IJESKA contribution to science

International Journal of Earth Sciences Knowledge and Applications
journal homepage: http://www.ijeska.com/index.php/ijeska

Research Article e-ISSN: 2687-5993

Improvement of Forest Change Maps Based on Normalized Difference Vegetation Index (NDVI)

Anwar Sidahmed[1], Rashid Jalal[2], Elyas Ahmed[3], Rémi d'Annunzio[2], Marekae Sandker[2]

[1]*Remote sensing and Seismology Authority, Khartoum, Sudan*
[2]*University of Khartoum Faculty of Forestry, Khartoum, Sudan*
[3]*Food and Agriculture Organization, Rome, Italy*

INFORMATION

Article history
Received 23 November 2022
Revised 15 December 2022
Accepted 15 December 2022

Keywords
Normalized Difference Vegetation Index (NDVI)
Accuracy assessment
Google Earth
Response design
Object-based
Image analysis

Contact
*Anwar Sidahmed
nanosid25@gmail.com

ABSTRACT

Normalized Difference Vegetation Index (NDVI) is one of the most widely used numerical indicator that uses the visible bands (VIS) and near-infrared bands (NIR) of the electromagnetic spectrum, its use as an indicator for vegetation and vegetation health based on how plants reflect certain ranges of the electromagnetic spectrum. The development of applications such as Google Earth and Microsoft Bing Maps, very high resolution (VHR) satellite imagery can be viewed over many parts of the world. The study used already created change maps based on Landsat and Aster and estimated NDVI to improve the accuracy of the data and estimate the accuracy assessment of these maps using available VHR in Google Earth. The area of the classes changed after the improvement on these maps using NDVI and the accuracy of the change maps was 0.83.

1. Introduction

Forests worldwide are critically important habitats in terms of their biological diversity, ecological functions and the ecological services provide habitat for more than half of the world's species (Groombridge and Jenkins, 2000), they regulate local and global climate, ameliorate weather events. Forests provide a number of ecosystem services including carbon dioxide capturing and storing, and are an important component of the global carbon cycle, which could add multiple benefits to climate change mitigation. The human impacts on forest lands are still very great and increasing, due to the growth of the population which is directly linked to urbanization, which is generally responsible for depletion of nearby forest ecosystems.

In addition to pressures on forest ecosystems, changes in forest cover have had important effects on biodiversity, soil conservation, water quantity and quality and world climate (Iida and Nakashizuka, 1995; Johnson et al., 1997; Chen et al., 2001; Dupouey et al., 2002; Upadhyay et al., 2007; Liu et al., 2006).

Remote sensing plays a critical role in the forest mapping and monitoring of disturbances and changes in forest cover. Different types of satellite sensors (i.e., optical, RADAR LiDAR and hyperspectral at varying spatial and temporal resolutions) play different and complementary roles in forest monitoring. With Landsat as an optical remote sensing archived data in including a time series of more than 40 years, it has become possible to monitor long-term changes in forests at a higher resolution (Sexton et al., 2013; Hansen et al., 2013).

More recently, with changes in Web 2.0 technology (Hudson-Smith et al., 2009) and the development of applications such as Google Earth and Microsoft Bing Maps, very high resolution (VHR) satellite imagery can be viewed over many parts of the world. Moreover, the land area now covered by VHR imagery has also grown dramatically over recent years (Lesiv et al., 2018b). This opens up the possibility to visually identify land cover and land use features, as well as the structure of forests. For the purpose of this paper, we define VHR imagery as having a spatial resolution of less than 2 m, while high-resolution (HR) imagery refers to the resolutions of Landsat (30 m) or Sentinel 2 (10 m).

Accuracy and consistency of forest area and forest area change information is important reporting requirement for countries in the context of accessing results based payments for REDD+ (FAO, 2016). Decision 4/CP.15, paragraph 1(d) (ii) asks Parties to provide estimates that are transparent, consistent, accurate (as far as possible), and that reduce uncertainties, taking into account national capabilities and capacities. This study aims to use Landsat 8 NDVI for improving the accuracy of forest cover change maps and then estimate the accuracy of these maps.

