

References

Achard, F., Brown, S., Brady, M., DeFries, R., Grassi, G., Herold, M., Mollicone, D., Mora, B., Pandey, D., & Souza, C. (2012). GOFC-GOLD REDD Sourcebook, COP-19 release - A sourcebook of methods and procedures for monitoring and reporting anthropogenic greenhouse gas emissions and removals associated with deforestation, gains and losses of carbon stocks in forests remaining forests, and forestation. In. Wageningen: GOFC-GOLD

Anderson, J.R. (1976). *A land use and land cover classification system for use with remote sensor data*: US Government Printing Office

Bernard, F., & Minang, P.A. (2011). *Strengthening Measurement, Reporting and Verification (MRV) for REDD*: International Institute for Sustainable Development

Boschetti, L., Flasse, S.P., & Brivio, P.A. (2004). Analysis of the conflict between omission and commission in low spatial resolution dichotomic thematic products: The Pareto Boundary. *Remote Sensing of Environment, 91*, 280-292

Colson, F., Bogaert, J., Nelson, B., Pinagé, E.R., & Ceulemans, R. (2009). The influence of forest definition on landscape fragmentation assessment in Rondônia, Brazil. *Ecological Indicators, 9*, 1163-1168

Congalton, R.G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment, 37*, 35-46

Congalton, R.G., & Green, K. (2008). *Assessing the accuracy of remotely sensed data: principles and practices*: CRC press

DeFries, R., Achard, F., Brown, S., Herold, M., Murdiyarso, D., & Schlamadinger, B. (2007). Earth observations for estimating greenhouse gas emissions from deforestation in developing countries. *environmental science & policy, 10*, 385-394

Di Gregorio, A. (2005). *Land cover classification system: classification concepts and user manual: LCCS*: Food & Agriculture Org.

FAO (2002a). Expert Meetin g on Harmonizing forest-related definitions for use by various stakeholders. Food and Agricultural Organisation of the United ations (FAO), Rome, Italy 23-25 January 2002. Available at: http://www.fao.org/forestry/15533-0cb816e82c09c14873ce9226dd13910b9.pdf (accessed on 04 July 2014)

FAO (2002b). second expert meeting on harmonizing forest related definitions for use by various stakeholders. FAO – IPCC-CIFOR- IUFRO-UNEP , 11 - 13 september 2002 - Rome, Italy. Available at: http://edepot.wur.nl/144587. (Accessed on 29 June 2014)

FAO (2006). Global Forest Resource Assessment 2005. Progress towards sustainable forest managment. Food and Agricultural Organization of the United Nations, Rome, 2006. www.fao.org/forestry/fra2005/pdf (Accessed on May 14, 2014)

FAO (2009a). The FRA 2010 Remote Sening Survey. An outline of objectives, data, methods and approach.Forest Resources Assessment Working Paper 155 Available at: http://www.fao.org/docrep/012/k7023e/k7023e00.pdf (Accessed on 2 May 2014)

FAO (2009b). The FRA 2010 Remote Sensing Survey. An outline of the Objectives, data, methods and approach

FAO (2010a). *Global Forest Resources Assessment 2010 Main Report*: FAO

FAO (2010b). Global forest resources assessment. Country report Ethiopia. Food and Agriculture Organisation (FAO), Rome, Italy. Available at: http://www.fao.org/docrep/013/al501E/al501e.pdf (accessed on 25 April 2014).

FAO (2012). Food and Agricultural Organization of the United Nations, Global Forest Land-Use Change 1990-2005, FAO Forestry Paper No. 169.Rome, 2012. Available at: http://www.fao.org/docrep/017/i3110e/i3110e.pdf (accessed on 28 April 2014)

FAO (2013). National Forest Monitoring Systems: Monitoring and Measurement, Reporting and Verification (M & MRV) in the context of REDD+ Activities.
www.unredd.net/index.php?option=com_docman&task=doc.(Accessed on 8 May 2014)

FDRE (2011a). Ethiopia's Climate-Resilient Green Economy: Green economy strategy. Addis Ababa: Federal Democratic Republic of Ethiopia. Available at: http://www.uncsd2012.org/content/documents/287CRGE%20Ethiopia%20Green%20Economy_Brochure.pdf. (Accessed on 30 June 2014)

FDRE (2011b). Readiness Preparation Proposal (R-PP) of Federal Democratic Republic of Ethiopia (FDRE), submitted to Forest Carbon Partnership Facility (FCPF). Addis Ababa, Ethiopia.

