# Assessment of the accuracy of University of Maryland (Hansen et al.) Forest Loss Data in 2 ICF project areas – component of a project that tested an ICF indicator methodology

**FINAL REPORT, July 2015**

**Authors:**

Edward Mitchard, The University of Edinburgh

Karin Viergever, Ecometrica

Veronique Morel , Ecometrica

Richard Tipper, Ecometrica

**Acknowledgements:**

This study forms a key component of a larger project that tested the ICF Indicator Methodology, funded by the European Space Agency (ESRIN Contract No.4000112345/14/I-NB: Earth Observation Support for Assessing the Performance of UK government's ICF Forest Projects), with additional support from NERC (Innovation Voucher Scheme).

# Table of contents

# 1 Introduction

This document describes the methods and results of an exercise to estimate the accuracy of the University of Maryland (UMD, Hansen et al.) Forest Loss Data. This report is was produced as one of the project deliverables for "*Earth Observation Support for Assessing the Performance of UK Government's ICF Forest Projects*" funded by ESA and contracted to Ecometrica.

In our report on the ICF Hectares Indicator in May 2014 [1], we concluded that free-of-charge, standardised global deforestation products derived from satellite data would be of use in monitoring the past and current performance of ICF projects. In particular we thought that such products could enable a low-cost, automated and consistent means to provide annual estimates of actual deforestation in ICF project areas, in order to form one side of the calculation of Key Performance Indicator #8, the Hectares Indicator.

The only global deforestation dataset of a sufficient resolution currently available is the Global Forest Loss dataset described in a paper in the journal Science by Matthew Hansen of the University of Maryland (UMD) and colleagues in 2013[2]. The data maps annual forest loss per year between 2001 and 2013 at a spatial resolution of 30m and is freely available to view and download via the University of Maryland data portal[3]. The data are also available via the Global Forest Watch (GFW)[4] online forest monitoring and alert system, although resampled to a coarser resolution of approximately 90m. The data were produced from a time-series analysis of over 655 000 Landsat 8, ETM+ and TM images from 2001 through 2013, led by scientists at the University of Maryland but with significant support from Google, with the actual product produced using their Google Earth Engine.

While this UMD dataset is a major advancement in the understanding and quantification of global forest change research and conservation planning, a thorough understanding of its key limitations as well as uncertainties and inaccuracies within specific forest types and different canopy densities is vital in order to ensure its appropriate use for specific applications and in local contexts. This study aims to estimate whether significant areas of deforestation are missed or incorrectly detected and mapped by the UMD forest loss per year product within Brazilian cerrado vegetation and Ghana high forest, using both a visual interpretation and quantitative analysis of multi-date very high resolution (5 m) RapidEye and SPOT satellite data. It is important to note that this accuracy assessment does not aim to quantify errors of omission and commission strictly according to the Hansen et al study definition of forest cover and loss, but rather to measure the performance of the product for the purposes of assessing ICF forest conservation and management projects within varying landscapes and forest types.

While some accuracy assessment was done in the original paper (Hansen et al. 2013), finding accuracy greater than 90 % for its forest/non-forest delineation when tested against independent test datasets, such tests do not provide a robust assessment of its use for detecting change for the ICF's purposes. There are several reasons in particular for necessitating an independent assessment of accuracy specifically targeted at the type of change normal in ICF projects:

---

[1] Tipper et al. (2014) The ICF Hectares Indicator: a review and suggested improvements to the indicator methodology (Download)

[2] Hansen, M. C., P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, S. V. Stehman, S. J. Goetz, T. R. Loveland, A. Kommareddy, A. Egorov, L. Chini, C. O. Justice, and J. R. G. Townshend. 2013. "High-Resolution Global Maps of 21st-Century Forest Cover Change." Science 342: 850–53.

[3] http://earthenginepartners.appspot.com/science-2013-global-forest.

[4] http://data.globalforestwatch.org/datasets/93ecbfa0542c42fdaa8454fa42a6cc27

- There have been questions over the accuracy of the UMD product, stating that the internal accuracy assessment in the Hansen et al. (2013) paper overstates its accuracy. These concerns are typified in a Comment on the paper by Tropek and colleagues[5], though a Response by Hansen et al.[6] stated that many of the supposed errors were caused by differences in the definition of forest and forest loss. These questions mean an independent accuracy assessment in landscapes relevant to the ICF is necessary.

- For the purpose of the UMD data, "tree cover" is defined as all vegetation taller than 5 meters in height. Under this structural definition, plantations such as oil palm, soy beans, tea and monoculture crops or trees are included as "forests", though they do not have the carbon or biodiversity value of natural forest, and indeed may not be considered forest according to natural definitions. Furthermore, plantation harvesting and management as well as fire and storm damage are interpreted as forest loss within the dataset, and Hansen et al emphasise that "loss" does not always equate to deforestation6. It should be possible to use other datasets to mask out areas not considered as forest by the ICF project in question, but as many ICF landscapes include large areas of plantation or agricultural activities, and the definition of 'forest' differs dramatically between countries, this creates a requirement to test whether the product remains useful.

- The resolution of the Hansen et al. (UMD) forest loss product, at 30 m, should be suitable in most cases to see the majority of deforested (and majorly degraded) areas. However, in some areas it may be that small areas of deforestation dominate, and thus the UMD forest loss product may underestimate total loss. In this study we included an example of Ghana to test the effect of resolution, as we know that much forest loss in Ghana occurs at a very small scale.

- In the Hansen et al. study, errors were only reported in terms of user's and producer's accuracy, not errors of omission and commission. Converting between these numbers is non-trivial as neither classification accuracy nor deforestation processes are randomly distributed in space. In order to confirm the suitability of the UMD data for use in calculating the Hectares Indicator it is important to calculate errors of omission and commission on an annual basis.

- Accuracy of the UMD forest loss product appears to have been assessed mostly with regards to changes from tall and closed canopy tropical forest to non-forest. This does not represent the type of changes occurring in many ICF landscapes however, with many involving smaller-scale changes in less high biomass forest types. It is important therefore to test in real project ecosystems, in particular those in woodlands/savannas, forest mosaics, or already degraded forest.

[5] 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. Available at www.sciencemag.org/content/344/6187/981-d.

[6] Hansen, M., Potapov, P., Margono, B., Stehman, S., Turubanova, S., and Tyukavina, a (2014). Response to comment on "High-resolution global maps of 21st-century forest cover change". Science (New York, N.Y.) 344, 981. Available at: http://www.sciencemag.org/content/344/6187/981.5.full.pdf.

# 2 Methods

## 2.1 UMD (Hansen *et al.*) classification

Matt Hansen and colleagues published a seminal paper in the journal *Science* in 2013[7]. The study is a collaboration between various US-based scientists (from the University of Maryland and other institutions) and Google, using the latter's Google Earth Engine to process the thousands of terabytes of the complete Landsat 7 archive from 2000-2012, covering the whole world excluding Antarctica, and has since been updated to include Landsat 8 data and extend the forest loss layer to 2013.

The method involved first pre-processing all Landsat scenes for the growing season (654,000 scenes in total), correcting and normalising them so all were equivalent regardless of calibration or atmospheric conditions, and developing an automated process to remove all cloud and cloud shadow. Then a set of variables are extracted from all valid observations for each pixel, including features related to average greenness, and trends in that greenness through time.

Using an extensive network of training data gleaned mostly from manual interpretation of hyperspatial (very high resolution, ≤5 m pixels) data, automated decision trees were set up to enable predictions of the percentage tree cover (in the year 2000), forest loss, and forest gain per pixel. The forest loss layer returned either 'no change', or a single year from 2001-2013 where loss occurred. By contrast the forest gain layer returns either 'no change' or 'gain', but does not offer a year for this. Both loss and gain can occur in the same pixel, for example where a pixel has been deforested in 2001 but regrows and is at some point reclassified as forest, or due to error (both are produced independently). However, loss can occur only once using this algorithm, so such a pixel could never again be flagged as deforested.

Note that the 'gain' product does not allocate a specific year, only that gain occurred over the period 2001-2013; only the 'loss' product specifies a year of change. This makes subsetting the time period and calculating net change in forest area impossible: net change can only be calculated for the full period, and even then with difficulties due to a subset of pixels featuring both gain and loss, with no information as to whether the gain predates or postdates the loss event.

Forest loss is defined in the paper as 'stand-replacement disturbance or the complete removal of tree canopy cover at the Landsat pixel scale'. It is unclear whether this is meant to include a more subtle change whereby trees are removed from an area that remains forest ('degradation'), however it is clear that at least some pixels flagged as Forest Loss have undergone degradation. No initial sift is made for canopy cover: a pixel with a starting canopy cover of 0 % can still be flagged as deforested (there is an inherent assumption here that canopy cover will have increased prior to the loss event, but the canopy cover and forest loss layers are produced entirely independently). Similarly artificial plantations and natural forest are not differentiated. Therefore some 'forest loss' events classified by the product would not be technically deforestation: some will occur in pixels that do not meet local definitions of forest (due to canopy cover, height or area criteria), and thus may not be deforestation depending on the local definition of forest; and some will be degradation, the removal of some trees from an area that remains classified as forest: again due to the area still meeting cover, height or area criteria after the forest loss event to meet the definition of forest in that area. The forest cover layer is unfortunately not produced annually, and thus it is impossible using this dataset alone to convert the 'forest loss' layer into layers that would approximate maps of deforestation and degradation based on

[7] Hansen, M. C., P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, S. V. Stehman, S. J. Goetz, T. R. Loveland, A. Kommareddy, A. Egorov, L. Chini, C. O. Justice, and J. R. G. Townshend. 2013. "High-Resolution Global Maps of 21st-Century Forest Cover Change." Science 342: 850–53.

local definitions. This complicates any analysis using these data, and assessing the impact of these complications is a major driver for this study.

It is possible to subset the results by initial (year 2000) canopy cover, as there is a 30 m canopy cover product produced for the year 2000. This is performed extensively by Hansen et al. in their results, with loss and gain widely subsetted into pixels with a treecover in 2000 ≤25 %, 26-50%, 51-75%, and 76-100%.

Hansen et al. performed an internal accuracy assessment based on an independent dataset. This involved collecting data from 1,500 x 120m x120m blocks, distributed across each biome. In each 120m block an assessment was made of canopy cover and change using very high spatial resolution imagery, the best available. No ground truth data was used. Producer's and User's accuracy for Loss data were ~87 %, and for gain data ~75%, with a good balance between Producer's and User's accuracies (suggesting low bias). Overall accuracies were stated at over 99%, but this represents the fact that the vast majority of pixels did not change over the time period, and were correctly flagged as not changing. Of particular relevance for this study the figures for the tropics were lower, with accuracies for Loss of ~85 %, and for gain a Producer's accuracy of just 48 %. These are overall errors over the whole time period, these figures are not available for the accuracy of a particular year. These results are mixed, and do not truly allow an assessment of whether the Hansen et al. (UMD) dataset is suitable for use to calculate unbiased figures for the Hectares Indicator.

We therefore decided to further, independently assess the accuracy of this dataset in two contrasting sites - Cerrado in Brazil, and a forest-savanna matrix in Ghana. In both cases a combination of RapidEye and SPOT data (all but one scene at 5 m resolution) were used, providing a resolution 36 x higher than the Hansen *et al.* (UMD) dataset. In order to avoid methodological or producer bias, these independent high resolution data were classified by two different operators using different methods, and their products then independently assessed by a third operator.

## 2.2 High resolution validation dataset methods

In both cases three hyperspatial (5 m) scenes were used per site, with at least a decade separating the combined span of three images. This allowed a thorough assessment of the accuracy of the annual Hansen product, without the potential errors that accrue from the use of just two images.

Automated classification was performed using the support vector machine classifier in ENVI 4.8 (Exelis) software, and statistical results calculated using *R*, by Edward Mitchard (EM) of the University of Edinburgh. Manual classification based on careful visual interpretation of the hyperspatial optical data was performed by Veronique Morel (VM) of Ecometrica using ArcGIS software. Independent verification of both these classifications was performed (see Appendix 1, Brazil carried out by Karin Viergever (KV) of Ecometrica; Appendix 2, Ghana, carried out by EM).

The automated method produced classified images of 'forest' and 'non-forest' for each of the three dates, and then these classified images were compared to the UMD forest loss data both in terms of total hectares deforestation detected per period, but also using direct pixel comparisons to produce estimates of errors of commission and omission. The manual method compared the images directly to the UMD forest loss data to obtain estimates of errors of commission and omission, as well as correctly classified forest loss.