2. Methodology

In this study forest cover maps were created for 2006 using ASTER (15-meter spatial resolution) and 2010, 2014 and 2018 using Landsat images (30-meter spatial resolution) were used. The images were downloaded from United State Geological Survey (USGS) www.usgs.gov/landsat during the dry and wet seasons, with maximum cloud cover of 30%.

Aster images were already combined on the website, the bands used for Landsat 7 were 4, 3, 2, and for Landsat 8 the bands were 5, 4, 3. For the creation of individual Forest cover map Global Land Cover Network (FAO/GLCN) approach was followed (GLCN/FAO) (http://www.fao.org/geospatial/projects/detail/en/c/1035672). Each single image was processed, interpreted, validated using available very high resolution images from the Bing map in QGIS.

2.1. Image segmentation

Object-based image analysis (OBIA) approach was used for image segmentation, in which objects were defined by spectral, textural and border properties. The resulted vector layer of objects (i.e., image segments) represent regions with similar pixel values with respect to some characteristic or computed property such as colour, intensity or texture and pattern. Segmentation processing was done using eCognition, with a minimum mapping unit (MMU) of 0.4 hectares (ha) based on the Sudan national forest definition.

Table 1. Sample size allocation to strata

(A) Map class	Number of samples		
	(B) Proportional	(C) Adjusted	(D) Reference data included in analysis*
Gain_06_10	9	100	62
Gain_10_14	6	100	68
Gain_14_18	9	100	72
Loss_06_10	27	100	87
Loss_10_14	8	100	72
Loss_14_18	16	100	69
Stable forest overall	265	265	200
Stable non-forest overall	634	634	510
Total	**974**	**1499**	**1140**

** 359 samples were excluded from analysis due to unavailability of suitable image and low confidence in interpretation*

Objects smaller than 0.4 ha were merged to comply with the defined requirements for MMU. Then overlapping areas were corrected and the layer was made ready for visual interpretation. Because of the difference in images resolutions (Landsat 30m and ASTER 15m) different scales were applied for segmentation suitable to each resolution in the segmentation process.

However, it is not excluded that these different resolutions could have an impact on the map areas and statistics, even if the above measures are expected to result in these differences being minimal. In case some difference remains in the map areas, these are corrected for with the spatial assessment units (MMU).

The image segments developed were used as the basic unit of classification (labelling and assigning each segment to the target land cover class. The labels were manually assigned to each polygon (i.e., image object) during the visual interpretation using LCCS 3 Basic Coder plugin in QGIS.

2.2. Forest change map

The forest maps of 2006, 2010, 2014 and 2018 were overlaid to obtain a change map in which each polygon contains: - Forest / non-forest class in 2006 - Forest / non-forest class in 2010 - Forest /non-forest class in 2014 - Forest / non-forest class in 2018 - Area in hectares (ha).

2.3. Improvement of change map

Quality of land cover interpretation by the photo interpreters were under continuous checking (low quality interpretation sent back to photo interpreters for reinterpretation) by more experienced experts. Such quality checking of land cover interpretation was an integral part while developing the individual land cover map. After the submission in January and before starting the accuracy assessment, the change map was further checked for potential areas of improvement. This improvement work involved identifying and checking potential areas of misclassification. Polygons for checking were identified based on NDVI analysis for the months of January and February of the mapping years based on Landsat imagery. January and February were considered to separate the effect of grass and crops from forest to the extent possible considering the phenology and cropping pattern in the area. For each of the map polygons median NDVI was calculated.

In the first round of checking, polygons with area greater than two hectares and NDVI values lower or higher than the threshold for forest or non-forest class, respectively, in all years were selected for checking. About 0.65 million ha of area (about 5% of total area) were checked and reclassified where deemed necessary using visual interpretation at this stage2.5. Accuracy assessment of change map. The objectives of the map accuracy assessment were to assess the accuracy of the change map from 2006.

