

Article

Historical aerial surveys map long-term changes of forest cover and structure in the central Congo Basin

Koen Hufkens^{1,2,*}, Thalès de Haulleville³, Elizabeth Kearsley¹, Kim Jacobsen^{1,3}, Hans Beeckman³, Piet Stoffelen⁴, Filip Vandelook⁴, Sofie Meeus⁴, Michael Amara⁵, Leen Van Hirtum⁵, Jan Van den Bulcke¹, Hans Verbeeck¹, Lisa Wingate²

¹ Ghent University, Ghent, Belgium;

² INRA, UMR ISPA, Villeneuve d'Ornon, France;

³ Royal Museum for Central Africa, Tervuren, Belgium;

⁴ Botanic Garden Meise, Meise, Belgium;

⁵ National Archives of Belgium, Brussels, Belgium;

* Correspondence: koen.hufkens@gmail.com

Version January 24, 2020 submitted to Remote Sens.



Abstract: Given the impact of tropical forest disturbances on atmospheric carbon emissions, biodiversity and ecosystem productivity accurate long term reporting of Land-Use and Land-Cover (LULC) change in the pre-satellite era (< 1972) is an imperative. Here, we use a combination of historical (1958) aerial photography and contemporary remote sensing data to map long-term changes in the extent and structure of the tropical forest surrounding Yangambi (DR Congo), in the central Congo Basin. Our study leveraged Structure-from-Motion and a novel Convolutional Neural Network based LULC classifier, using synthetic landscapes based image augmentation, to map historical forest cover across a large sub meter resolution orthomosaic (~82800 ha) geo-referenced to $\sim 4.7 \pm 4.3$ m. A comparison with contemporary LULC data showed a shift from previously highly regular industrial deforestation of large areas, to discrete smallholder farming clearing, increasing landscape fragmentation but also providing opportunities for substantial forest regrowth. Efforts to quantify long term canopy texture changes and a link to above ground carbon had limited to no success. Our analysis provides methods and insights into key spatial and temporal patterns of deforestation and reforestation at a multi-decadal scale, providing a historical context for past and ongoing forest research in the area.

Keywords: Aerial survey, data recovery, CNN, Deep Learning, SfM, Congo Basin

17 1. Introduction

18 Tropical ecosystem services are severely impacted by deforestation and forest degradation [1–3].
19 Not only does tropical forest Land-Use and Land-Cover Change (LULCC) constitute 10 to 15% of
20 the total global carbon emissions [4], changes in forest fragmentation affect the forest structure and
21 function [5–7]. Strong fragmentation effects decreases the number of large trees along forest edges
22 [8,9], while species composition and biodiversity are equally negatively affected [10–12]. Estimates
23 show that 31% of carbon emissions are caused by edge effects alone [6].

24 Accurate estimates of LULCC and forest canopy structure are therefore imperative to estimate
25 carbon emissions and other ecosystem services [1,2]. Remote sensing products have been key inputs in
26 LULCC assessments as they provide accurate spatial information to help estimate carbon emissions
27 [1,13]. More so, high resolution aerial images provide scientists tools to monitor forest extent, structure
28 and carbon emissions as canopy texture is linked to above ground biomass [14–16]. Yet, most of these
29 estimates are limited in time to recent decades [1,2,17,18].

30 Historical estimates of **Land-Use and Land-Cover (LULC)**, in the pre-satellite era (< 1972), exist
31 but generally rely on non-spatially explicit data (i.e socio-economic data) [2,17,19,20]. Efforts have
32 been made to use other geo-spatial data sources such as historical maps [21], declassified CORONA
33 satellite surveillance data across the US and central Brazil [22] as well as aerial surveys in post World
34 War II German [23]. Survey data across the African continent is less common, inaccessible or both.
35 Some studies do exist, as Buitenwerf *et al.* [24] and Hudak and Wessman [25] used aerial survey images
36 to map vegetation changes in South African savannas, whilst Frankl *et al.* [26] and Nyssen *et al.* [27]
37 mapped the Ethiopian highlands of 1930's.

38 Across the central Congo Basin most of these historical images were collected within the context
39 of national cartographic efforts by the “Institut Géographique du Congo Belge” in Kinshasa (then
40 Léopoldville), DR Congo. Despite the existence of large archives of aerial survey imagery (Figure 1,
41 Appendix Figure 3), as of yet, no studies have valorized these data. The lack of a consistent valorization
42 effort is unfortunate as the African rainforest is the second largest on Earth and covers ~630 million ha,
43 representing up to 66 Pg of carbon storage [28], and currently loses forest at an increasing pace [29].
44 Given the impact of LULCC on the structure and functioning of central African tropical forests, and
45 their influence on both carbon dynamics [30] and biodiversity [12], accurate long term reporting of
46 historical forest cover warrants more attention [21].

47 Here, we use a combination of historical aerial photography and contemporary remote sensing
48 data to map long-term changes in the extent and structure of the tropical forest surrounding Yangambi
49 (DR Congo), in the central Congo Basin, effectively linking the start of the anthropocene [31] with
50 current assessments. Yangambi was, and remains, a focal center of forest and agricultural research and

51 development in the central Congo Basin. Past research in the region allows for thorough assessment of
52 land-use and land-cover change from a multi-disciplinary point of view, confronting us with complex
53 deforestation and land-use patterns.