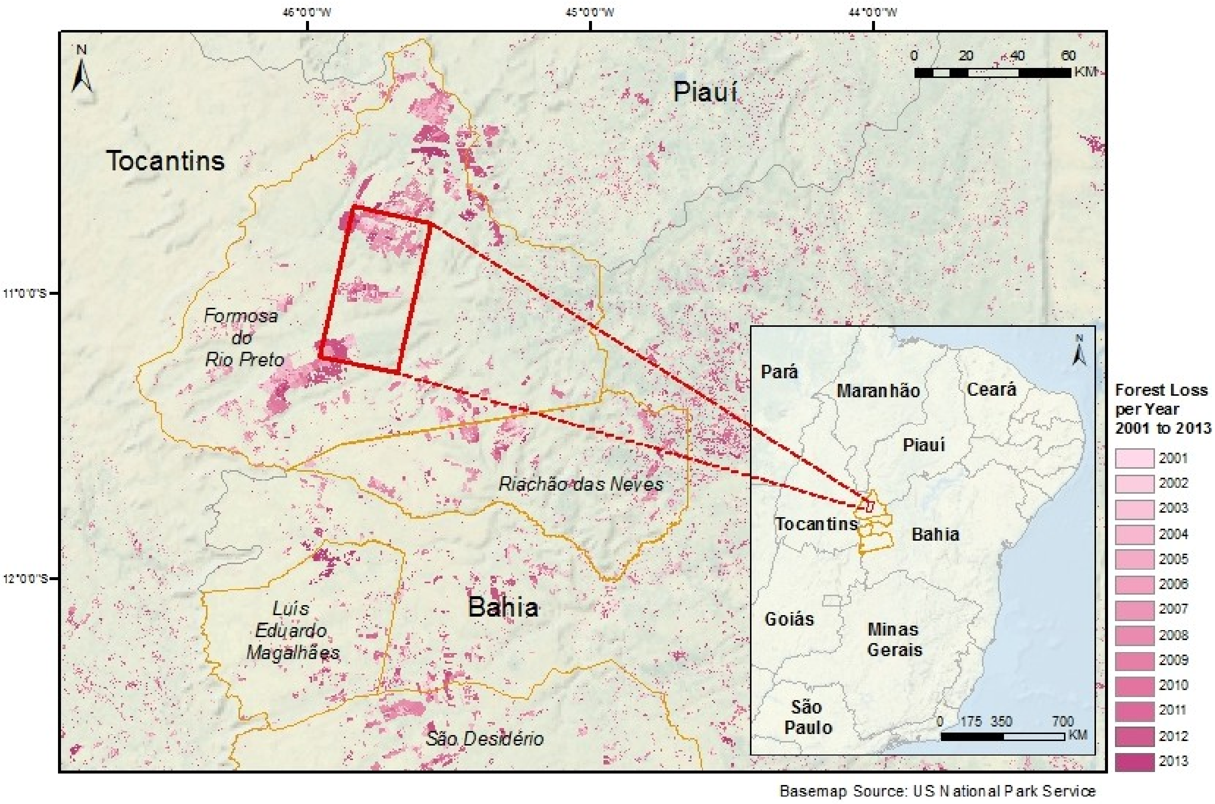
The specific details of the data and methods applied to each site follows.

### 2.2.1 Brazilian Cerrado - Methods

**Scenes**

An area located at the northern tip of Formosa do Rio Preto Municipality in Bahia state was chosen that has been subject to considerable deforestation activity in the past (as identified by the UMD forest loss per year dataset) and where cloud-free high resolution optical imagery was available for three relevant time periods between 2001 and 2013.

**Figure 1 - Brazil study area overlaid on UMD forest loss dataset**

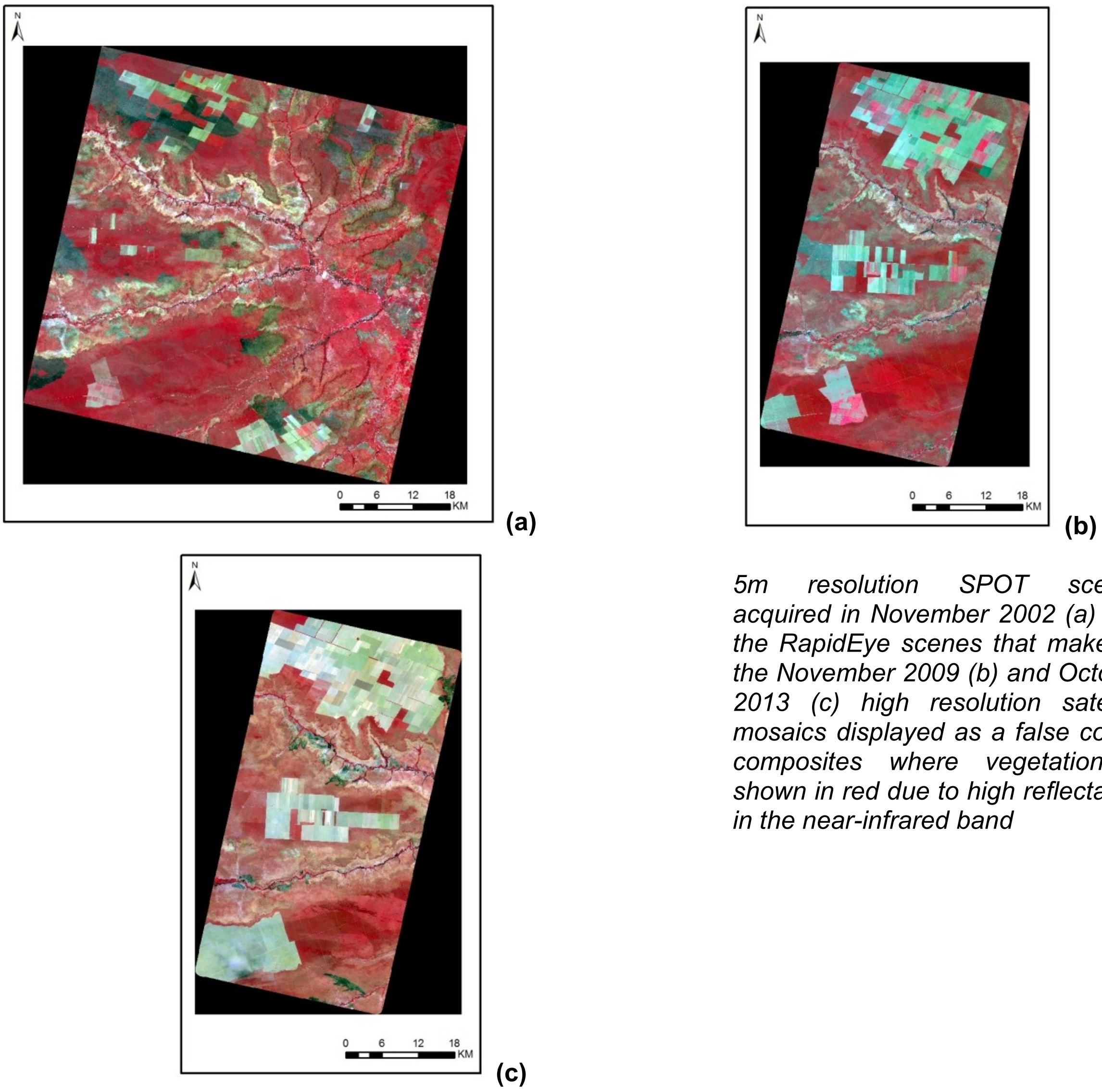



Care was taken to select scenes in the dry season, where the contrast between grass, crops and trees should be at its greatest. Three scenes were selected covering 11 years at a 5 m resolution (Table 1, Figure 2). The area covered was 374 547 ha in size for the 2002 SPOT imagery, and 193 568 ha for the 2009 and 2013 RapidEye imagery.

**Table 1 - Brazil high resolution data**

| Date | Data Type | Resolution |
|---|---|---|
| 13/11/2002 | SPOT | 5 m |
| 14/11/2009 | RapidEye | 5 m |
| 01/10/2013 | RapidEye | 5 m |

**Figure 2 - Brazil high resolution imagery - false colour composites**




*5m resolution SPOT scenes acquired in November 2002 (a) and the RapidEye scenes that make up the November 2009 (b) and October 2013 (c) high resolution satellite mosaics displayed as a false colour composites where vegetation is shown in red due to high reflectance in the near-infrared band*

**Data preparation**

The RapidEye scenes overlaid precisely when compared at a 10 x zoom. A slight offset was noticed with the SPOT scene, which was corrected using the manual selection of 50 Ground Control Points by eye, and a 1-degree warp, which had a Root Mean Square error of 0.3 pixels (1.5 m). It can therefore be assumed that the pixels overlapped precisely.

The Hansen et al. data (Version 1.1) was downloaded and warped to match the UTM, 5 m projection of the RapidEye and SPOT data. No offset was detectable. No removal was performed for pixels classified as deforested in the Hansen forest loss dataset that were below the canopy cover or forest area threshold in the Hansen tree cover dataset, i.e. to be counted as 'forest' in Brazil.

**Automated classification methods**

In the scenes, confirmed by looking at higher-resolution data available in Google Earth and geo-located photos on the same system, there appear to be three types of major landcover: Crops, Forest and natural-non-forest ('Shrub'). A dataset of 10,000 pixels for each of these classes was created for each of the three time points, and used to train a classifier. The final classifier used was a Support Vector Machine with 2 pyramid levels and a Radial Basis Function, using all bands as well as a Standard Deviation 5x5 textural filter. This produced User's and Producer's accuracies over 98 % in all cases compared to the input dataset. These maps were then compared to produce maps giving deforestation for 2003-2009 and 2010-2013, and these forest loss maps directly compared to the UMD data to produce maps showing errors of omission and commission.

**Manual classification methods**

The individual scenes were carefully colour balanced so colours and contrast levels matched. The three data mosaics dated 2002, 2009 and 2013 were first compared visually in detail. Areas of change (forest loss and regeneration) were identified and compared to the UMD Hansen data and assessed for differences which could indicate (i) areas incorrectly mapped as deforestation, i.e. errors of commission, and (ii) areas of deforestation that were missed by the Forest Loss per Year product, i.e. errors of omission. For the latter, care was taken to exclude from the analysis areas that had changed from non-forest vegetation cover to bare soil, which can occur due to seasonal changes and agricultural practices but which do not represent deforestation. Such areas were carefully digitised on screen.

**Verification**

A point-based assessment of the automated and manual classification results was carried out independently by a third interpreter. In the absence of field data, the assessment is based solely on the interpretation and opinion of the third assessor using the same high resolution optical data, and is presented as Appendix 1. Although the verification was done in the form of a traditional point-based accuracy assessment, the results should not be interpreted as an accuracy assessment. The outcome of this verification exercise points out possible errors in the two classification results and gives insight into the possible causes of errors in all datasets.

### 2.2.2 GHANA - METHODS

**Scenes**

It was very difficult to find an area in Ghana with cloud-free data available for three points in time. Eventually one area was found on the border of the Western & Central Region in southern Ghana, including several intact forest patches and large areas of forest-savanna mosaic (Figure 3). Care was taken to select scenes in the dry season, however due to limited data availability these span a wider range of months than the Brazil dataset. Rainfall for the sites in advance of the images were compared and no large differences were noted (previous 2 months total rainfall within 30 % in all three cases). Unfortunately only 10 m data was available for the earliest time point - though it should be noted this still offers 9 pixels for every 1 Hansen pixel, and therefore still provides a reasonable dataset for assessment, it is not ideal. For 2013 a composite of two Rapideye scenes captured within 4 days was used to most closely replicate the area of the SPOT scenes (Table 2). The area of overlap between the three scenes was quite low, at only 89 410 ha (Figure 4).