Foody, G.M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, *80*, 185-201

Friedl, M., Woodcock, C., Gopal, S., Muchoney, D., Strahler, A., & Barker-Schaaf, C. (2000). A note on procedures used for accuracy assessment in land cover maps derived from AVHRR data

Fritz, S., See, L., McCallum, I., Schill, C., Obersteiner, M., van der Velde, M., Boettcher, H., Havlík, P., & Achard, F. (2011). Highlighting continued uncertainty in global land cover maps for the user community. *Environmental Research Letters, 6*, 044005

Giri, C., Zhu, Z., & Reed, B. (2005). A comparative analysis of the Global Land Cover 2000 and MODIS land cover data sets. *Remote Sensing of Environment, 94*, 123-132

Hansen, M., Potapov, P., Moore, R., Hancher, M., Turubanova, S., Tyukavina, A., Thau, D., Stehman, S., Goetz, S., & Loveland, T. (2013a). High-resolution global maps of 21st-century forest cover change. *science, 342*, 850-853

Hansen, M., Potapov, P., Moore, R., Hancher, M., Turubanova, S., Tyukavina, A., Thau, D., Stehman, S., Goetz, S., & Loveland, T. (2013b). Supplementary Materials for High-resolution global maps of 21st-century forest cover change. *science, 342*, 850-853

Herold, M., Mayaux, P., Woodcock, C., Baccini, A., & Schmullius, C. (2008). Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets. *Remote Sensing of Environment, 112*, 2538-2556

Hollister, J.W., Gonzalez, M.L., Paul, J.F., August, P.V., & Copeland, J.L. (2004). Assessing the accuracy of National Land Cover Dataset area estimates at multiple spatial extents. *Photogrammetric Engineering & Remote Sensing*, *70*, 405-414

IPCC (2000). *Land use, land-use change and forestry: a special report of the Intergovernmental Panel on Climate Change*: Cambridge University Press

IPCC (2003). Good practice guidance for land use, land-use change and forestry. In: IPCC National Greenhouse Gases Inventory Programme, Institute for Global Environmental Strategies (IGES) Hayama

IPCC (2006). Draft 2006 Intergovernmental Panel On Climate change (IPCC) Guidelines for National Greenhouse Gas Inventories. Twenty-fifth session, Mauritius, 26-28 April 2006. Available at: www.ipcc.ch/meetings/session25/doc4a4b/vol1.pdf

Ismail, M.H., & Jusoff, K. (2008). Satellite data classification accuracy assessment based from reference dataset. *International Journal of Computer and Information Science and Engineering, 2*, 96-102

Langford, W.T., Gergel, S.E., Dietterich, T.G., & Cohen, W. (2006). Map misclassification can cause large errors in landscape pattern indices: examples from habitat fragmentation. *Ecosystems, 9*, 474-488

Lund, H.G. (2006). *Definitions of forest, deforestation, afforestation, and reforestation*: Forest Information Services. Available at: http://home.comcast.net/~gyde/DEFpaper.htm. Last updated: 28 March 2014 (online accessed on 02 July 2014)

Magdon, P., Fischer, C., Fuchs, H., & Kleinn, C. (2014). Translating criteria of international forest definitions into remote sensing image analysis. *Remote Sensing of Environment, 149*, 252-262

Maniatis, D., & Mollicone, D. (2010). Options for sampling and stratification for national forest inventories to implement REDD+ under the UNFCCC. *Carbon balance and management, 5*, 1-14

McCallum, I., Obersteiner, M., Nilsson, S., & Shvidenko, A. (2006). A spatial comparison of four satellite derived 1km global land cover datasets. *International Journal of Applied Earth Observation and Geoinformation, 8*, 246-255

MoA (2013). Terms of Reference for Developing Capacities for a national Measuring, Monitoring, Reporting and Verification System to support REDD+ participation of Ethiopia. Background, Capacity Assessment and Roadmap. Prepared by Minstry of Agriculture (MoA), Environlmental Protection Agency (EPA), with support from the Norwegian embassy and Wageningen University, Version 4.0

Moges, Y., Eshetu, Z., & Nune, S. (2010). Ethiopian forest resources: Current status and future management options in view of access to carbon finances. *Addis Ababa: Literature Review Prepared for the Ethiopian Climate Research and Networking Groups and United Nations Development Program*

Neeff, T., Von Luepke, H., & Schoene, D. (2006). Choosing a forest definition for the Clean Development Mechanism. *Forests and Climate Change Working Paper (FAO)*

Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., & Wulder, M.A. (2014). Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment, 148*, 42-57

Potapov, P., Hansen, M.C., Gerrand, A., Lindquist, E.J., Pittman, K., Turubanova, S., & Løyche Wilkie, M. (2011). The global Landsat imagery database for the FAO FRA remote sensing survey. *International Journal of Digital Earth, 4*, 2-21

Rindfuss, R.R., Walsh, S.J., Turner, B., Fox, J., & Mishra, V. (2004). Developing a science of land change: challenges and methodological issues. *Proceedings of the National Academy of Sciences of the United States of America, 101*, 13976-13981

Romijn, E., Ainembabazi, J.H., Wijaya, A., Herold, M., Angelsen, A., Verchot, L., & Murdiyarso, D. (2013). Exploring different forest definitions and their impact on developing REDD+ reference emission levels: a case study for Indonesia. *environmental science & policy, 33*, 246-259

Romijn, E., Herold, M., Kooistra, L., Murdiyarso, D., & Verchot, L. (2012). Assessing capacities of non-Annex I countries for national forest monitoring in the context of REDD+. *environmental science & policy, 19*, 33-48

Schoene, D., Killmann, W., von Lüpke, H., & Wilkie, M.L. (2007). *Definitional Issues Related to Reducing Emissions from Deforesation in Developing Countries*: Food and Agriculture Organization of the United Nations

Steven, R., & Brown, J. (2012). Region one Vegetation classification, mapping, inventory and analysis report Flathead National Forest VMap Accuracy Assessment; Version 12 (www.fs.usda.gov/Internet/FSE_DOCUMENTS/stelprdb5363894.pdf). Numbered Report 12-36-v2

Tropek, R., Sedláček, O., Beck, J., Keil, P., Musilová, Z., Šímová, I., & Storch, D. (2014). Comment on "High-resolution global maps of 21st-century forest cover change". *science, 344*, 981-981

UNEP/CBD (2001). Main Theme: Forest Biological Diversity. Report of the Ad Hoc Technical Expert Group on Forest Biological Diversity. Subsidiary Body for Scientific, Technical and Technological Advice, Seventh Meeting, Montreal, 12-16 November 2001.

Available at: www.biodiv.org/programmes/areas/forest/definitions.asp. (Acessed on 22 July 2014)

UNFCCC (2001). The Marrakesh Accords and the Marrakesh Declaration. In, *adopted by the Conference of the Parties at its seventh session*

Vietnam-redd.org (2011). Measurment, Reporting and Verification (MRV) Framework Document. With reference to safeguards information and monitoring of polices and measures under the Viet nam National REDD+ PROGRAM. Vertion 1, Draft 3/September 2011. Available at: http://vietnam-redd.org/Upload/Download/File/MRV_Framework_Document_V1_Draft_3_4046.pdf. (Accessed on 25/10/2014)

Watson, C. (2009). Forest carbon accounting: overview and principles. *Forest carbon accounting: overview and principles*

Watson, R.T., Noble, I.R., Bolin, B., Ravindranath, N., Verardo, D.J., & Dokken, D.J. (2000). *Land use, land-use change and forestry: a special report of the Intergovernmental Panel on Climate Change*: Cambridge University Press

WBISPP (2004). Woody Biomass Inventory and strategic planning Project (WBISPP). Phase 2, Termonal Report. Addis Ababa. . *Forest Resources of*

WHRC (2013). First detailed map of global forest change is focus of Science magazine report. Press Release. Available at: http://whrc.org/news/pressroom/index.html#sthash.2wLJ4UZE.dpuf. In: Woods Hole Research Center

Wulder, M.A., Franklin, S.E., White, J.C., Linke, J., & Magnussen, S. (2006). An accuracy assessment framework for large-area land cover classification products derived from medium-resolution satellite data. *International Journal of Remote Sensing, 27,* 663-683

www.fcmcglobal.org REDD+ Measurment, Reporting and Verification (MRV) manual. Summary for policy makers (http://www.fcmcglobal.org/documents/mrvmanual/MRV_Manual_Policymakers_Summary.pdf) (Accessed on May 14, 2014)

Zhu, Z., Yang, L., Stehman, S.V., & Czaplewski, R.L. (2000). Accuracy assessment for the US Geological Survey regional land-cover mapping program: New York and New Jersey region. *Photogrammetric Engineering and Remote Sensing, 66*, 1425-1438