54 We leverage Structure-from-Motion to generate a large orthomosaic of historical imagery and
55 develop a Convolutional Neural Network based forest cover mapping approach, based upon a
56 semi-supervised generated dataset extensively leveraging data augmentation. Our methodology
57 aims to provides a historical insight into important LULCC spatial patterns in Yangambi, such as
58 fragmentation and edge complexity. We further contextualize the influence of changes in the forest's life
59 history on past and current research into Above Ground Carbon (AGC) storage [30] and biodiversity
60 [12] in the central Congo Basin. Our fast scalable mapping approach for historical aerial survey data,
61 using limited supervised input, would further support long-term land-use and land-cover change
62 analysis across the central Congo Basin.

63 2. Methods

64 2.1. Historical data acquisition

65 Data for the central Congo Basin region, surrounding Kisangani, were collected in several flights
66 during the dry season of 1958 and 1959 (from 8/01/1958 to 20/02/1958 and from 28/12/1958
67 to 9/01/1959 respectively, see Appendix Figure 1) to generate topographic maps of the area,
68 supervised by the "Institut Géographique du Congo Belge" in Kinshasa, DR Congo (then Léopoldville).
69 Black-and-White infrared images (0.4 - 0.9 μm) were gathered along flight paths running mostly from
70 west to east, between **09:00 and 11:00** local time. Along a flight path continuous images were taken
71 using a Wild Heerburg RC5a (currently Leica Geosystems) with an Avigon lens assembly (114.83mm
72 / f 5.6, with a 90°view angle) resulting in square photo negatives of 180 by 180 mm. Flights were
73 flown at an average absolute altitude of ~5200 m above sea level, covering roughly 18 530 km² at an
74 approximate scale of 1/40 000. The use of the integrated autograph system ensured timely acquisition
75 of pictures with a precise overlap (~1/3) between images. This large overlap between images together
76 with flight parameters would allow post-processing, using stereographs, to create accurate topographic
77 maps. Original data from this campaign are stored in the [Royal Museum for Central Africa](#) in Tervuren,
78 Belgium (Figure 1).

79 2.2. Site selection

80 We prioritised flight paths and images that contained current day permanent sampling plots,
81 larger protected areas, and past agricultural and forest research facilities (Figure 2). This selection
82 provides a comprehensive mapping of the Yangambi area and the life history of the forest surrounding

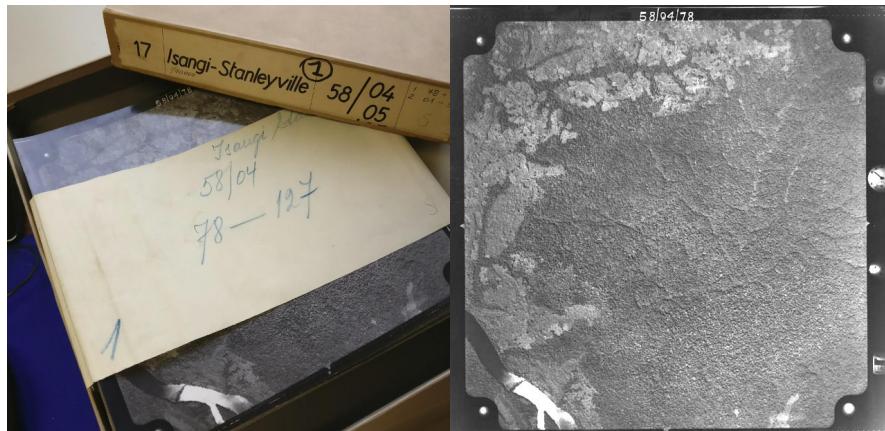


Figure 1. A box of historical aerial photographs (left) and a single aerial photograph (right) showing part of the Congo river. Note the meta-data provided in the right side margin of the image, such as acquisition time and flight height.

83 it. Thereafter, we selected flight paths 1 through 11 for digitization. From this larger dataset of 334
 84 images we selected 74 survey images for orthomosaic compositing and further analysis. All the
 85 selected images stem from the flight campaign made during January and February of 1958. The area
 86 includes the Yangambi village, 20 contemporary permanent sampling plots [30], past and present
 87 agricultural experimental plots [32] and large sections of the Yangambi **UNESCO Man and Biosphere**
 88 **reserve** surrounding to the west and east of the village. Although not formally mosaicked we provided
 89 a full dataset of pre-processed images using the cropping and normalization routines described below.
 90 The latter data was not used in subsequent LULCC analysis, but has been archived and made available
 91 to the public separately (see code & data availability statement below).

92 2.3. Digitization and data processing

93 All selected images, covering the Yangambi area, were contact prints as original negatives of the
 94 prints were not available. Images were scanned at a resolution exceeding their original resolution
 95 (or grain) at the maximal physical resolution of an Epson A3 flatbed scanner (i.e. 2400 dpi or 160 MP
 96 per image) and saved as lossless tiff images. Data were normalized using contrast limited histogram
 97 equalization [33] with a window size of 32 and a clip limit of 1.5. Fiduciary marks were used to rectify
 98 and downsample the images into square 7700x7700 pixel images (~1200 dpi, 81 MP). This resulted in a
 99 dataset with digital images at a resolution that remained above the visible grain of the photographs,
 100 whilst the reduced image size facilitated easier file handling and processing speed.