**Figure 3 - location of the SPOT & RapidEye satellite data in Ghana**

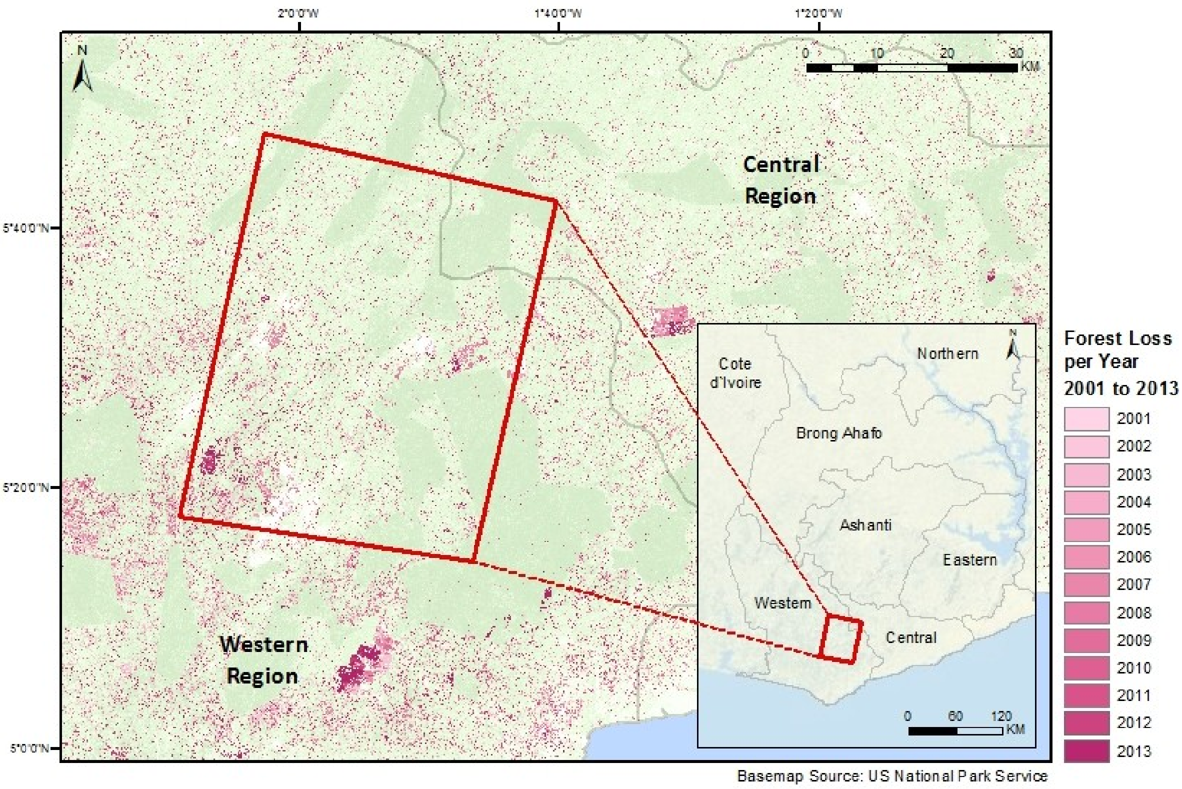
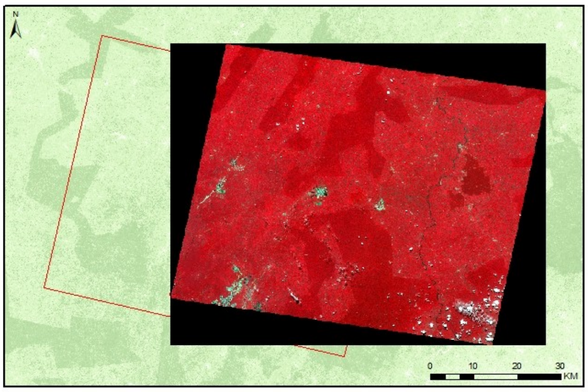
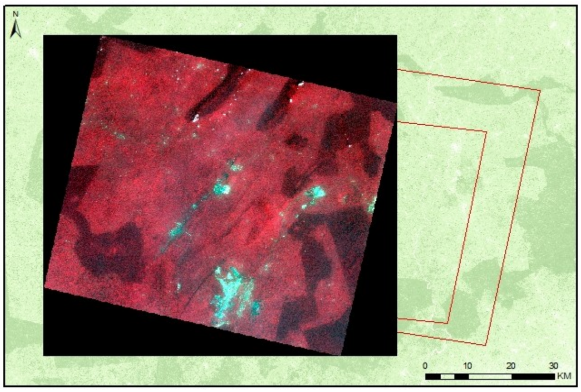


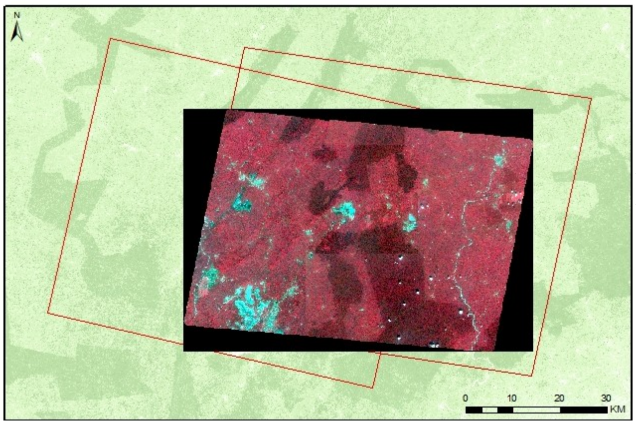
**Figure 4 - Footprints of SPOT and RapidEye scenes, false colour composites**




**(a)**




**(b)**




**(c)**

*a) 10 m resolution SPOT (b) 5 m resolution SPOT scene (c) 5 m resolution RapidEye mosaic. The mosaics are displayed as false colour composites where vegetation is shown in red due to high reflectance in the near-infrared band.*

**Table 2 - Ghana high resolution data**

| Date | Data Type | Resolution |
|---|---|---|
| 15/04/2001 | SPOT | 10 m |
| 12/01/2007 | SPOT | 5 m |
| 17/12/2013 & 21/12/2013 (composite scene) | RapidEye | 5 m |

**Data Preprocessing**

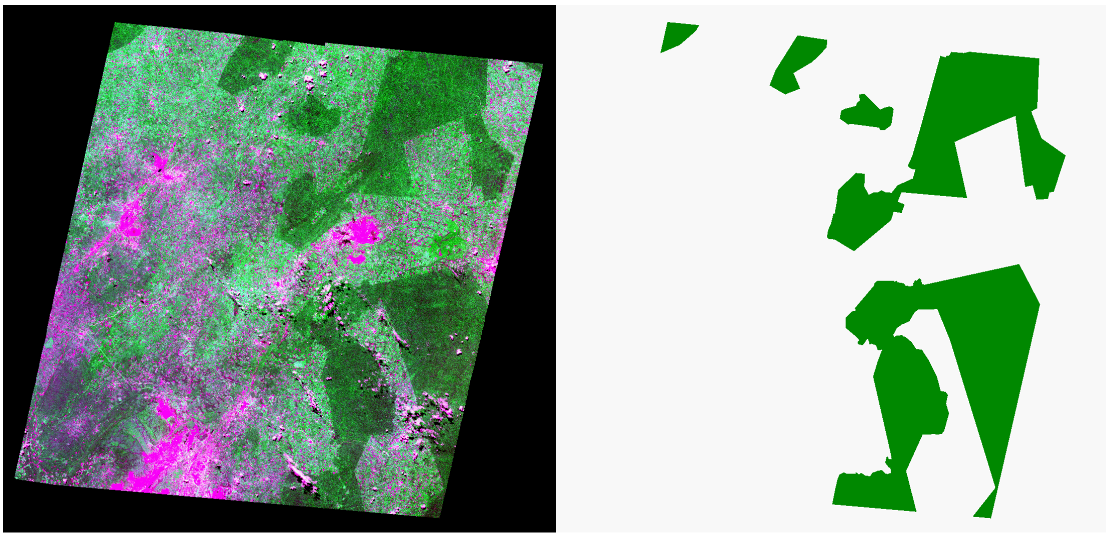
When compared at a 10 x zoom no offsets could be seen between the images. They were therefore warped to each other using their existing georeferencing, but non-overlapping areas were masked. The 10 m resolution data from 2001 was upsampled to 5 m in this process using a cubic convolution function.

The forest definition for Ghana is very broad - any patch of trees with a canopy cover of at least 15 %, a minimum potential height of 2 m, and a minimum area of 0.1 ha, is technically forest. The study area is covered by large blocks of forest, which clearly meet this definition, but also large areas of mixed forest cover, featuring small patches of trees around a landscape of heavily-human-influenced savanna. The UMD dataset sees much of this forest-savanna-agriculture matrix as forest (i.e. above the 15 % canopy cover threshold) and sees rapid deforestation and reforestation throughout.

We tried many methods using a number of different approaches to classify this forest-savanna-agriculture matrix as forest and non-forest. An additional three-class approach was taken, with forest, non-forest and scrub; and a further as forest, agriculture, non-forest and scrub. In no cases were automated classification accuracies greater than 75 % achieved. In order to assess the accuracy of a 2nd dataset, it was felt that the primary dataset accuracy had to exceed 95 %, or at the very least 90 %, to be able to state any conclusions. Therefore a decision was taken to only compare the maps around the main forest blocks, which were identified by a classification of the 2001 SPOT image using a 13x13 median filter and a ground truth dataset based on point from within or outside the national parks, to produce a broad Intact-Forest vs Non-forest Classification.

These forest blocks were extended by 500 m, in order to include dynamics around their boundaries, and the resulting layers were then used to mask all three images as well as the UMD classification, and further analyses took place only within these forest blocks. The total area analysed was thus 45 591 ha.

**Figure 5 - Landscape in 2002 & Forest blocks**



**Automated classification methods**

Within the intact forest blocks, a dataset of 5000 forest and 5000 non-forest pixels were used for classification at each time point. A 2-layer Neural Network was used to perform the classification using textural (standard deviation of a 5 x 5 window) and all spectral metrics available. User's and Producer's accuracies exceeded 96 % in all cases against the test metrics, though in a test these fell to below 80 % when textural metrics were excluded. Where cloud was confused with non-forest this was manually removed from the analysis for that period.

**Manual classification methods**

The three data mosaics were first compared visually in detail having been prepared so their colours and contrast levels matched. Areas of change (forest loss and regeneration) were identified and compared to the UMD Hansen data and assessed for differences which could indicate (i) areas incorrectly mapped as deforestation, i.e. errors of commission, and (ii) areas of deforestation that were missed by Forest Loss per Year product, i.e. errors of omission. Any areas identified as deforested between the acquisition of the 2001 SPOT imagery and the 2007 SPOT imagery, and subsequently the between 2007 SPOT and 2013 RapidEye image that were not included in the UMD Hansen data set for any year up to 2013 were mapped by means of on-screen digitization, and then quantified and summarized.

**Verification**

For the Ghana study area, a point-based method would not give meaningful results since the areas of forest loss were so small that a grid containing many thousands of points would have been necessary to capture a sufficient number of change pixels. Instead a direct visual comparison of the two maps was performed, allowing a qualitative assessment of the differences between the two interpretations.

# 3 Results

## 3.1 Results summary

The UMD dataset performed variably: it appeared to detect forest loss well in the Brazil study site, but poorly in Ghana. In Brazil errors of omission and commission were both reasonably low and balanced: the two classification methods produced slightly different estimates, but the balance of probabilities suggests that overall errors of omission and commission were below 15 % overall. Total deforestation rate estimates between the two products were very similar in Brazil, with the UMD figures in between those estimated from the manual and automated classification of the high resolution data. In Ghana, however, errors of omission dominated, with the two classification methods producing very similar results suggesting >80 % of forest loss was missed by the UDM dataset.

## 3.2 Detailed results: Brazil

The two independent analyses produced broadly similar results: the UMD data appears to perform well in this landscape comprising cerrado vegetation with transitions to large fields of arable crops. Some forest loss events were falsely detected by the UMD dataset as having occurred in fields, and some loss events were observed later than they occurred, but in general the majority of forest clearance was detected.

The automated classification detected more errors of commission (change detected where no changed actually occurred) than the manual classification (24 % to 3 % over the whole time period), apparently caused by different interpretations as to what is or is not forest. A third independent point-based assessment concluded that it is likely that the automated classification overestimates commission errors for the period 2003-09 (Appendix 1). Both methods detected similar rates of omission (where real change was missed), at about 13-14 % over the whole time period.

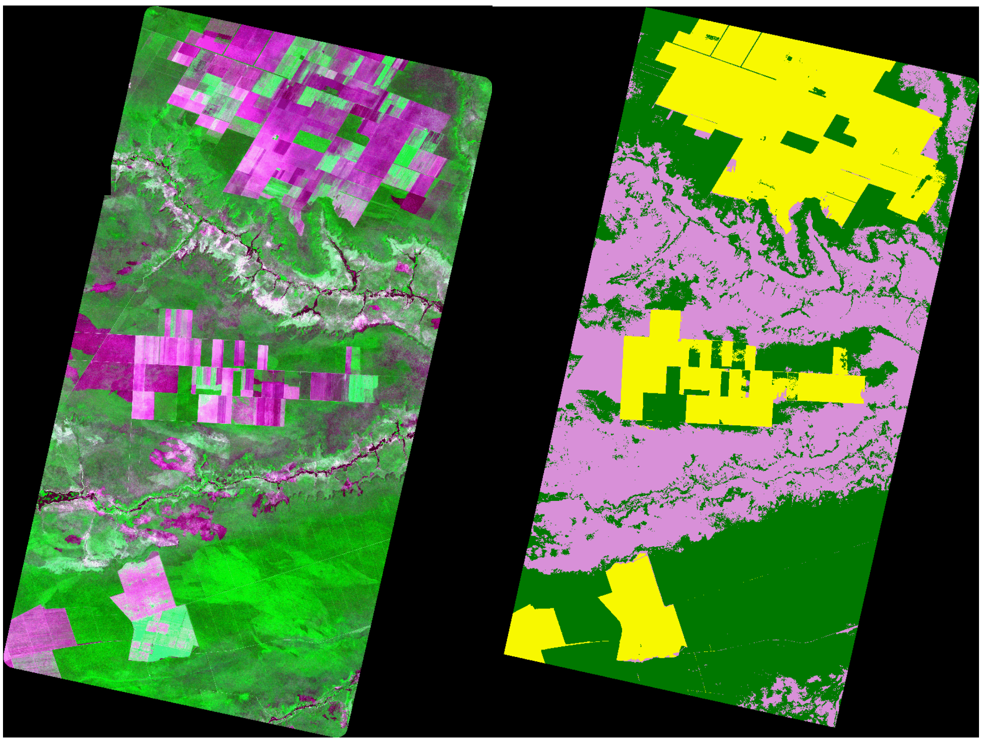
In terms of area-summary statistics (i.e. deforestation rates), the UMD datasets predicts deforestation rates slightly lower than those in the manual classification, and slightly higher than those predicted in the automated classification, so in all likelihood these are approximately correct.

### 3.2.1 Brazil - Automated classification

The classification procedures dividing the imagery into shrubs, non-forest and forest classes appeared to work very well, as shown in Figure 6 below.