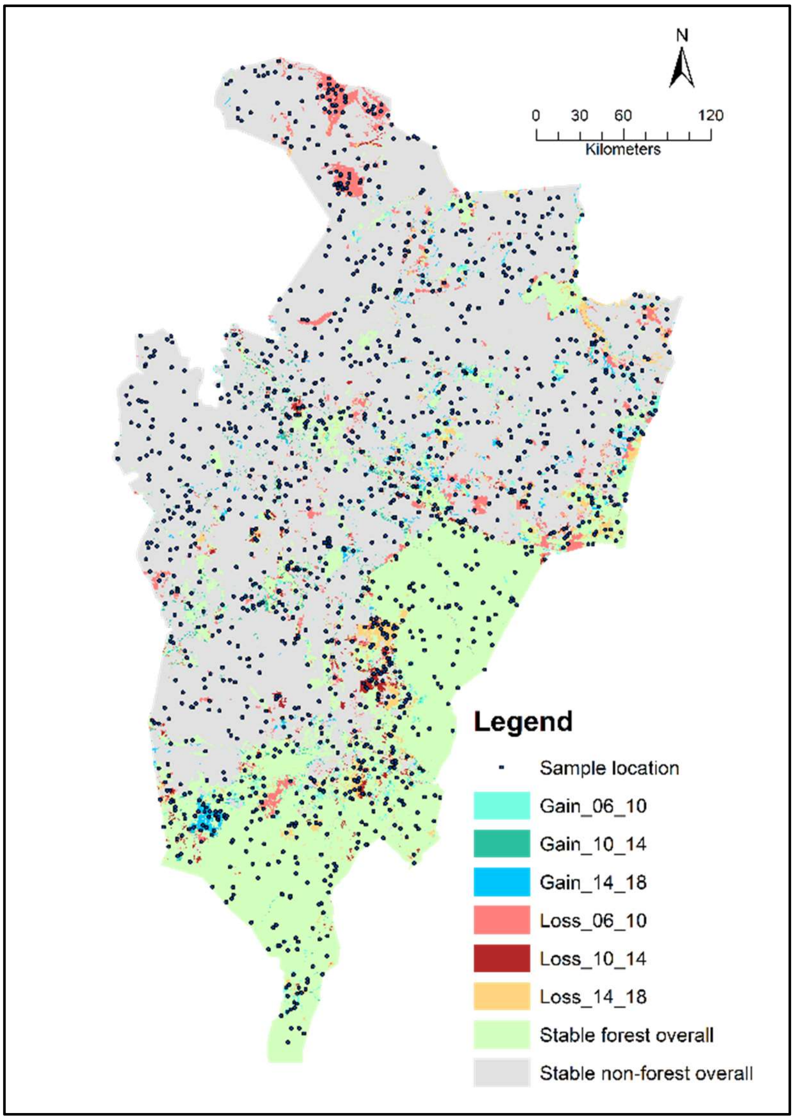

Fig. 1. Distribution of samples over change classes

2.4. Sampling design

A probability sampling design i.e., stratified random sampling was implemented. The classes of change map were used to construct strata. The following equations (Cochran, 1977) were used to calculate an adequate overall sample size (n) for stratified random sampling.

$$n = \frac{(\sum W_i S_i)^2}{[S(\hat{O})]^2 + (\frac{1}{N}) \sum W_i S_i^{\,2}} \quad (1)$$

$$S_i = \sqrt{EUA_i * (1 - EUA_i)} \quad (2)$$

where; i is activity class, N is number of units in the area of interest, $S(\hat{O})$is the standard error of the estimated overall accuracy, W_i is the mapped proportion of area of class i, $S_i$is the standard deviation of stratum i, and $EUA_i$is expected user accuracy of stratum i.

The standard error of the estimated overall accuracy ($S(\hat{O})$) was set to 0.01. Stable and rare classes (i.e., change classes) are expected to have high and low user accuracy, respectively (FAO, 2016). Accordingly, for stable classes (i.e., stable forest and stable non-forest) expected user accuracy was set to 0.9 and for change classes (i.e., gain and loss) this was set to 0.7. Overall minimum sample size was found to be 974.