101 Data was processed into a georeferenced orthomosaic using a Structure-from-motion (SfM, Ullman
 102 [34]) approach implemented in **Agisoft Metashape** version 1.5.2 (Agisoft LLC, St. Petersburg, Russia).
 103 An orthomosaic corrects remote sensing data to represent a perfectly downward looking image, free
 104 from perspective distortions due to topography and camera tilt. Using the SfM technique features,

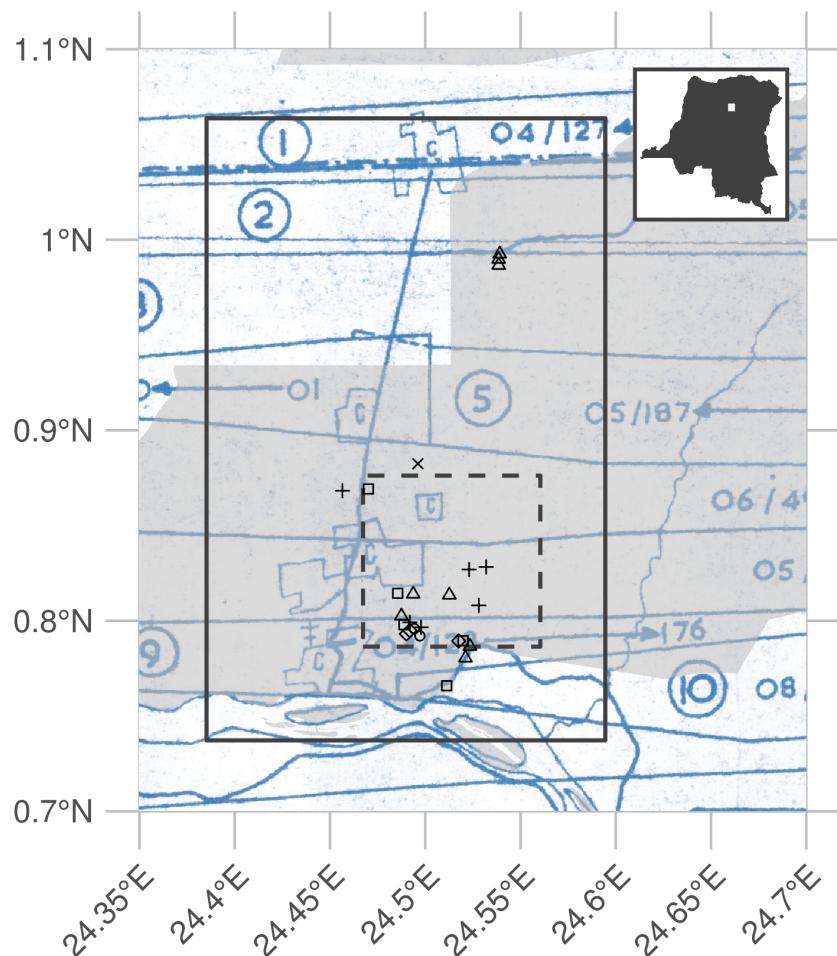


Figure 2. Overview of the historical flight paths during aerial photo acquisition and ancillary data used in this study. The bounding box of the orthomosaic data presented in this study is shown as a rectangle (23x36 km). The outline of a recent high-resolution Geo-eye panchromatic image is shown as a dashed dark grey rectangle (10x10km). The location of various permanent sampling plots are shown as x, +, and open squares and triangles for the mixed, mono-dominant and edge plots respectively. The grey polygon delineates the current day Yangambi Man and Biosphere reserve. The inset, top right, situates the greater Yangambi region (white rectangle) with the DR Congo. The full flight plan and details are shown in Appendix Figure 1 and 2

105 areas in images with a large degree of similarity, are matched across various images to reconstruct
106 a three dimensional scene (topography) from two-dimensional image sequences. During the SfM
107 analysis we masked clouds, glare or large water bodies such as the Congo river.

108 We calculated the orthomosaic using a low resolution point cloud and digital elevation map
109 (DEM). Additional ground control points were provided to assist in the referencing of image and
110 constrain the optimization routine used in the SfM algorithm. Ground control points consisted of
111 permanent structures which could be verified in both old and new aerial imagery (i.e. ESRI World
112 Imagery) and consisted of corner points of build structures (e.g. a building, bridge or swimming
113 pool etc.). Although clouds were removed during the SfM routine we did not mask all clouds in the
114 final orthomosaic to maximize forest coverage. The final scene was cropped, to provide consistent wall
115 to wall coverage of the reconstructed scene. The orthomosaic was exported as a geotiff for further
116 georeferencing in QGIS [35] using the georeferencer plugin (version 3.1.9) and additional ESRI World
117 Imagery high resolution reference data. We used 3rd degree polynomial and 16 ground control points
118 to correct the final image. Ground control points, raw image data and final processed image are
119 provided in addition to measures of uncertainty such as mean and median error across all ground
120 control points. All subsequent analysis are executed on the final geo-referenced orthomosaic or subsets
121 of it.

122 2.4. Land-Use and Land-Cover Change

123 2.4.1. Classifying Land-Use and Land-Cover

124 Model training

125 We automatically delineated all natural forest in the historical data, thus excluding tree plantations,
126 thinned or deteriorated forest stands which showed visible canopy cover loss, fields and buildings.
127 We used the Unet Convolutional Neural Net (CNN, Ronneberger *et al.* [36]) architecture implemented
128 in Keras [37] with an efficientnetb3 pre-processing backbone [38] running on TensorFlow [39] to train
129 a binary classifier (i.e. forest or non-forested). Training data were collected from the orthomosaic by
130 randomly selecting 513 pixel square tiles from locations within homogeneous forested or non-forested
131 polygons in the historical orthomosaic (Figure 5). Separate polygons were selected for training,
132 testing and validation purposes. Validation polygons were sampled 300 times, while both testing
133 and validation polygons were sampled at 100 random locations. Tiles extracted from locations
134 close to the polygon border at times contained mixed cover types. Tiles with mixed cover types
135 were removed from the list of source tiles (Table 1). Homogeneous source tiles were combined in
136 synthetic landscapes using a random gaussian field based binary mask (Figure 3). We generated 5000