***Figure 6 - 2009 Classification - Brazil - automated classification***

**a) R-G-B Original image** **b) Classification**

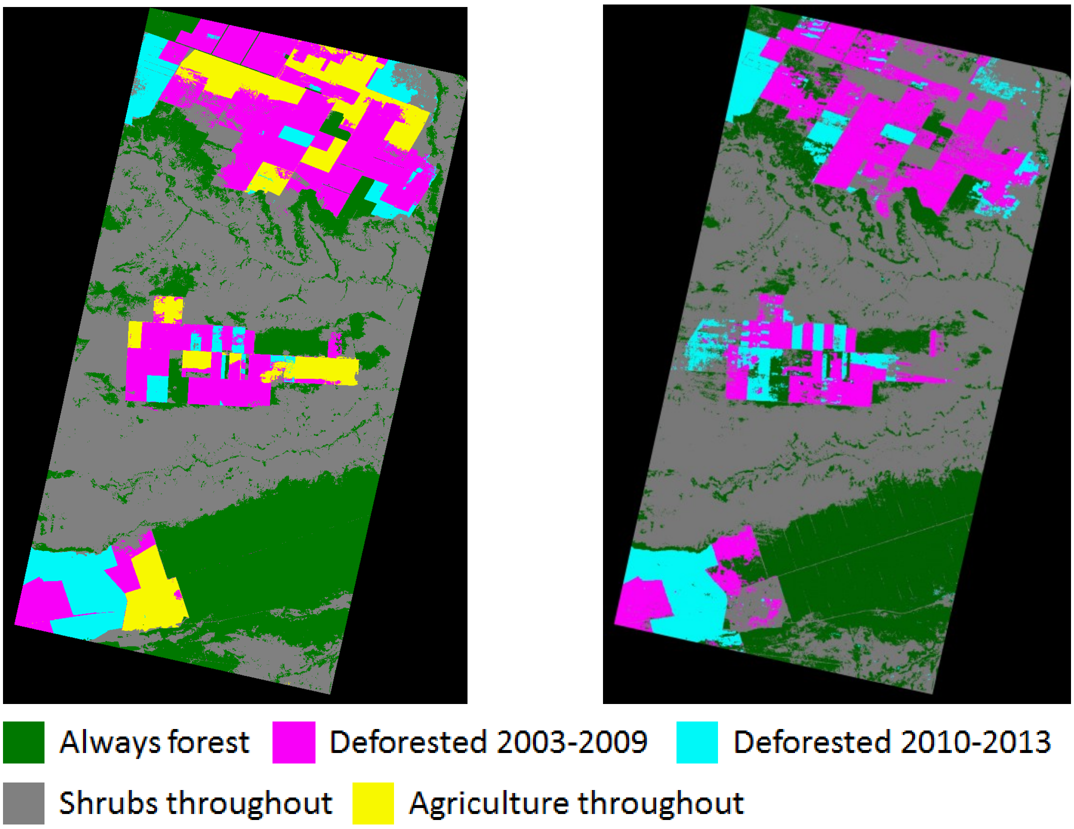


***Key: Green-Forest; Yellow-Agriculture; Mauve-Shrub***

The timeline of deforestation according to the SPOT and RapidEye data is shown in Figure 7 below - it can be seen that the forest in this area is steadily eroded over time by new large fields for agriculture. The Hansen dataset shows a broadly similar pattern, with a similar deforestation rate, but with a more patchy pattern.

***Figure 7 - Deforestation 2003-2013 - automated classification***

**a) SPOT-RapidEye** **b) UMD**




***For the UMD dataset no 'agriculture' and 'shrub' classes exist, but for comparison any pixels not deforested with canopy cover in 2000 between 0-29 are coloured grey, and >30 green.***

In the above imagery there is a distinction made between agriculture and shrubs to assist with interpretation - errors are thus visible even at this stage as there are clear areas (e.g. towards the bottom left) where the UMD data detects deforestation in areas flagged as agriculture in the SPOT data in 2002. No differentiation is made between shrubs and agriculture in the UMD dataset, so no such distinction is made for UMD in Figure 7.

In general the UMD and automated SPOT/RapidEye analyses match very well: the pink and cyan areas in Figure 7 mostly overlap. Additionally the difference between forest and non-forest in the final classification appears well matched between the two classifications, giving confidence that future detections will match.

Deforestation appears to centred around three clusters - the north, centre and south-west of the image. In the north and SW the detections match closely in general, though with some confusion and some misidentification of changes, in particular detecting change in areas that were already deforested in 2002 (commission errors), as shown in Figure 8.

**Figure 8 - Comparison of raw imagery, classifications, and detected deforestation for a 3 x 3 km subset using the automated classification**

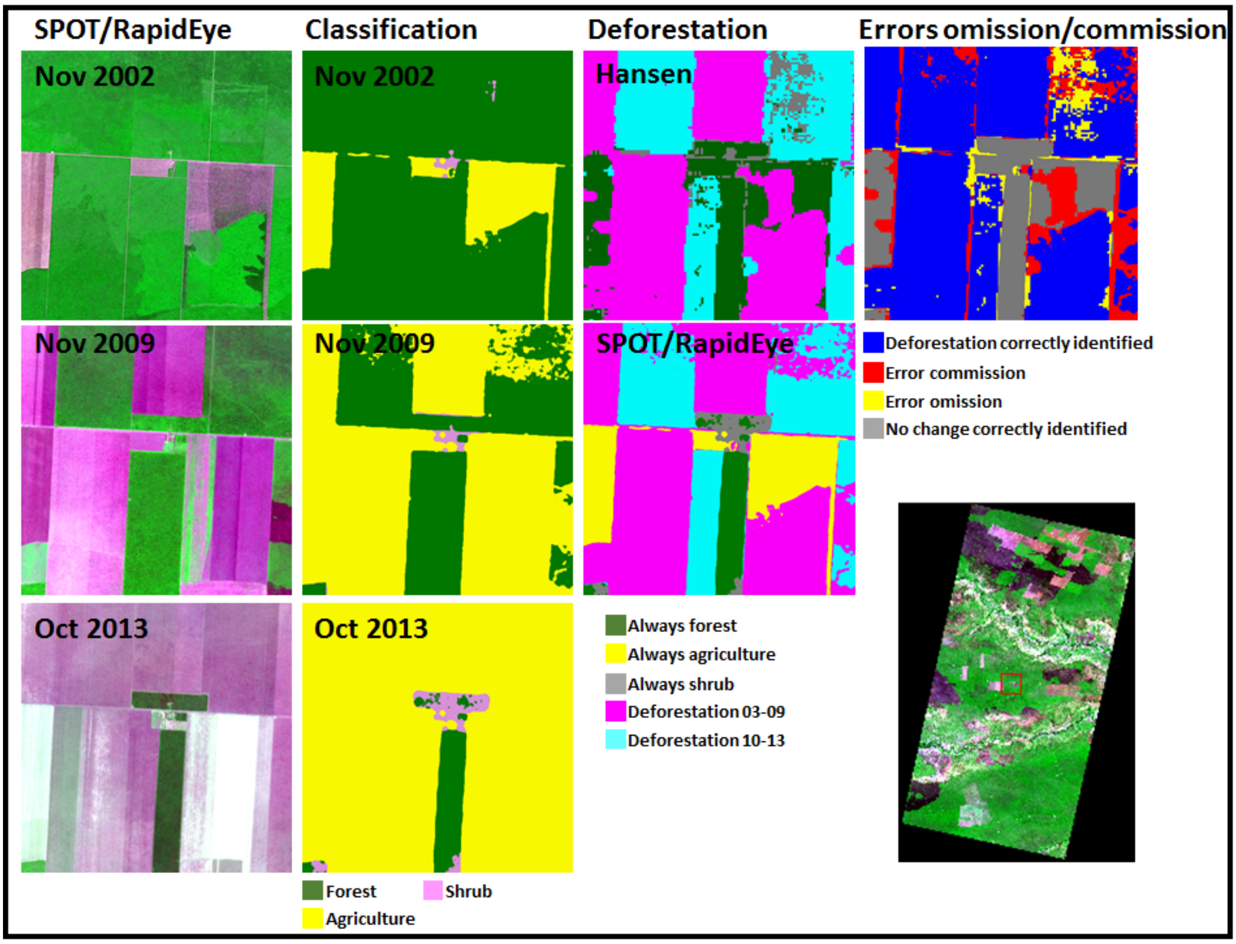



Figure 8 shows that the automated classification of the remote sensing imagery appears to have worked well, identifying forest, agricultural fields and a small patch of regrowing forest apparently correctly. The UMD forest loss data detects the main changes well, but also detects deforestation in some areas that were cleared prior to the start of the study period in 2002: the red pixels in the centre-right of the errors omission/commission box were clearly non-forest in the Nov 2002 SPOT scene, but seen as deforested between 2003 and 2009 by the UMD forest loss data. This is a theme throughout the Brazil case study, with errors of commission dominating over errors of omission in the results of the automated classification. It is possible that this area was deforested earlier in 2002 (prior to the Nov 2002 acquisition of the scene), in which case the error of commission in this case is caused by a misallocation of forest loss to the wrong year. This has been seen elsewhere by the two independent interpreters (Appendix 1).

The complete results comparing the UMD data with the automated classification pixel by pixel are summarised in these Table 3:

**Table 3: Total area and per-pixel comparisons between automated classification and UMD data for Brazil case study**

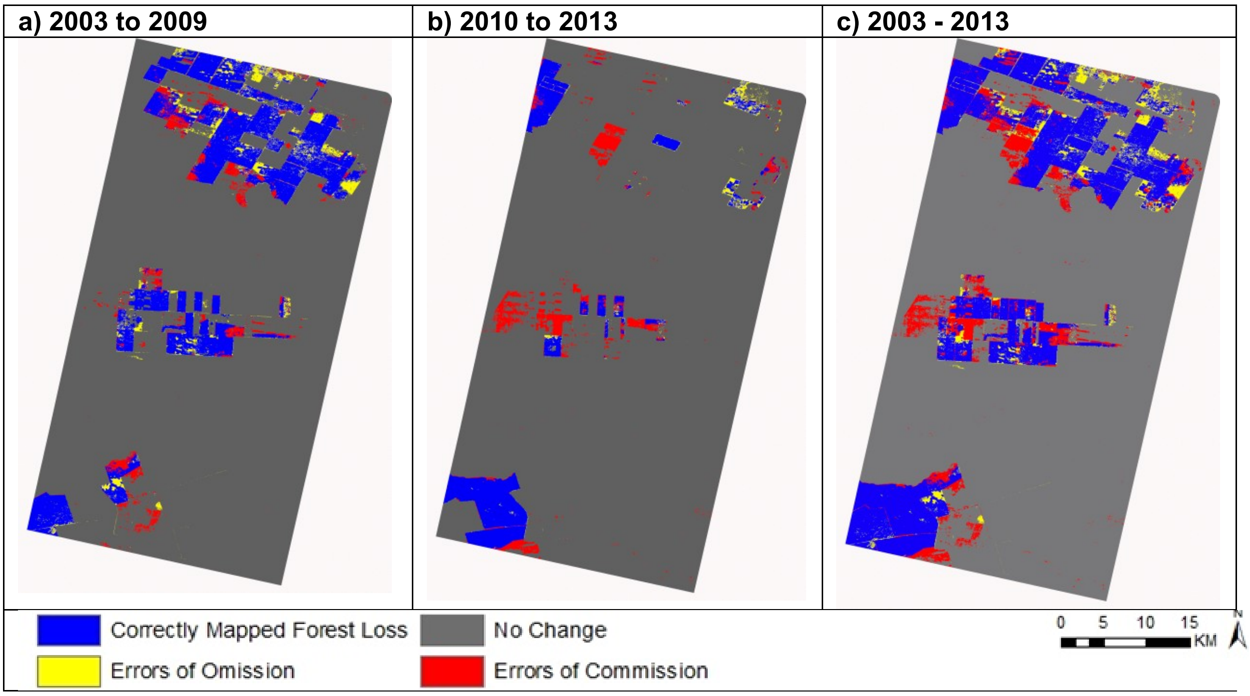
| **Area comparison** | **Spot-Rapideye (ha)** | **UMD(ha)** | |
|---|---|---|---|
| Deforestation 2003-2009 | 29,142 | 27,538 | |
| Deforestation 2010-2013 | 10,731 | 14,490 | |
| Total deforestation 2003-2013 | 39,873 | 42,027 | |
| Area forest stays as forest | 45,455 | 45,849 | |
| | | | |
| **Deforestation rate comparison** | | | |
| Annual deforestation rate 2003-2009 | 4.88% | 4.48% | *per year* |
| Deforestation rate 2010-2013 | 4.77% | 6.00% | *per year* |
| Total deforestation rate (2003-2013) | 4.25% | 4.35% | *per year* |

| **Per pixel change comparison (ha)** | **2003-2009 (ha)** | **2010-2013** | **2003-2013** |
|---|---|---|---|
| Change detected correctly | 23,173 | 9,474 | 32,647 |
| Change detected where no change (commission) | 4,365 | 5,015 | 9,380 |
| Changed not detected where change occurred (omission) | 3,981 | 1,024 | 5,005 |
| No change detected where no change occurred | 137,039 | 137,039 | 137,039 |
| | | | |
| Rate commission | **16.08%** | **47.77%** | **24.91%** |
| Rate omission | **14.66%** | **9.75%** | **13.29%** |

It can be seen that the errors of commission are particularly high in the 2010-2013 period, where the SPOT/Rapideye automated classification estimated there were about 10,500 ha of deforestation, whereas UMD saw 14,500, with over 5,000 ha detected in error. This reflects in the UMD dataset reporting a deforestation rate in that period of 6 % per year, compared to 4.77 % in the automated classification.