Zomer, R.J., Trabucco, A., Verchot, L.V., & Muys, B. (2008). Land area eligible for afforestation and reforestation within the Clean Development Mechanism: a global analysis of the impact of forest definition. *Mitigation and Adaptation Strategies for Global Change, 13*, 219-239

Appendices

Appendix I List of Tables

Table 13. Some major examples of misclassification (disagreement, ommission and commission errors) observed from Hanasen et al. and FAO RSS data sets in Ethiopia

Sample point	Latitude (N)	Longitude (E)	Hansen et al.	FAO RSS	GEI	Remark
22	4°57'57.87"	35°57'48.45"	NF	NF	OWL	No gain is observed in both
396	5° 0'0.72"	36° 2'29.56"	Forest	Forest	Forest	OE in Hansen et al.
26	5°59'0.80"	35°58'30.09"	OWL	OWL	NF	No forest loss is observed in both
62	11°59'22.17"	37° 0'51.51"	OWL	NF	NF	No any tree cover
31	7°59'9.34"	35°58'33.74"	Forest	NF	NF	CE Hansen et al.
79	8°57'44.29"	38° 2'20.08"	Forest	OWL	OWL	CE in Hansen
41	11°58'48.50"	35°58'23.01"	Forest	Forest	Forest	Difference in labelling of forest in RSS for the sample sites
59	11° 1'38.70"	37° 1'39.71"	Forest	OWL	Forest	
86	10°59'3.29"	37°58'38.30"	Forest	OWL	Forest	Forest misclassified as OWL in RSS
112	7° 0'38.59"	39°58'32.37"	NF	NF_Bare land	NF_ urban	RSS misclassified as Bare land
94	6°57'29.38"	38°59'47.21"	OWL	NF	NF	Hansen et al. misclassification
134	8°57'31.25"	41° 0'17.69"	Forest	OWL	NF	It is settlement with agricultural tree cover
138	9°59'39.73"	40°58'38.40"	NF	OWL	Forest	Both disagree also no gain is recorded
127	12° 0'9.20"	39°59'46.94"	NF	FL	FL	Loss disagreement b/n Hansen and others
188	7°28'17.99"	36°29'17.68"	FL	-	FL	Loss Hansen with CE
199	6°32'33.42"	38°29'11.60"	FL	-	FL	Loss Hansen CE
208	9°31'45.88"	39°28'4.98"	FL	-	NF	No change is observed
320	12°45'50.63"	37°34'13.00"	FG	-	FG	Mach more OE
342	9°50'5.85"	39°45'18.20"	FG	-	FG	Mach more OE

Adapted from Tropek et al. (2014)

NB. NF represent non forest class and similarly OWL=other woodland, FL=forest loss, FG=forest gain, OE=omission error, CE=commission error and SNNPR=Sothern Nations and Nationalities of peoples Region (southern Nations)

Table 14. Number of tiles (at 1° x1° intersection) having a reference data from Google earth imagery between 2000 and 2006.

year	2000	2001	2002	2003	2004	2005	2006	Total
No of tiles	0	1	1	4	8	11	26	51

NB. Some tiles have reference data in 2007. This was also used to identify whether the area experienced forest cover change or not.

Table 15 Conceptual Frame work in assigning sample sights to land cover and land cover change classes

Class		2000	2001-2005	Decision
Change		F	NF	FL
		F	OWL	FL
		NF	F	FG
		OWL	F	FG
No	change	F	F	F
		OWL	OWL	OWL
		NF	NF	NF

Time

This framework was employed for FRA RSS. Change or no change was assessed from 2000-2005. If time series imagery is limited in these periods imagery of 2006 and 2007 were checked also for consistency.

Class		2000	2001-2005	2006-2012	Decision
Change		F	NF	NF	FL
		F	OWL	OWL	FL
		NF	F	F	FG
		OWL	F	F	FG
		NF	F	NF	NF
No	change	F	F	F	F
		OWL	OWL	OWL	OWL
		NF	NF	NF	NF