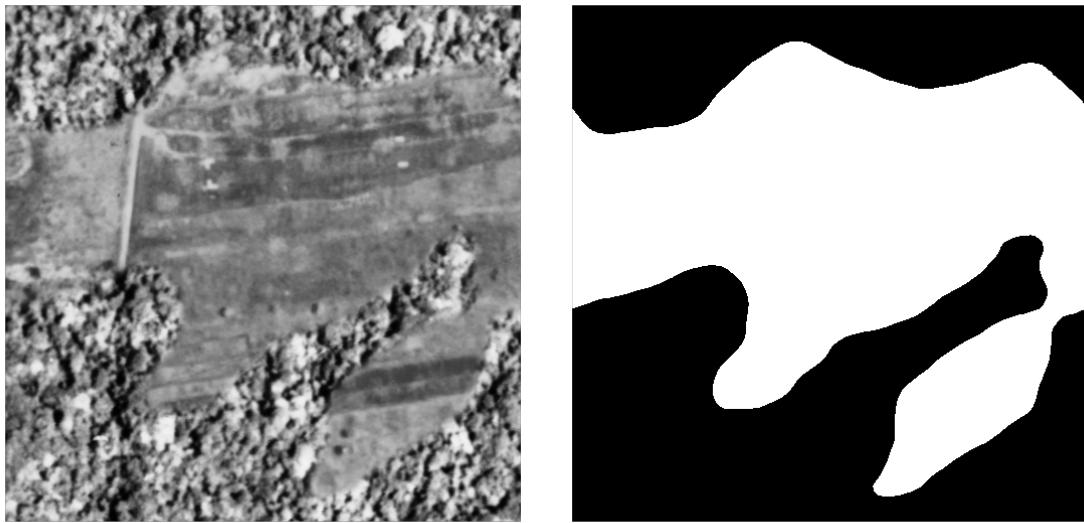


Figure 3. An example synthetic landscape, combining homogeneous forest and non-forest images into a patchy landscape using random gaussian field based masks. The left panel shows a combined synthetic landscape, while the right panel shows the corresponding forest (black) and non-forest (white) labels.

137 synthetic landscapes for training, while 500 landscapes were generated for both the validation and the
 138 testing datasets for a total of 6000 synthetic landscapes. In order to limit stitch line misclassifications,
 139 along the seams of mosaicked images, we created synthetic landscapes with different forest tiles
 140 to mimick forest texture transitions. We applied this technique to 10% of the generated synthetic
 141 landscapes (**across training, validation and testing data**).

Table 1. Number of source tiles used for the generation of synthetic landscapes.

	training	testing	validation
forest	300	100	100
non-forest	294	92	84

142 The CNN model was trained for 100 epochs with a batch size of 30 using Adam optimization [40],
 143 maximizing the Intersect-over-Union (IoU) using Sørensen–Dice [41] and categorical cross-entropy
 144 loss functions. Data augmentation included random cropping to 320 pixel squares, random orientation,
 145 scaling, perspective, contrast and brightness shifts and image blurring. The optimized model was used
 146 to classify the complete orthomosaic using a moving window approach with a step size of 110 pixels
 147 and a majority vote (> 50% agreement) across overlapping areas to limit segmentation edge effects.
 148 **In addition, we provide raw pixel level classification agreement data for quality control purposes**
 149 **(see Data availability below)**. We refer to Figure 4 for a synoptic overview of the full deep learning
 150 workflow.

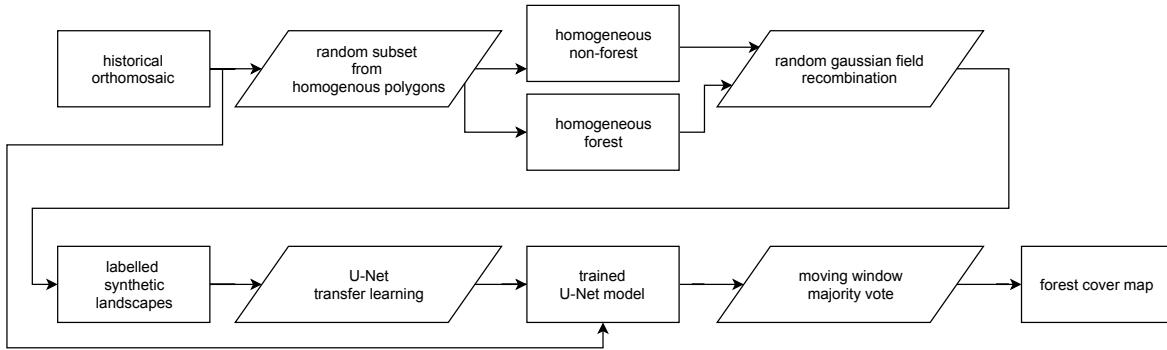


Figure 4. A diagram of the deep learning workflow followed in training a binary forest / non-forest cover convolutional neural net U-Net model to generate our forest cover map.