It should however be noted that when area averaged over the whole scene (allowing errors of omission and commission to average out) this does not have such a big effect: over the whole period (deforestation 2003-2013) the automated image analysis detects an average deforestation rate of 4.25 %, whereas UMD detects 4.35%: this suggests robustness in the UMD analysis, but this type of comparisons tends to flatter datasets, by averaging considerably in time and space.

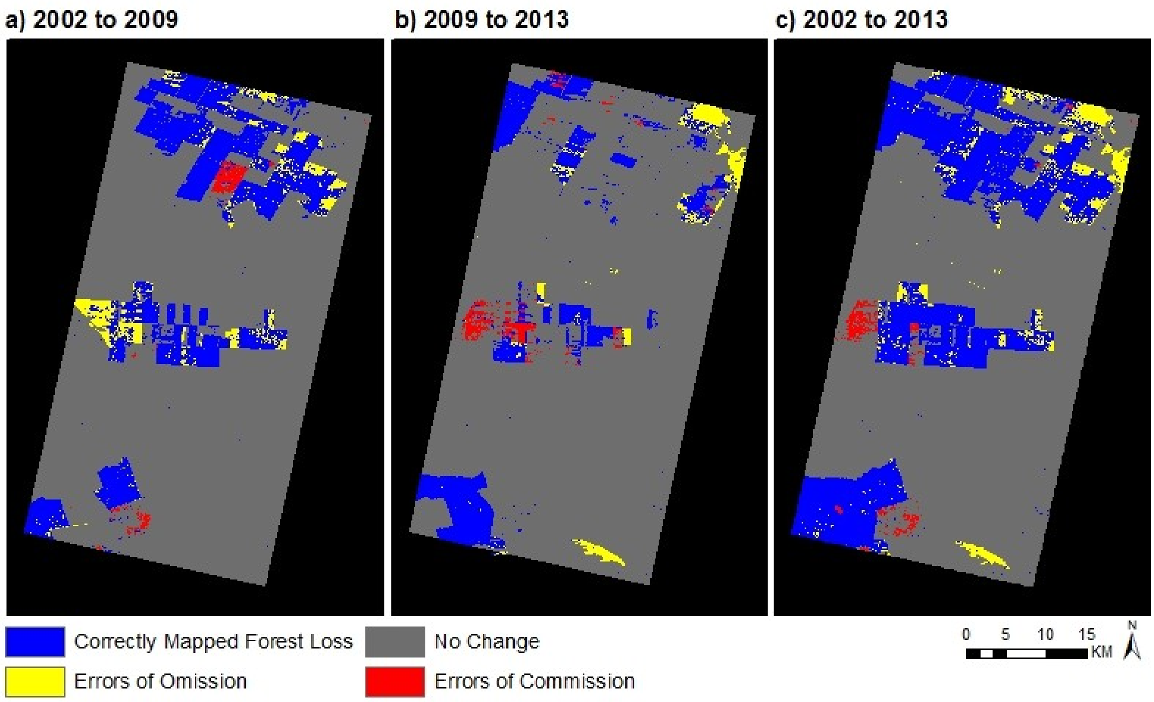
**Figure 9: Errors of Omission and Commission between UMD and automated classification**



### 3.2.2 BRAZIL - MANUAL CLASSIFICATION

Unlike the automated approach, the manual classification directly produces errors of omission and commission (Figure 10).

**Figure 10: Errors of Omission and Commission between UMD and manual classification**




Comparing figures 9 and 10 it can be seen that the two methods produce broadly similar results. However, when looking at the detail it is clear that the manual classification has estimated a much lower rate of commission errors. The independent point-based verification suggests that the automated classification overestimates the commission errors for the period 2003-09, causing a lower estimate of the areas classified as “Correct change”, while the manual classification underestimates the commission errors for the period 2010-13 causing a slight overestimation in the areas classified as “correctly mapped forest loss”. Similarly, the manual classification has estimated a slightly higher rate of omission errors than the automated classification (Table 4). The QA suggests that the automated classification underestimates omission errors for the period 2010-13, which causes an overestimation of the area classified as “Correct no change”. The independent point-based verification suggests that the omission errors in the manual classification are underestimated for the period 2003-09 and overestimated for the 2010-13 period, in both cases this affects the category “Correct no change”.

Both classification methods generally agree that the UMD forest loss classification is reasonably accurate, with both estimating that normally over 70 % of change pixels are correctly detected in all time points. Reasons and examples of the errors, and reasons for the differences between the two classifications, are covered in the Discussion section.

**Table 4: Errors of omission and commission for the manual assessment - Brazil**

| **Area comparison** | **Spot-Rapideye (ha)** | **UMD(ha)** | |
|---|---|---|---|
| Deforestation 2003-2009 | 36,405 | 27,537 | |
| Deforestation 2010-2013 | 24,199 | 14,528 | |
| Total deforestation 2003-2013 | 51,923 | 46,205 | |
| Area forest stays as forest | 45,455 | 45,849 | |
| | | | |
| **Deforestation rate comparison** | | | |
| Annual deforestation rate 2003-2009 | 4.90% | 4.47% | *per year* |
| Deforestation rate 2010-2013 | 8.69% | 6.02% | *per year* |
| Total deforestation rate (2003-2013) | 4.85% | 4.56% | *per year* |

| **Per pixel change comparison (ha)** | **2003-2009 (ha)** | **2010-2013** | **2003-2013** |
|---|---|---|---|
| Change detected correctly | 30,351 | 17,956 | 44,449 |
| Change detected where no change (commission) | 1,379 | 2,246 | 1,670 |
| Changed not detected where change occurred (omission) | 6,054 | 4,018 | 7,475 |
| | | | |
| Rate commission | **3.79 %** | **10.22%** | **3.22%** |
| Rate omission | **16.63%** | **18.29%** | **14.40%** |

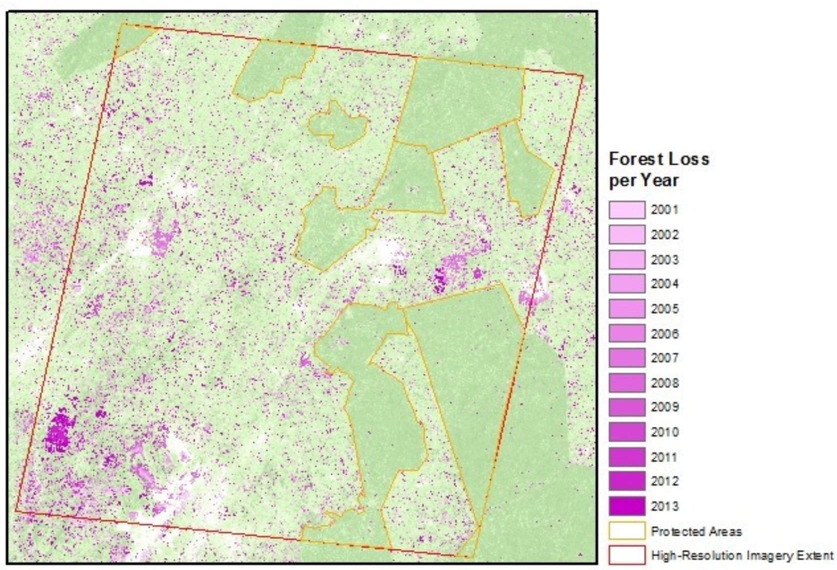
## 3.3 Detailed results - Ghana

The automated and manual classification produce very similar results here: the area of forest loss detected is far higher in the analysis of the 5 m resolution data than predicted by the UMD dataset. Errors of omission estimated through the automated and manual classification methods are of the order of 80 - 90 %. The analysis was only performed in the tall forest blocks, and the UMD data sees very little change within these blocks: the change it does see is mostly correct, with errors of commission just 2-3 %. However, fundamentally less than 10 % of the forest loss or disturbance detected in the SPOT and RapidEye datasets is correctly detected in the UMD dataset. This is also reflected in much lower total deforestation rate estimates in the UMD dataset than the high resolution analyses: unlike in Brazil where total deforestation rates were near-identical between UMD and high resolution analyses, here estimated rates of deforestation are ten times larger in the high resolution analysis.

The results are unequivocal: the forest disturbance clearly visible in the high resolution optical data is not detected by the UMD dataset, making it probably unsuitable for use in reporting against the Hectares Indicator in this country. This does not mean that the UMD dataset is incorrect as such, just that the resolution of its input dataset, and definition of forest change, make it unsuitable for the ICF reporting in this landscape.

It should be noted that the UMD classification does detect a lot of change in this landscape, just not in the main forest blocks (Figure 11). We were unable to create consistent classifications from the high resolution data outside the forest blocks, so could not assess the accuracy of these detected changes. However, it is quite possible that the UMD performs well in the forest-savanna-farmland mosaic of Ghana, just not in the forest block areas.

**Figure 11: UMD Forest loss data showing changes throughout Ghana scene**

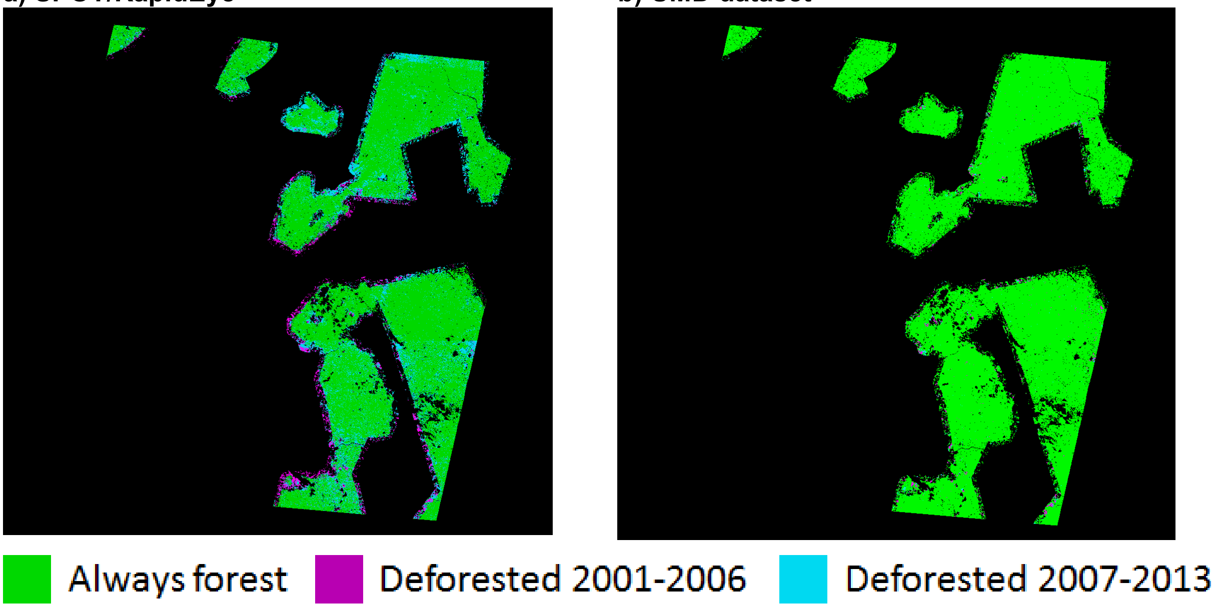


### 3.3.1 Ghana - Automated Classification

Comparing the deforestation maps for the two periods directly (Figure 12), it is clear that the high resolution analysis detects far more deforestation than the Landsat-based UMD dataset. The greater changes exist both in the 500 m buffer around the forest blocks, but also within the tall forest blocks, especially in the 2007-2013 period.

***Figure 12 - Deforestation map Ghana (automated classification)***

**a) SPOT/RapidEye** **b) UMD dataset**




Unsurprisingly given the above, a much higher rate of deforestation is estimated from the high resolution data than the UMD data, and errors of omission dominate (Figure 13, Table 5).