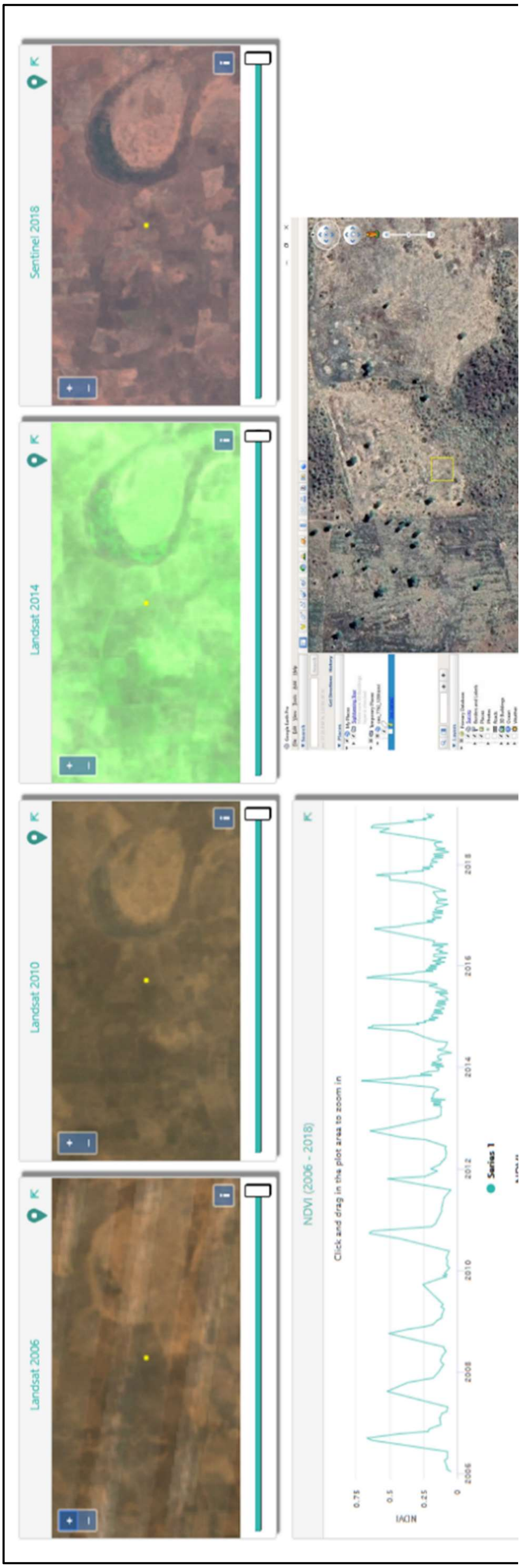

Fig. 2. Reference data collection employing Collect Earth Online and Google Earth

The minimum sample size was distributed proportionally among the classes, with an increase of minimum sample size of at least 100 samples per class to ensure that rare change classes were sufficiently sampled. This resulted in total 1499 samples for which reference data was to be collected.

Table 1 shows the allocation of sample size to strata along with the distribution reference data included in analysis (discussed in response design). Column B presents the proportional distribution of minimum sample size. Column C presents the allocated samples with an increase of minimum sample to 100. Column D presents the distribution of reference data included in the analysis. **Hata! Başvuru kaynağı bulunamadı.** shows the distribution of allocated samples over map classes.

Landsat 5 was used for 2010. Landsat 8 was used for 2014 and 2018. NDVI for 2006 was not possible to be calculated due unavailability of suitable Landsat images. Landsat 5 does not have coverage in 2006 and Landsat 7 was not considered due to data gap resulting from scan line corrector (SLC) failure.

2.4.1. Response design

The response design encompasses all steps of the protocol that lead to a decision regarding agreement or disagreement of the reference and map classifications (Olofsson et al., 2014). The four major features of the response design (i.e., the spatial assessment unit, the sources of information used to determine the reference classification, the labeling protocol for the reference classification, and a definition of agreement) are discussed in the following subsections.