151 *Model validation*

152 We report the CNN accuracy based upon the IoU of our out-of-sample validation dataset
 153 of synthetic landscapes. In addition, we report confusion matrices for all pixels across the
 154 homogeneous validation polygons, as well as the training and testing polygons (see Figure
 155 5). Furthermore, we used the first acquisition of a recent pan-chromatic Geo-Eye 1 stereo pair
 156 (Geo-Eye, Thornton, Colorado, U.S.A., order 737537, 2011-11-11 08:55 GMT or 09:55 local time)
 157 to classify and assess the robustness of the CNN algorithm on contemporary remote sensing
 158 imagery with similar spectral and spatial characteristics. We used the Global Forest Change
 159 version 1.6 (GFC, tile 10N-020E) [1] map data, capturing forest loss up to 2011, to quantify accuracy
 160 on downsampled CNN Geo-Eye classification results. Once more, we report the confusion matrix
 161 of between the GFC and CNN derived forest cover maps, masking clouds and cloud shadows. To
 162 summarize confusion matrices we report accuracy as:

$$163 \quad Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)}$$

164 in which TP, TN, FP, FN are True Positive, True Negative, False Positive and False Negative,
 165 respectively.

166 2.4.2. Characterizing long term change

167 To map long term LULCC in the Yangambi region we used the contemporary Global Forest
 168 Change version 1.6 (GFC, tile 10N-020E) map data [1]. Using the GFC data we calculated the latest
 169 state of the forest with respect to the conditions at the start of 1958, 60 years earlier. In our analysis we
 170 only included GFC pixels which recorded no forest loss throughout the whole 2000 - 2018 period. As
 171 the resolution of the historical forest classification exceeds that of the GFC map we downsampled our
 172 historical forest cover data to 30 m GFC resolution **using a nearest neighbour approach**. We masked
 173 out all water bodies using the Global Forest Change survey data mask layer, and limited the analysis
 174 to the right bank of the Congo river. We provide summary statistics of historical and contemporary

175 forest loss and regrowth. We map persistent forest loss after 1958, regrowth after loss in 1958, recent
176 forest loss and long term (stable) forest cover.

177 2.4.3. Landscape fragmentation & Above Ground Carbon estimates

178 To quantify changes in the structure of forest cover and its disturbances we used spatial pattern
179 analysis landscape fragmentation metrics [42]. We report the ratio of edge to area and the fractal
180 index to quantify landscape complexity of forest disturbances. Statistics were calculated for all forest
181 disturbance patches larger than 1 ha and smaller than the 95th percentile of the patch size distribution
182 using the R package landscapemetrics [43]. We provide mean and standard deviation on edge, area,
183 their ratio and fractal dimension for both the historical and contemporary Hansen *et al.* [1] forest cover
184 maps.

185 We estimated above ground carbon (AGC) losses and gains over time using plot based averages
186 of recent inventory data at permanent sampling plots in the area (Figure 2). We refer to Kearsley *et al.*
187 [30] for the survey method and allometric relations used to scale the survey data. Unlike standard
188 square 1 ha plots, edge plots were set back 200 m from forest edges and were 50x200 m, with the
189 50 m side of the plot along the forest edge and continuing 200 m into the forest (Appendix Table 1).
190 We further confirmed that forest edge plots did not show a significantly different AGC compared to
191 those of non-edged / mixed forest plots (Mann Whitney U test, $p < 0.05$). Thus it was not necessary to
192 explicitly quantify changes in AGC caused by edge effects. Moreover, we used the mean values of the
193 mixed forest as representative for potential AGC losses. Despite the challenges inherent in quantifying
194 AGC for forest edges we mapped the total extent of the edges in the contemporary landscape. To
195 align our landscape analysis with exploratory analysis of the survey data we used a buffer of 200 m to
196 estimate the extent of forest edges and patches, up to the location of forest edge plots.

197 Surveys of old plantations show a large variation in AGC, depending on age and the crop type.
198 For example, the AGC values varied from 168.67 to 86.55 Mg C ha⁻¹, for *Hevea brasiliensis* (rubber
199 tree) and *Elaeis guineensis* (oil palm) plots respectively (Bustillo *et al.* [44], personal communications).
200 These higher values are in line with the mixed AGC estimates in the area, while the palm plantations
201 resemble old-regrowth values (81.8 Mg C ha⁻¹, see Appendix Table 1). We therefore use both the
202 estimates of old-regrowth and mixed forest to estimate AGC for regrowth. We did not have sufficient
203 data to account for individual changes in AGC across plantations.

204 2.5. Canopy structure & FOTO texture analysis

205 We compared the structure of the canopy both visually and using Fourier Transform Textural
206 Ordination (FOTO, Couturon [45]). FOTO uses a principal component analysis (PCA) on radially
207 averaged 2D Fourier spectra to characterize canopy (image) texture. The FOTO technique was first

described by Couturon [45] to quantify canopy stucture in relation to biomass and biodiversity, and can be used across multiple scenes using normalization [16].

We used the standard FOTO methodology with fixed zones, instead of the moving window approach. The window size was set to the same size (187 pixels or ~150 m) as used in the moving window analysis above. To account for illumination differences between the two scenes we applied histogram matching. No global normalization was applied, as the scene was processed as a whole. PC values from this analysis for all permanent sampling plots in both image scences were extracted using a buffer with a radius of 50 m.

For both site based and scene analysis we correlated PC values with permanent sample plot inventory data such as stem density, above ground biomass and tree species richness. Additional comparisons are made between contemporary Geo-Eye data and the historical orthomosaic derived PC values. Due to the few available permanent sampling plots in both scenes we used a non-parametric paired signed rank (Wilcoxon) test [46] to determine differences between the PC values of the Geo-Eye and historical orthomosaic image scenes across mono-dominant and mixed forest types. In all analysis, mono-dominant site 4 was removed from the analysis due to cloud contamination.