**Figure 13 - Errors of Omission and Commission for Ghana (automated classification)**

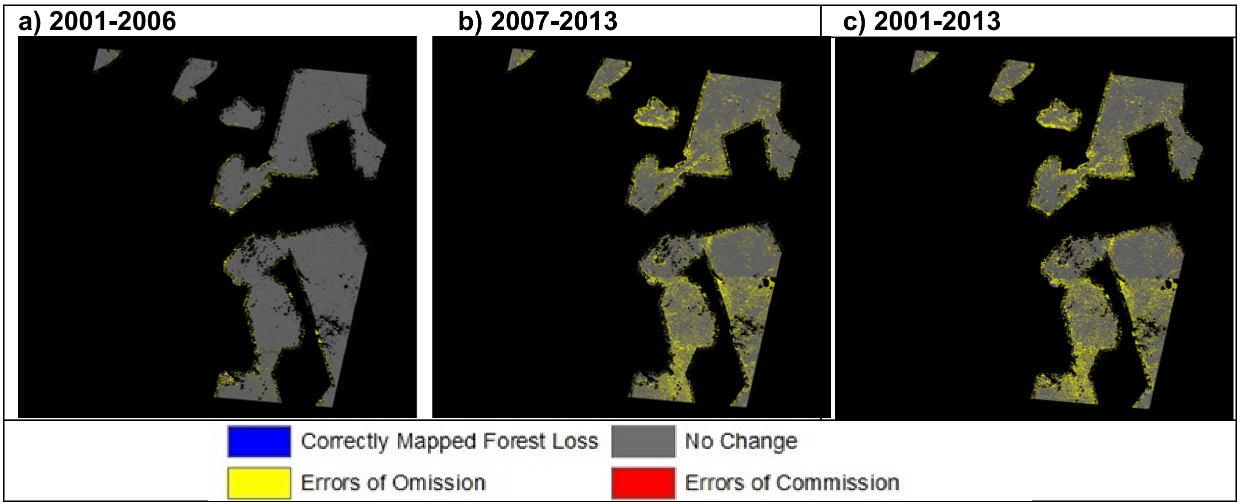



**Table 5 - comparison of UMD and automated classification results, Ghana**

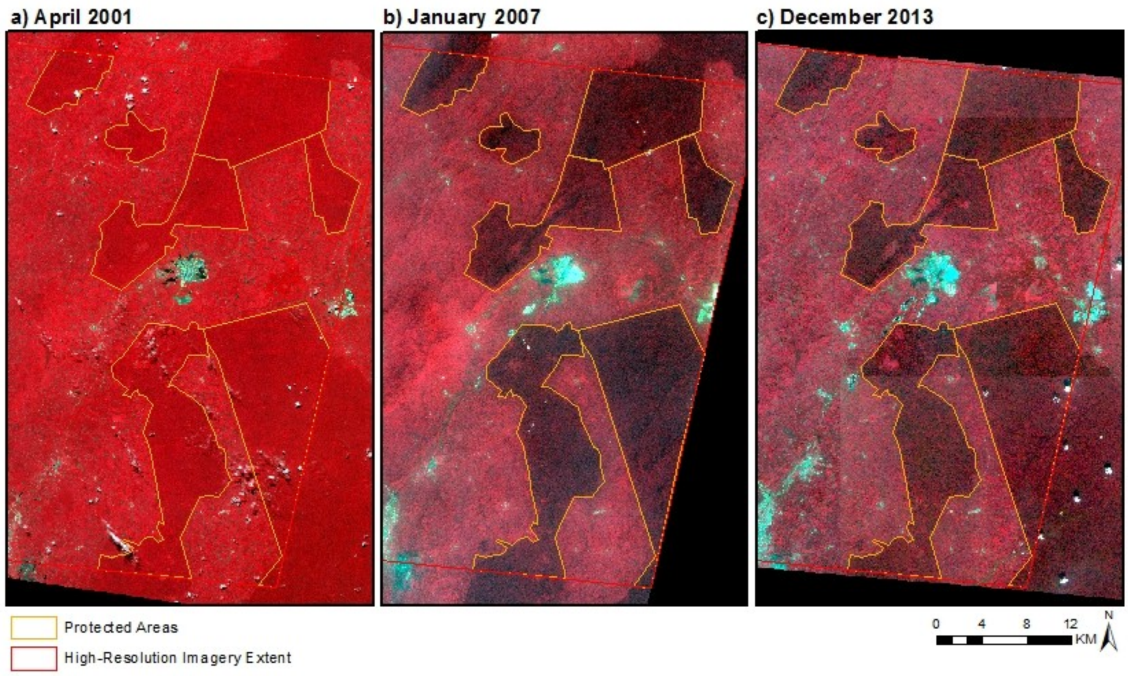
| Area comparison | Spot-Rapideye (ha) | Hansen (ha) | |
|---|---|---|---|
| Deforestation 2001-2006 | 1,637 | 542 | |
| Deforestation 2007-2013 | 7,742 | 322 | |
| Total deforestation 2001-2013 | 9,378 | 864 | |
| Area forest stays as forest | 35,684 | 44,540 | |
| | | | |
| Deforestation rate 2001-2006 | 0.61% | 0.20% | per year |
| Deforestation rate 2007-2013 | 2.55% | 0.10% | per year |
| Deforestation rate 2001-2013 | 1.60% | 0.15% | per year |

| | 2001-2006 | 2007-2013 | 2001-2013 |
|---|---|---|---|
| Change detected correctly | 222 | 162 | 701 |
| Change detected where no change (commission) | 76 | 63 | 139 |
| Changed not detected where change occurred (omission) | 1,325 | 7,352 | 8,677 |
| No change detected where no change occurred | 43,286 | 37,181 | 35,545 |
| | | | |
| Rate commission | **4.94%** | **0.84%** | **1.49%** |
| Rate omission | **85.66%** | **97.85%** | **92.52%** |

### 3.3.2 GHANA - MANUAL CLASSIFICATION

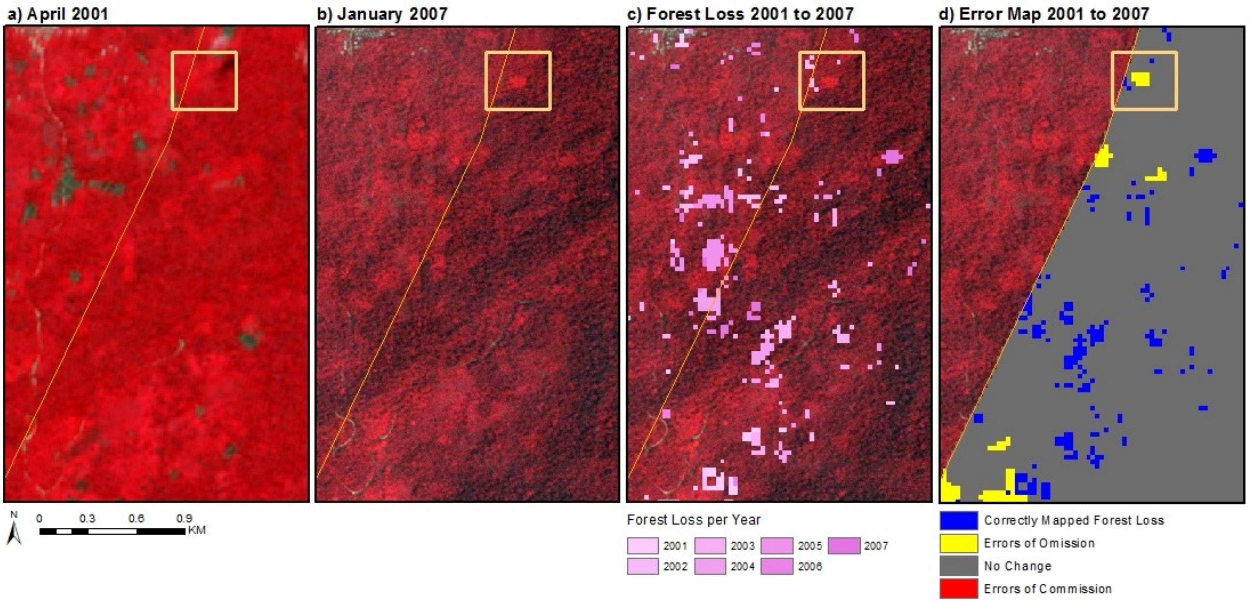
There were difficulties in matching the colours and contrast between scenes, and clear cloud cover visible in the 2001 and 2013 time points. The best colour balance in shown in Figure 14, with this set of images used to perform the classification.

**Figure 14 - input scenes for manual classification, false colour composite**



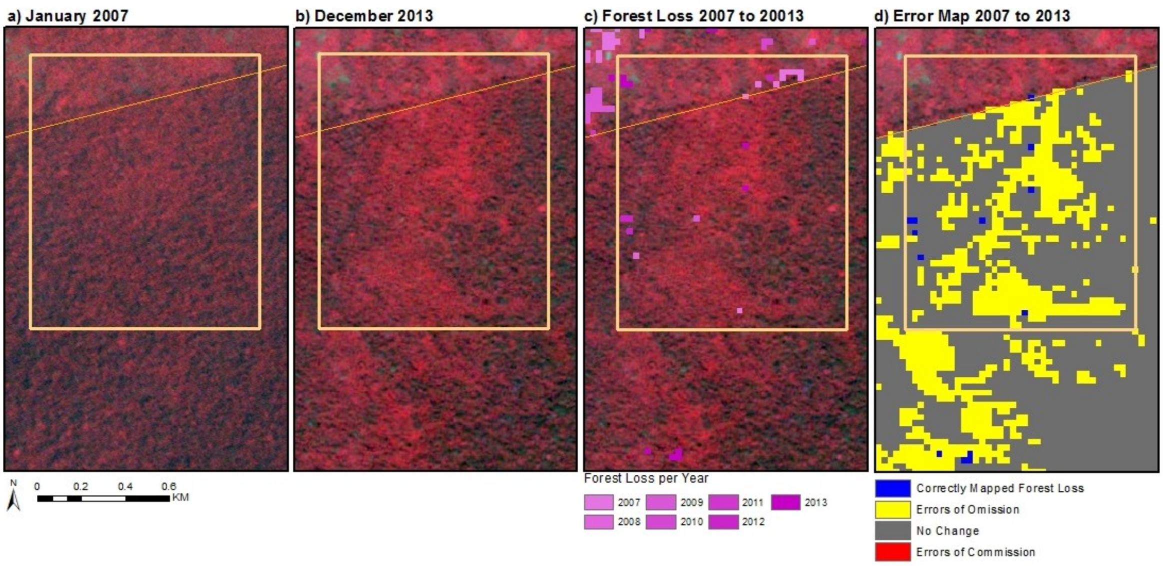

Changes in the data are difficult, but not impossible to see. Figure 15 shows an example of a small area of deforestation that was missed by the UMD dataset, surrounded by others that were detected. Figure 16 shows a much larger area of deforestation/degradation that was undetected by the UMD dataset.

**Figure 15 - Example of a small error of omission in UMD data**




***Comparison of the high resolution data for (a) April 2001 and (b) January 2007 with (c) deforested areas as mapped by the Forest Loss per Year product for the years 2001 to 2007. The error map for 2001 to 2007 (d) shows this as an error of omission as this area is not clearly identifiable as an area of loss in 2001.***

**Figure 16 - Example of large area of omission in UMD data**




***Comparison of the high resolution data for (a) January 2007 and (b) December 2013 (c) deforested areas as mapped by the UMD Hansen Forest Loss per Year product for the years 2007 to 2013. The error map for 2007 to 2013 (d) shows this area of loss as patchy errors of omission where clear deforestation and high levels of degradation can be identified.***

As with the automated classification, errors of omission dominated (Figure 17, Table 6). The manual classification suggests that just 9 % of actual loss was detected and mapped correctly by the UMD data.

**Figure 17 - Ghana errors maps - manual classification compared with UMD data**

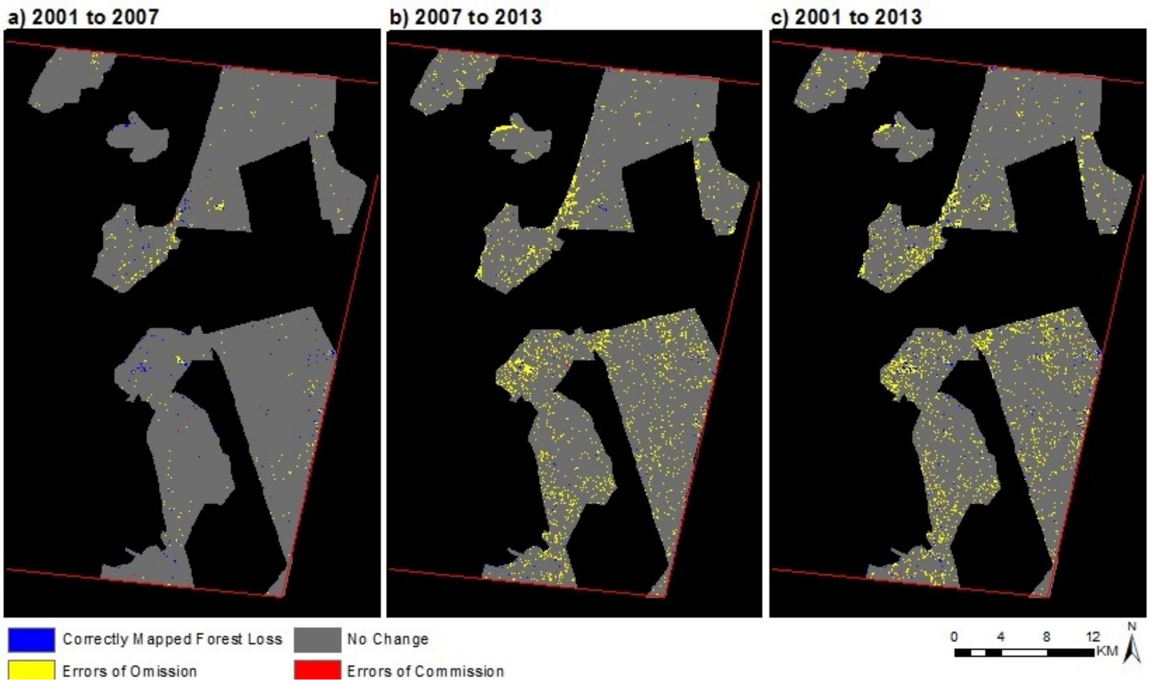



**Table 6 - Errors of omission and commission for the manual assessment - Ghana**

| | 2001-2006 | 2007-2013 | 2001-2013 |
|---|---|---|---|
| Change detected correctly | 230 | 203 | 339 |
| Change detected where no change (commission) | 20 | 11 | 9 |
| Changed not detected where change occurred (omission) | 471 | 2,980 | 3,350 |
| No change detected where no change occurred | 43,286 | 37,181 | 35,545 |
| | | | |
| **Rate commission** | **2.85%** | **0.33%** | **0.24%** |
| **Rate omission** | **67.19%** | **93.62%** | **90.81%** |

# 4 Discussion

The two case studies produced highly contrasting results, with the UMD data performing well against the high resolution data in Brazil, and poorly in Ghana.