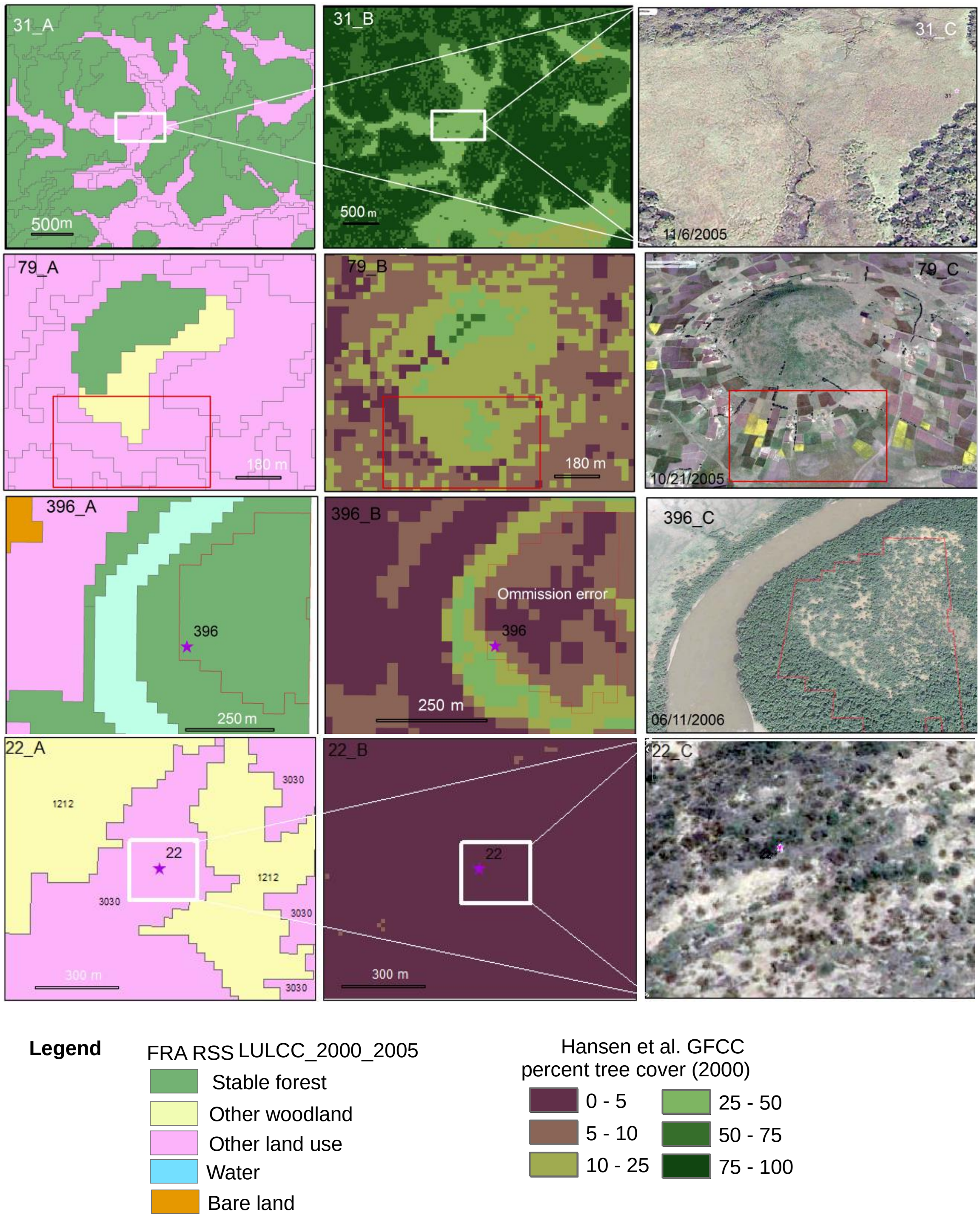
Time 1 Time 2

This framework is developed for assignment Hansen et al. in to assign the five land cover classes. Time 1 indicates the assessment period, while time 2 is used only to check the consistency of data obtained in the study period (2000-2005).

Legend	
FL	Forest loss
FG	Forest gain
F	Forest
OWL	Other woodland
NF	non-forest

In the no change classes namely stable forest, Other Woodland Land (OWL) and stable non-forest throughout the time series, decision was made if there was no change of forest to non-forest and Vice versa between the time period or later time. The sample point under assessment was labelled under one of the no change classed following the above framework and the methodology on Figure 2. In change class, the sample point is labelled as forest loss if there was change of forest/other woodland covered area to non-forest class or open woodland in the time period. It was labelled as forest gain if change from non-forest class to forest or closed non-agricultural wood land covered is observed (Figure 2 and Table 15).