3. Results

3.1. Orthomosaic construction

Our analysis provides a first spatially explicity historical composite of aerial survey images **in support of mapping land-use and land-cover** within the Congo Basin. The use of high resolution historical images combined with SfM image processing techniques allowed us to mosaic old imagery across a large extent. The final orthomosaic composition of the Yangambi region **provided** an image scene covering approximately 733 million pixels across **~82800 ha** (~23x36 km, Figure 2). The overall accuracy of the SfM orthomosaic composition was 0.88 m/pixel using the sparse cloud DEM for corrections at 45.8 m/pixel. The resulting georeferenced scene had a spatial accuracy of approximately **23 m**. Further georeferencing outside the SfM workflow reduced the mean error at the ground control points to 5.3 ± 4.9 px ($\sim 4.7 \pm 4.3$ m), with a median error of 2.9 px (2.6 m). The orthomosaic served as input for all subsequent LULCC analysis with all derived maps provided with the manuscript repository (see data & code availability statements below).

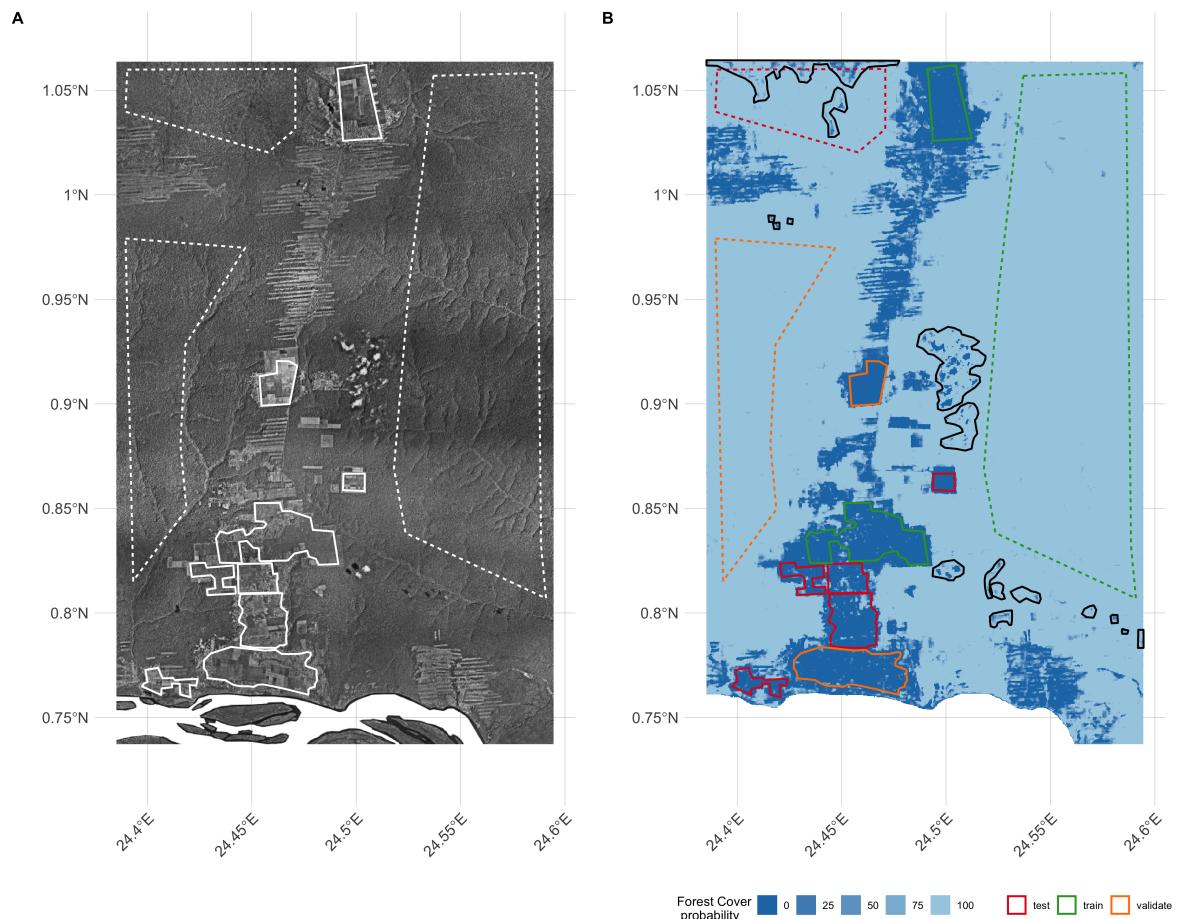


Figure 5. Overview of the final orthomosaic of the greater Yangambi region (A) as well as the forest cover classification uncertainty (B) used to generate the final Land Land Cover map (see Figure 6). Homogeneous polygons used in the generation of the synthetic landscape for Convolutional Neural Network training, testing and validation are marked as dashed and full lines for forest and non-forest regions, respectively. Training, testing and validation regions are denoted in panel B in green, red and orange, respectively. Black polygon outlines denote cloud and image stitch line regions which were manually excluded from analysis, but retained in calculation of validation statistics (see Table 2).