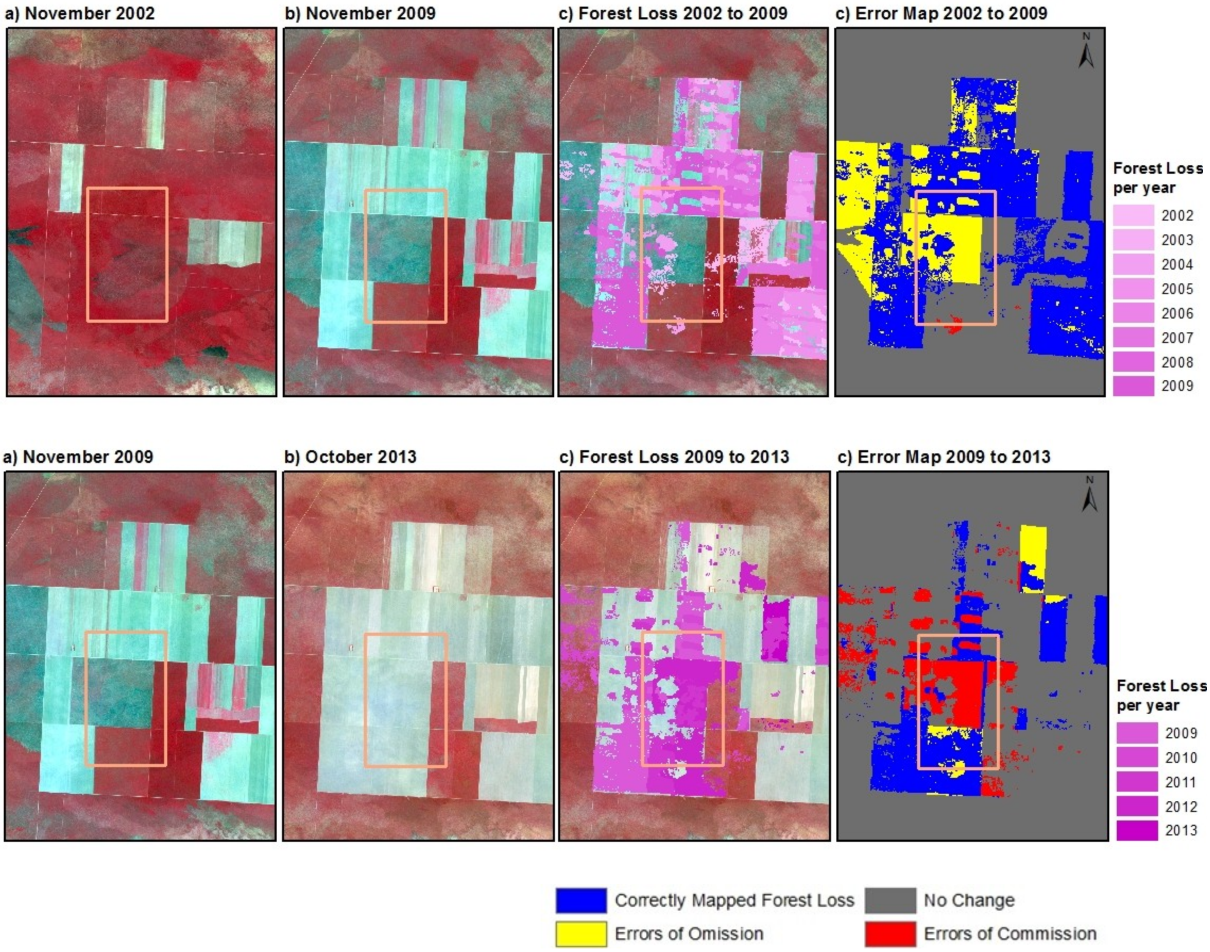
In the Brazilian Cerrado the UMD product performed well, predicting the total area deforested in each time period with high accuracy (Table 3), and with relatively low errors of omission and commission (Tables 3 & 4). There was some evidence of mis-allocated change, with in particular some changes that really occurred in the 2003-2009 period instead being detected in 2010-2013. It is impossible with only three time points to truly assess the proportion of such events: it could be that there were fewer than normal cloud-free Landsat scenes in 2009-10, for example. But the potential for allocating changes to the wrong year should be kept in mind when using the UMD data: this might suggest that reporting over 3-5 year cycles rather than annually would produce more accurate results.

The results in Ghana greatly contrast to the Brazil example, with the SPOT-RapidEye analysis detecting far more change than the UMD analysis, and errors of omission thus dominating. In fact, the differences are so severe that it appears different processes entirely are being detected: at 5 m the RapidEye and SPOT can see small-scale degradation that is invisible in the UMD dataset based on 30 m Landsat. It is known that the pattern of forest loss in Ghana is one driven by small-scale agroforestry, largely for cacao. In many cases farmers are encroaching into forest blocks, but do not clear the whole forest, instead removing only a subset of canopy trees while they clear the understory and small trees to allow cacao to grow. This creates a patchwork of canopy gaps which may be visible to RapidEye at 5 m, but impossible to detect with Landsat data. This may also relate back to the definitions of Forest Change in the UMD dataset: as there are still trees in many of the areas detected as cleared by the automated and manual classifications, the UMD algorithm may have correctly not flagged such pixels as forest loss by its definition. However, this would suggest that the fundamental definitions driving the UMD analysis make it unsuitable for monitoring forest change for the ICF in landscapes such as Ghana.

## 4.1 Specific discussion and lessons from Brazil example

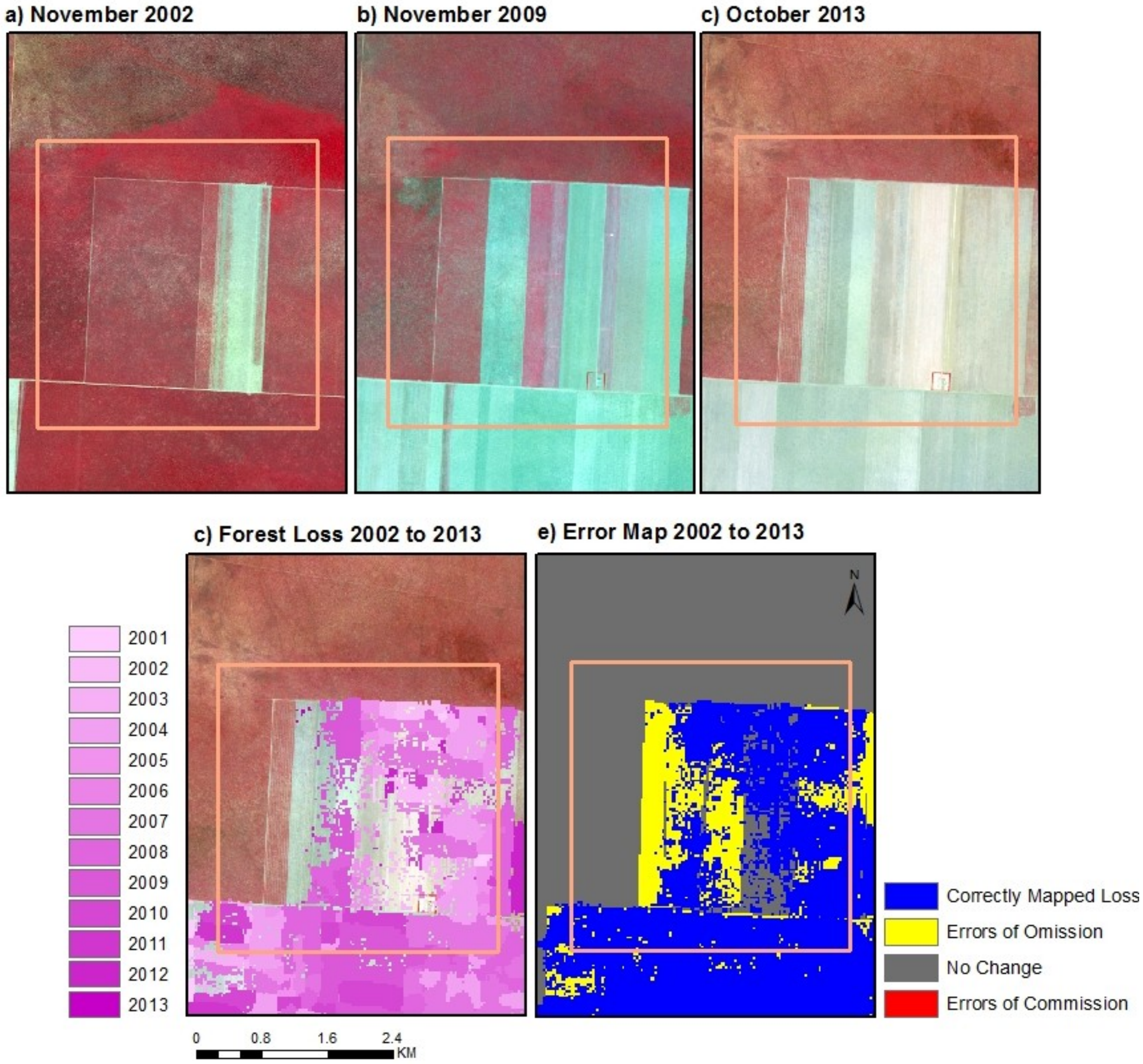
Many of the errors detected in the Brazil cerrado example were a case of mis-allocation. In Figure 18 we show an example where a change that occurred prior to 2009 was not detected until 2012, therefore becoming an error of omission in the first period, and commission in the second.

**Figure 18 - example of deforestation detected several years late in UMD data**



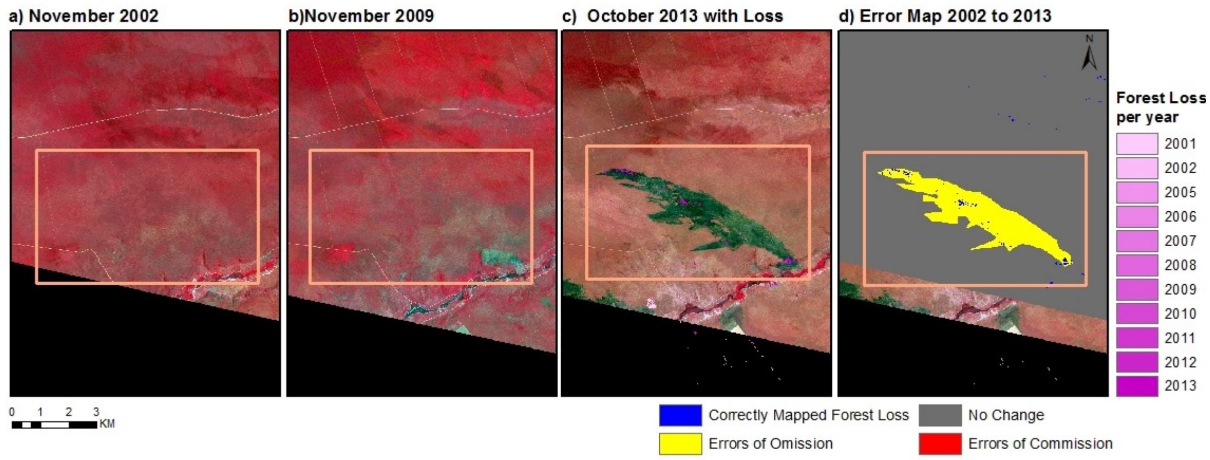
Some errors of omission were however caused by an underestimate of the area lost. Quite often the UMD data is quite patchy, only showing clearance from part of a field for example. Figure 19 shows an example of a case, where a large field (3 km wide) cleared gradually throughout the period is only partially flagged as deforested in the UMD dataset.