2.4.2. Spatial assessment unit

Pixels, blocks of pixels and polygons are all potentially viable spatial assessment units for conducting an accuracy assessment. Stehman and Wickham (2011) discuss several challenges associated with implementation and analysis of block and polygon-based accuracy assessment. Block of pixels and polygons are less likely to be homogeneous, so the response design and analysis protocols are more complex to account for within-unit heterogeneity. Pixel-based assessment (assuming within-unit homogeneity), on the other hand, can easily accommodate sampling designs employing strata. A traditional error matrix analysis can be readily applied to the case of homogeneous assessment units.

Moreover, for an area-based accuracy assessment, preservation of the areas of agreement and disagreement is one of the critical requirements which is comparatively well preserved by smallest possible spatial assessment unit. Considering these, 30m by 30m spatial assessment unit was used for reference data collection. Spatial units were randomly allocated to strata according to the adjusted sample size (as shown in **Hata! Başvuru kaynağı bulunamadı.**) using point sampling protocol.

2.4.3. Sources of reference data

The two ways to ensure better quality of reference classification than the map classification (Olofsson et al., 2014) are to ensure that the reference source is of higher quality (e.g., higher resolution satellite imagery) than what was used to create the map classification and in case of using the same source material for both the map and reference classifications (e.g., both classifications rely on Landsat data), to ensure that the process to create the reference classification is more accurate than the process used to create the classification being evaluated.

Table 2. Comparison of map areas before and after improvement of change map

Map Class	Area (ha)	
	First Forest Change	Improved Forest Change
Gain (2006 – 2010)	150,923	127,586
Gain (2010 – 2014)	124,327	86,727
Gain (2014 – 2018)	125,324	130,865
Loss (2006 – 2010)	362,543	383,797
Loss (2010 – 2014)	96,385	120,694
Loss (2014 – 2018)	264,139	230,709
Stable forest overall	4,059,585	3,657,541
Stable non-forest overall	8,308,567	8,753,874
Total	**13,491,793**	**13,491,793**