²³⁶ *3.2. Land-use and Land-Cover Classification*

²³⁷ *3.2.1. CNN model validation*

²³⁸ **The CNN deep learning classifier reached an Intersection-over-Union accuracy of 97% on the**
²³⁹ **detection of disturbed forest in the out-of-sample (validation) synthetic landscape data. Using all**
²⁴⁰ **pixels within the validation polygons (Figure 5) showed a similar accuracy value of ~98%. Using all**
²⁴¹ **polygons across the scene, including those used in the generation of testing and training synthetic**
²⁴² **landscapes, increased the accuracy to ~99% (Table 2). A comparison with recent pan-chromatic**
²⁴³ **Geo-Eye data shows good agreement, with an accuracy of ~87% across all pixels, between the**
²⁴⁴ **landsat based GFC data and downscaled CNN results (Table 2 and Figure 7).**

Table 2. Confusion matrix showing % agreement between forest / non-forest classes using a Convolutional Neural Network (CNN) across previously selected homogenous areas.

CNN	Geo-Eye (full scene)		Aerial Survey (all polygons)		Aerial Survey (validation polygons)	
	non-forest	forest	non-forest	forest	non-forest	forest
non-forest	10.17	8.55	97.4	0.19	98.59	0.0
forest	3.75	77.52	2.6	99.81	1.41	100.0
Accuracy	87.70		98.61		99.3	

²⁴⁵ *3.3. Long term changes in LULC and Above Ground Carbon*

²⁴⁶ Scaling our classifier to the whole historical orthomosaic we detected ~16200 ha (or ~20% of the
²⁴⁷ scene) of disturbed forests. A large fraction of the disturbed area was restored in the period between
²⁴⁸ the two acquisition periods. In total, 9918 ha, or little over half of the affected forest was restored
²⁴⁹ (Figure 5C-D, dark blue). Recent deforested areas, as registered through satellite remote sensing,
²⁵⁰ approximate 8776 ha (Table 3, Figure 6 - light green).

²⁵¹ Recent deforestation follows a distinctly different pattern compared to historical patterns.
²⁵² Historical deforestation showed a classical fishbone pattern for forest clearing with very sharp edges,
²⁵³ while current patterns are patchy and ad-hoc (figure 6C, dark blue and green colours respectively).
²⁵⁴ These differences are reflected in the analysis of landscape metrics of forest loss. Between the historical
²⁵⁵ and contemporary LULCC maps we see an increase in small disturbances, as indicated by the
²⁵⁶ decreasing area of the mean patch size, down to $\sim 1.86 \pm 0.75$ ha from $\sim 5.25 \pm 5.02$ ha historically.
²⁵⁷ Perimeter lengths were longer historically, at 1451 ± 943 m, compared to contemporary landscapes
²⁵⁸ $\sim 921 \pm 362$ m (Table 3). This shift in perimeter area ratio led to a similar change in the fractal index,
²⁵⁹ slightly increasing in value to 1.1 ± 0.05 from 1.09 ± 0.04 over time. Values closer to a fractal index of
²⁶⁰ 2 suggest a more complex (fragmented) landscape.