Appendix II List of Figures



NB: Each map is labelled based on Sample site number E.g. 31_A means the fierst map for sample point 31

Figure 8. Examples of Errors of commission and Errors of omission observed for no change classes

From the figure 5 samples 31 and 79 indicate errors of commission and sample 22 and sample 396 show error of omission. The white box in sample 31 shows that there was true classification to non-forest

(Other land use) in RSS and misclassification of non-forest area to forest with 25-50% canopy cover in Hansen data set. Similarly, sample 79 showed that relatively RSS is correctly classified as woodland where the sample point fall. While there was commission of non-forest with agricultural land areas without any tree cover as forest with 10 – 50%. Zooming to the area particularly south of the sample point (sample 79_C) we discovered that the area is settlement where agricultural activity is practiced and no more tree cover is observed in 2005, 2009 and till 2012. Sample 22 shows omission error of woodland class in both data sets, where both data sets reported stable non forest without any gain, while the area is really covered with dense woodland on 05/07/2006 (22_C) and no agricultural and habitation activity is there. Sample 396 showed that the area is forest with relatively less dense tree cover in the middle of the red polygon (396_C) evidencing the true classification of RSS (396_A) and negating the omission of forest class (396_B).

Moreover, form sample 22 and 396 and others (not showed here), we can deduce that Hansen et al. has more potential to pin point forest covered areas with more than 50% tree canopy cover. In such case we didn't observed misclassification of forest to other land cover classes and vice versa. Where there was less than 50% tree canopy cover most agricultural non tree cover areas are committed as forest, might be due to the data used is from growing season (Hansen 2013 supplementary) and most woodland covered areas of the country, particularly in south east part of Oromia Region and South west of Somali region are committed as non-forest with less than 5% tree canopy cover in Hansen et al.

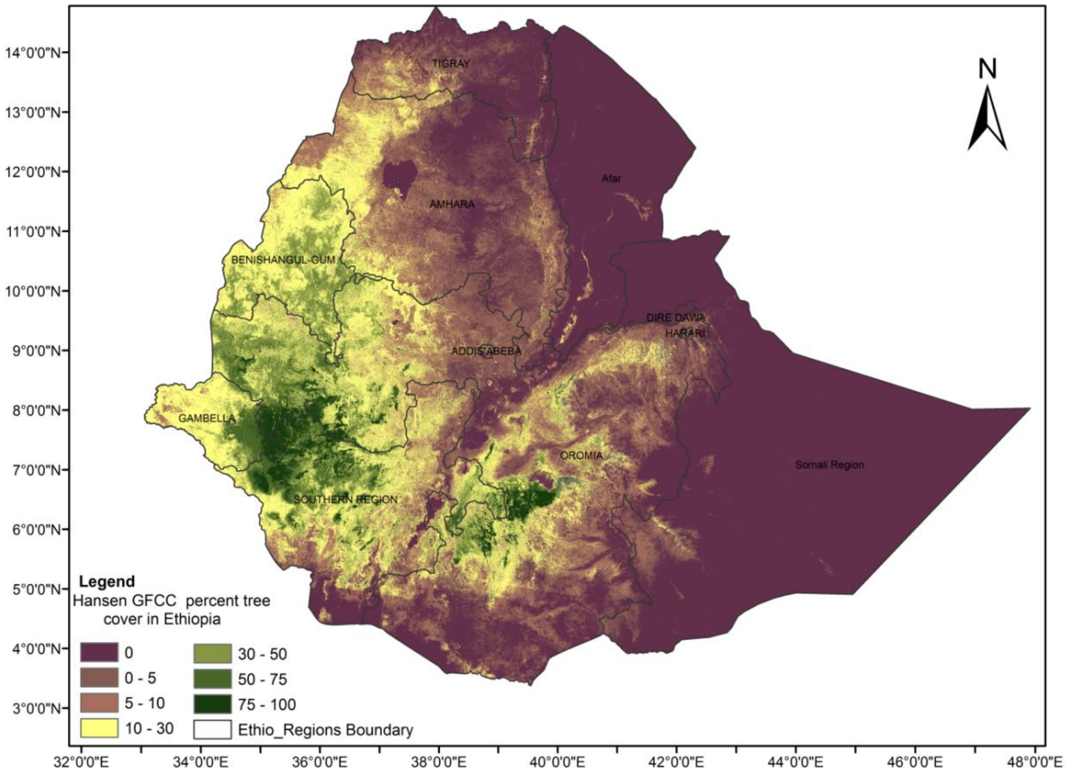


Figure 9. Forest extent of Ethiopia based on Hansen et al GFCC percent tree canopy cover.