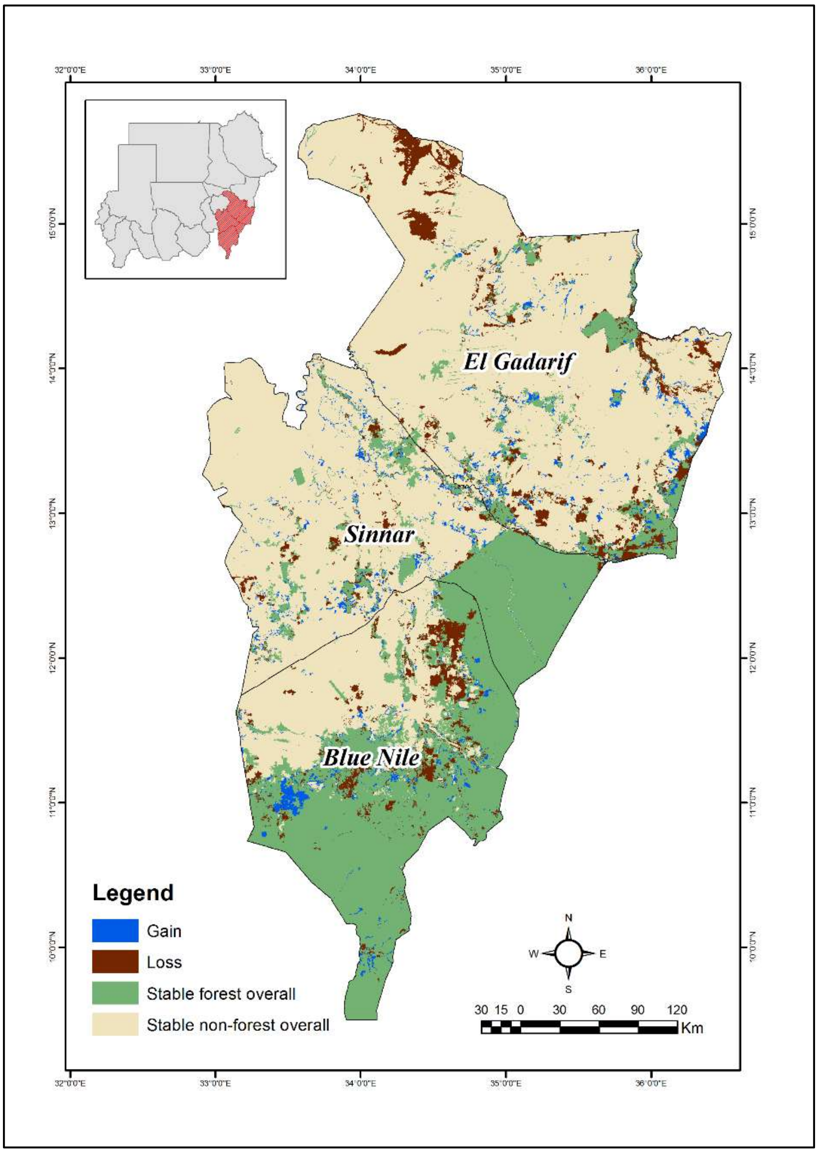

Fig. 3. Forest change map after improvement

Potential sources of reference classification can be ground visits to the sample locations or the use of aerial photography or satellite imagery. Practical considerations (e.g., costs) were influencing factors in the selection of sources of reference data for the accuracy assessment of change map. Collect Earth Online (CEO) platform (Saah et al., 2019) was used for collecting reference data (**Hata! Başvuru kaynağı bulunamadı.**). CEO is an open-source, web-based, crowd-sourcing technology for Earth Science analyses allowing users to collect reference data using a variety of imagery resources and processing capabilities. Very high-resolution imagery available through Google Earth (linked with CEO) historical imagery were used as primary source of information for reference classification. In addition, available images of Landsat (for 2006, 2010 and 2014) and Sentinel 2 (for 2018) and NDVI time series from 2006 to 2018 were used to facilitate reference classification

2.4.4. Labeling protocol

Each spatial assessment unit was assigned either forest or non-forest class for the years of 2006, 2010, 2014 and 2018 based on visual interpretation of available high-resolution image and local knowledge of the analysts. Availability of high-resolution image for specific year was a major concern for collecting reference data. In case of unavailability of high-resolution image for a specific year, images (if available) for years immediate before or after were used for interpretation***2.***
3.

If a spatial unit was found impure (i.e., representing an area of more than one class), the majority class was assigned. If a spatial unit could not be classified due to lack of suitable images, local knowledge, etc., the unit was noted as of no confidence, and hence excluded from analysis. In total reference data from 1140 spatial units (**Hata! Başvuru kaynağı bulunamadı.**) were included in the analysis.

2.4.5. Defining agreement

Consideration of high-resolution images from the years other than the mapping years for reference data collection has implications particularly for gains and losses which were disaggregated in three time periods (i.e., Gain 2006-10, Gain 2010-14, Gain 2014-18, Loss 2006-10, Loss 2010-14 and Loss 2014-18) in the change map – gain/loss of one-time period may fall in gain/loss in other time period resulting increase of omission/commission errors. Taking this into consideration, gains and losses were aggregated for the whole FRL time period of 2006 to 2018.

The reference data collected through Collect Earth Online were first translated into the map classes. Periodic gains/losses of both map and reference data were then aggregated for the whole FRL time period as follows - Gain 2006-10, Gain 2010-14 and Gain 2014-18 were aggregated as Gain; Loss 2006-10, Loss 2010-14 and Loss 2014-18 were aggregated as Loss. Agreement between reference and map data was then defined as when the respective classes matched.

3. Results

3.1. Forest change maps

Polygons with no change were classified as stable; stable forest overall, stable non-forest overall, loss (forest converted to non-forest) and gain (non-forest converted to forest) for three periods. A polygon with loss or gain in only one-time period was classified as loss or gain in that time period. The ears of first forest change map and the forest area changes after improvement using Mean of NDVI and Very high resolution images in Google Earth are presented in Table 2 and Fig. 3.