**Figure 19 - example of deforestation missed in the UMD dataset**



In broad terms the manual and automated classification of this area agree. However, there were significant differences: the automated classification detected much higher errors of commission, whereas the manual classification detected slightly higher rates of omission. The higher omission rates in the manual classification mostly relate to a single fire event (Figure 20): this was flagged as deforestation by the manual classification, but not in the UMD or automated classification. The independent verification also categorised the area as "correct no change", i.e. no forest loss as it appears that trees are still standing after the fire, but only a ground survey or later image could confirm if this genuinely represented conversion.

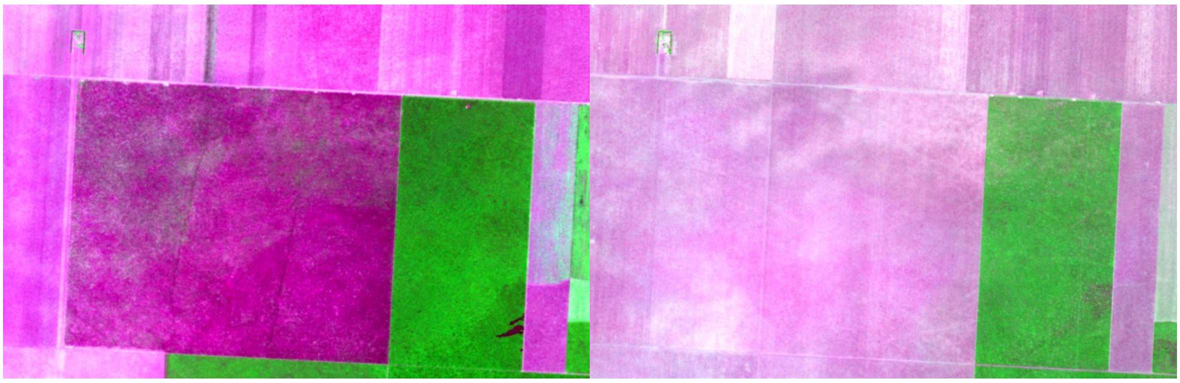
**Figure 20 - burn scar detected in manual classification.**



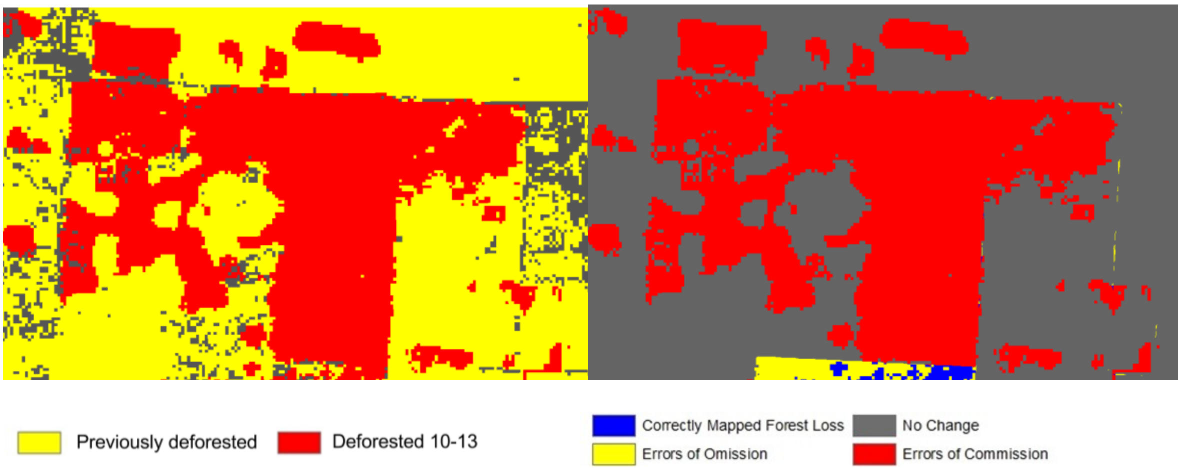
The high errors of commission observed in the automated classification (16 % in the first period, 48 % in the second period) are due to a detection of approximately 5 000 hectares per period where the UMD (and manual) classifications detect change and the automated classification does not. An example of such an area is given in Figure 21. Again these probably relates to differing definitions of forest and forest clearance being ‘booked’ in the wrong date, though there is limited evidence for either of those explanations for the case shown in Figure 21. There is evidence from the point-based independent verification that these errors of commission detected by the automated classification may be incorrect (Appendix 1).

**Figure 21 - example errors of commission detected in automated classification over Brazil.**

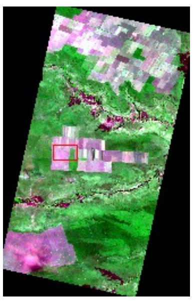
**a) 2009** **b) 2013**



**c) deforestation - UMD detection - 10-13** **d) Detected errors of commission**




**e) red box shows location**



The independent verification concluded that a source of mismatches in the 2 independent classification results were caused by different approaches for dealing with forest loss mapped by the UMD data for the years of image acquisition of the high resolution data used, specifically 2002 and 2009. The automated classification did not take into account areas that were classified by the UMD data as 2002 forest loss (e.g. Appendix 1, Fig A5, A13), while the manual classification did (e.g. Appendix 1, Figs A9, A13, A22, A23, A24). This affected the outcome of the verification results in different ways:

- For areas mapped by UMD as forest loss 2002, and where there was no forest visible on the 13-11-2002 image, we cannot determine when this area was deforested, and if in fact there was ever forest cover at this location. In such a case, the automated classification results of “Correct no change” were deemed correct by the verification. However, the approach of the manual classification gave the UMD data the benefit of the doubt and classified such areas as "correct change", causing a different result to the verification (e.g. Fig A22-A24), and potentially causing an overestimation of the areas classified as “Correct change” for the period 2003-09.
- For areas mapped by UMD as forest loss 2002, and where there is forest visible on the 13-11-2002 image which is replaced by non-forest on the 2009 image, the automated classification categorised the area as "Error of omission", which is potentially incorrect as there is a small possibility that the forest loss may have happened between 13-11-2002 and 31-12-2002.The manual classification mapped such areas as "Correct change" (e.g. Fig A24, A35). Since we have no way to check whether the forest loss did indeed occur in the short remaining time of 2002, the automated interpretation may potentially cause an overestimation of the area classified as “Error of omission”.

Furthermore, the manual classification double counted UMD forest loss mapped in 2009 by taking it into account for both the 2003-09 and 2010-13 periods, causing an overestimation of areas classified as “Correct change” in the period 2010-13 (for example see Appendix 1, Figs A25-30, A34). Further discussions of differences between the two high resolution datasets is in Appendix 1.

## 4.2 Specific discussion and lessons from the Ghana example

It should be noted that without ground data both analyses should be considered as preliminary estimates: neither the UMD nor either of the high resolution analyses represent the truth: that can only be provided by ground data. In the case of the Ghana analysis classification of forest dynamics from these optical data proved very difficult, and we have relatively low confidence in the results. This was despite removing the tricky forest-agriculture-savanna mosaics in between the tall forest blocks from our analysis. We would recommend an alternative method for detecting deforestation and degradation in this type of landscape, for example Radar satellite data (used successfully before in this type of landscape, e.g. in Mitchard *et al.* 2012[8]), aircraft LiDAR data (e.g. Boehm et al. 2013[9]), or ground-based observations, to provide more concrete results.

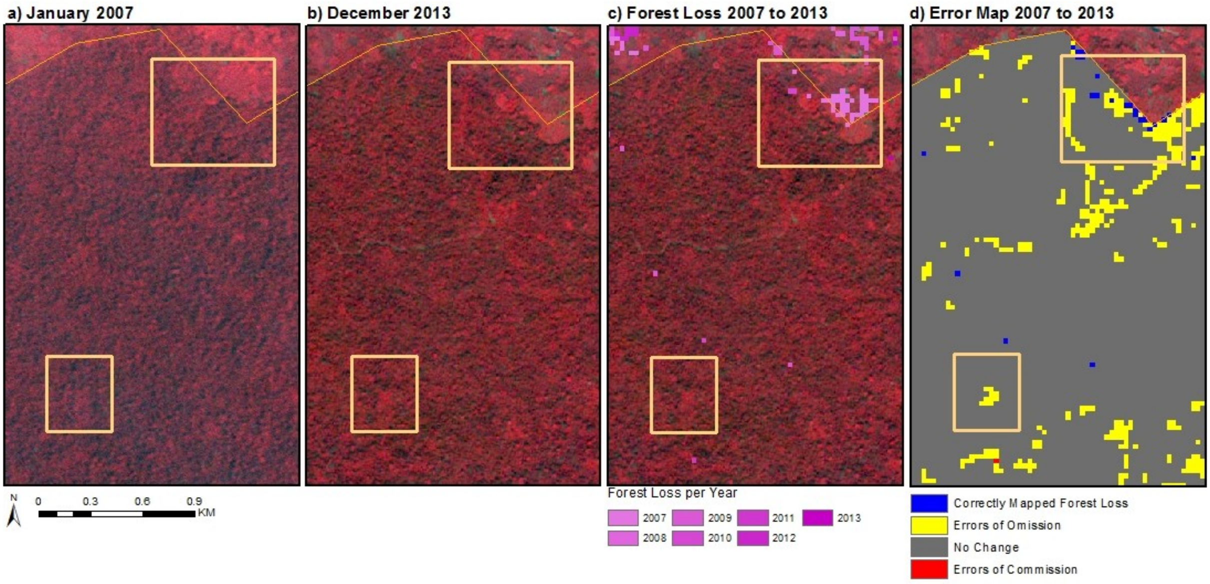
However, taken at face value these results appear to show a very significant underestimate of change for the UMD data, suggesting it is unsuitable for monitoring changes in these areas. Despite the caveats above, the high resolution image fragments shown in figures 15 and 16 look very real, and we are convinced examples such as these represent real changes on the ground. It should also be noted that, unlike in the Brazilian case, there was very strong agreement between the rates of commission and omission estimated by the manual and automated classifications here, adding confidence.

While most changes detected occurred near the boundaries of the forest blocks, suggesting encroachment for logging, agriculture or cacao development, worryingly some changes are observed far within the reserves (Figure 22). Ground verification would be required to confirm these areas as loss or degradation, as at such a small scale shadowing effects or climatic or seasonal changes can influence the interpretation of satellite imagery: but if real they suggest first that protection of these reserves is not currently effective, and secondly that the UMD data is not the correct tool to monitor these forests.

[8] Mitchard, E. T. A., P. Meir, C. M. Ryan, E. S. Woollen, M. Williams, L. E. Goodman, J. A. Mucavele, P. Watts, I. H. Woodhouse, and S. S. Saatchi. 2012. A novel application of satellite radar data: measuring carbon sequestration and detecting degradation in a community forestry project in Mozambique. Plant Ecology & Diversity.

[9] Boehm, H. D. V., V. Liesenberg, and S. H. Limin. 2013. Multi-Temporal Airborne LiDAR-Survey and Field Measurements of Tropical Peat Swamp Forest to Monitor Changes. Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of 6:1524-1530.

**Figure 22 - deforestation near and far from the end of a forest reserve in Ghana**




***Comparison of the high resolution data for (a) January 2007 and (b) December 2013 (c) deforested areas as mapped by the UMD Hansen Forest Loss per Year product for the years 2007 to 2013 showing both larger scale loss close to the boundaries of protected forests and smaller scale loss deep within the forest blocks. The error map for 2007 to 2013 (d) shows this area of loss as patchy errors of omission where clear deforestation and degradation can be identified.***

# 5 Conclusions

Summary points on the Brazil Cerrado study area are:

1. UMD forest loss data has overall good accuracy in detecting cerrado woodland conversion to large agricultural fields.

2. Errors consist of mis-allocation of deforestation into subsequent years (delayed pick-up). Therefore care is needed when interpreting changes in annual loss over short timeframes (maybe better over 3 to 5 years).

3. Some difficulty was experienced during the automated and manual classifications due to the heterogeneous nature and patchiness of the canopy cover in the study area. Differences in interpreter opinion on the canopy cover contributed to differences in the forest loss estimates shown in Tables 3 and 4.

4. The automated, and to a lesser extent the manual, classifications experienced some difficulty misclassifying agricultural changes as forest loss. Although the high resolution optical data adds useful texture and context, multi-temporal data adds useful information for separating agricultural changes from changes in forest cover.

5. Some question marks on shrubby lands and areas that are cut but allowed to regrow.

Summary points on Ghana study area are:

1. UMD data misses much forest change in this area: it is not seeing the same processes as the RapidEye data. This is due to both a resolution issue: it may be that the changes observed at 5 m resolution are just not visible in Landsat data; and due to the definition of forest and forest change. Calculations based on the UMD data in these areas of Ghana would therefore greatly underestimate forest disturbance and loss, therefore producing too high a value if used for reporting against the Hectares Indicator.

2. RapidEye data was not found suitable for determining forest change in the mosaic of forest/woodland/farmland that covers much of Ghana. Assessment of accuracy was only possible for the tall forest blocks.

3. Neither 5 m optical data nor UMD data is considered optimal for monitoring this landscape: we would recommend the use of active remote sensing products such as those from LiDAR or radar, or potentially very high resolution optical data (<1 m resolution). All of these options would involve high processing as well as data acquisition costs, so it may be that a sampling system, with significant ground component, would be necessary.