Table 3. Confusion matrix

Reference class / Map class	Gain	Loss	Stable forest overall	Stable non-forest overall
Gain	**20**	14	67	101
Loss	7	**43**	68	110
Stable forest overall	8	11	**158**	23
Stable non-forest overall	7	12	17	**474**

Table 4. Accuracy and area estimates

Class	Accuracy			Area (ha)			
	Producer's accuracy	User's accuracy	Overall Accuracy	Map area	Stratified area estimate	Standard error	95% confidence interval
Gain	0.106	0.099	0.83	345,178	323,201	68,871	134,986
Loss	0.243	0.189		735,200	569,718	85,765	168,099
Stable forest overall	0.822	0.790		3,657,541	3,515,013	128,970	252,782
Stable non-forest overall	0.896	0.929		8,753,874	9,083,861	132,142	258,998

3.2. Creating confusion matrix

The confusion matrix or error matrix - cross-tabulation of the class labels allocated by the classification of the map data against the reference data for the sample sites (Olofsson et al., 2014) is presented in Table 3. The diagonal cells (in bold) represent the correct classifications where map and reference data agree in their classification. All cells off-diagonal show omission and commission errors.

3.3. Estimating accuracies and bias-corrected areas

Three measures of accuracy (i.e., overall, producer's and user's accuracy) and adjusted areas were estimated using the formula provided by Olofsson et al. (2014) and was done in R based on scripts developed by FAO. Areas are estimated at State level using the class specific adjustment ratio of stratified area to map area of the whole region. Results are presented in Table 4.

4. Conclusions

NDVI thresholds was used for reclassification of some polygons in the forest change maps based on Landsat and Aster satellite data in this study. It was obvious that the area of all classes was changed and the accuracy was high and acceptable. The VHRI in google earth covered the all area of the study area and are available and free and they can easily have used for verification and accuracy assessment of the image classification instead of the field check as it is expensive and time consuming.

References

Cochran, W.G., 1977. Sampling Techniques (thrid edition ed.). New York: John Willey & Sons.

FAO, 2016. Map Accuracy Assessment and Area Estimation: A Practical Guide. Retrieved from Rome.

GFOI, 2016. Integration of remote-sensing and ground-based observations for estimation of emissions and removals of greenhouse gases in forests: Methods and Guidance from the Global Forest Observations Initiative, Edition 2.0, Food and Agriculture Organization, Rome.

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chını, L., Justıce, C.O., Townshend, J.R.G., 2013. High-resolution global maps of 21st-century forest cover change. Science 342 (6160), 850-853. https://doi.org/10.1126/science.1244693.

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.

Saah, D., Johnson, G., Ashmall, B., Tondapu, G., Tenneson, K., Patterson, M., Poortinga, A., Markert, K., Quyen, N.H., Aung, K.S., Schlichting, L., Matin, M., Uddin, K., Aryal, R.R., Dilger, J., Ellenburg, W.L., Flores-Anderson, A.I., Wiell, D., Lindquist, E., Goldstein, J., Chishtie, F., 2019. Collect Earth: An online tool for systematic reference data collection in land cover and use applications. Environmental Modelling & Software 118, 166-171.

Sexton, J.O., Song ,X-P., Feng, M., Noojipady, P., Anand, A., Huang, C., Kim, D-H., Collins, K.M., Channan, S., DiMiceli, C., Townshend, J.R., 2013. Global, 30-m resolution continuous fields of tree cover: landsat- based rescaling of MODIS vegetation continuous fields with lidar-based estimates of error. International Journal of Digital Earth 6 (5), 427-448. https://doi.org/10.1080/17538947.2013.786146.

Stehman, S.V., Wickham, J.D., 2011. Pixels, blocks of pixels, and polygons: Choosing a spatial unit for thematic accuracy assessment. Remote Sensing of Environment 115 (12), 3044-3055. https://doi.org/10.1016/j.rse.2011.06.007.