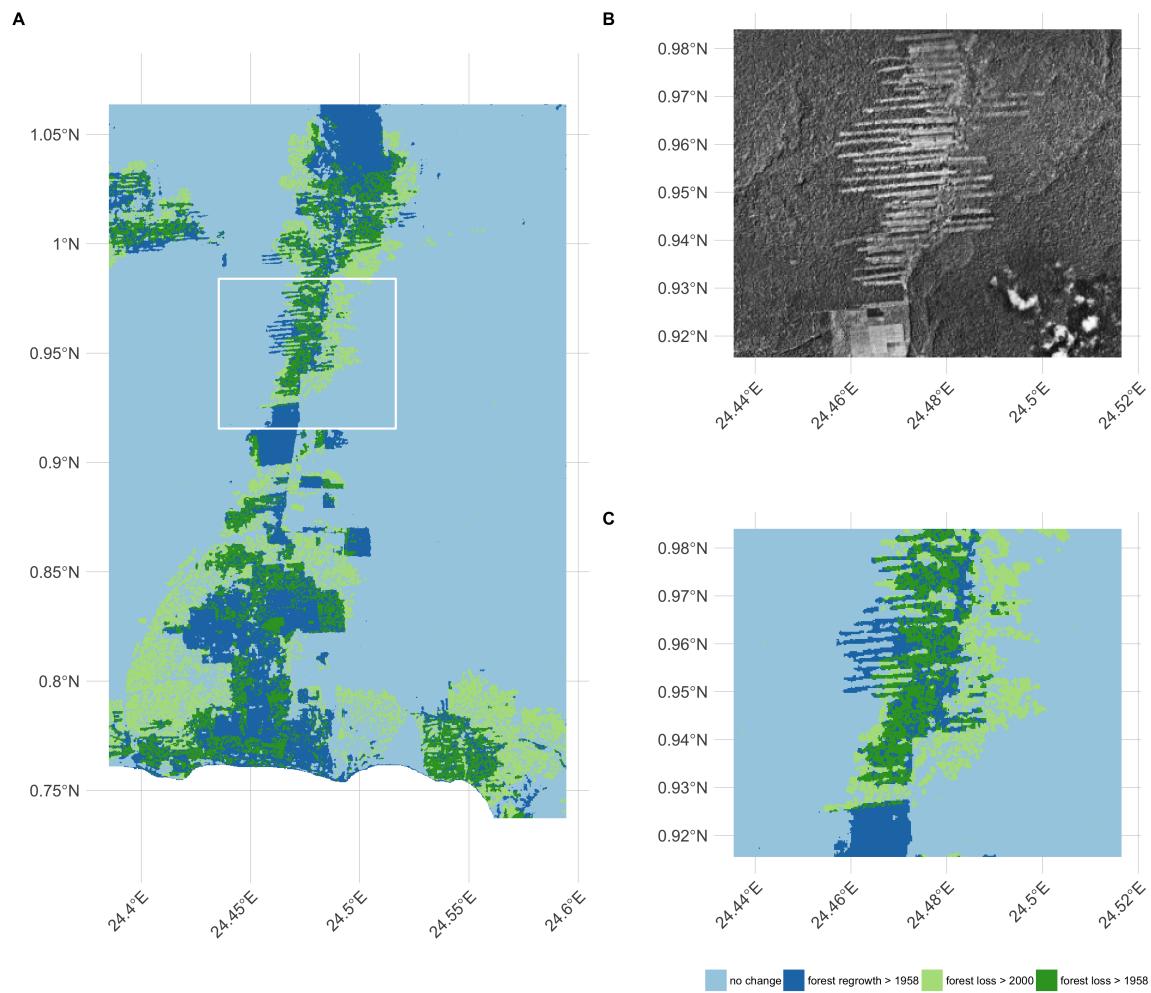


Figure 6. Overview of the final Land Use Land Cover Change map, a detailed inset of both the underlying orthomosaic (B) and the derived land use land cover change map displayed as the difference between the Convolutional Neural Network based classification orthomosaic and the recent Landsat based forest cover map by Hansen et al. 2013 (C).

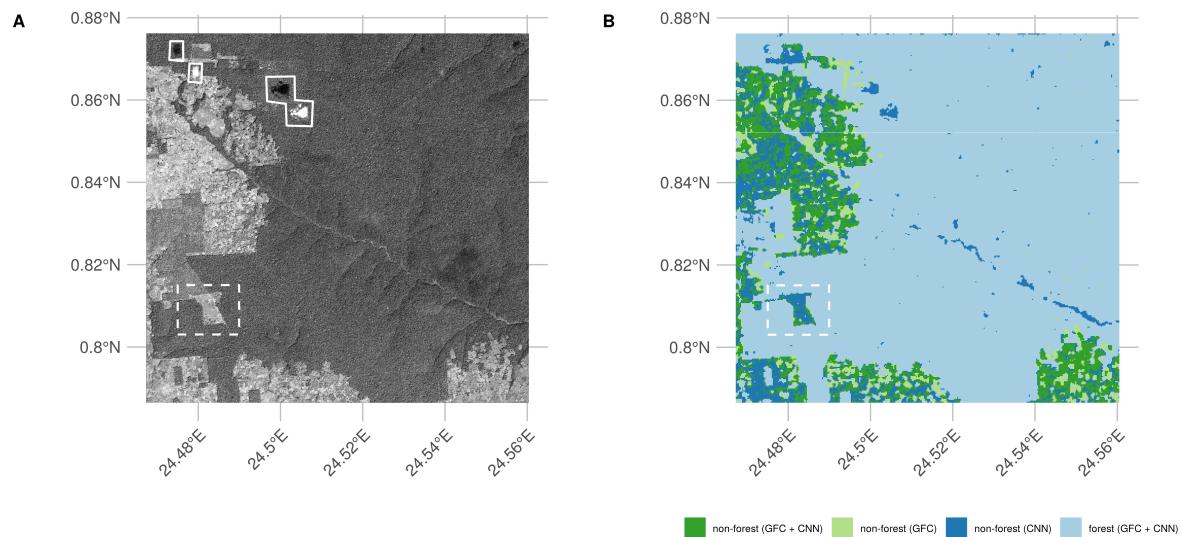


Figure 7. Convolutional Neural Network (CNN) based forest cover classification results (B) as run on a recent (2011) Geo-Eye panchromatic image (A). We show the difference between the Convolutional Neural Network based classification and a recent Landsat based forest cover map by Hansen et al. 2013. Full white outlines denote cloud contamination, the dashed rectangle shows a location where the CNN outperforms the Landsat based forest classification.

261 A comparison of forest edge plots with mixed forest plots showed no significant difference in
 262 AGC, or other reported values such as species richness, basal area or stem density (Mann Whitney U
 263 test, $p < 0.05$). Edge influence did not extend beyond 200 m from a forest edge, but still represented an
 264 area of **13151 ha**.

265 Changes in both land-use and land-cover led to concomitant changes in AGC stocks. Recovery
 266 throughout the region was characterized for patches of forest and plantations. Assuming high density
 267 stands, based on previous work, this could amount to a carbon gains of 1592 Gg C across our study
 268 area, offsetting more recent losses of approximately 1408 Gg C. On the other hand, at the low end, if
 269 we assume a lower carbon density of $81.8 \text{ Mg C ha}^{-1}$ this would result in a total carbon gain of 811 Gg
 270 C. Using our approach results indicate that overall deforestation around Yangambi has resulted in a
 271 loss of $\sim 2416 \text{ Gg C}$ in AGC stocks.

Table 3. Land use land cover change statistics of forest cover around Yangambi in the central Congo Basin. The data evaluates a difference between a historical (1958) aerial photography based survey and the Hansen et al. 2013 based satellite remote sensing data. Spatial coverage statistics are provided in square kilometers (km) and hectares (ha), rounded to the nearest integer as well as Above Ground Carbon (AGC) scaled using recent survey measurements.

	AGC	
	ha	Gg C ha^{-1}
Forest	68455	
- of which edges	13151	
Regrowth > 1958 loss	9918	811 - 1592
Loss > 2000	8776	1408
Loss > 1958 (permanent)	6282	1008

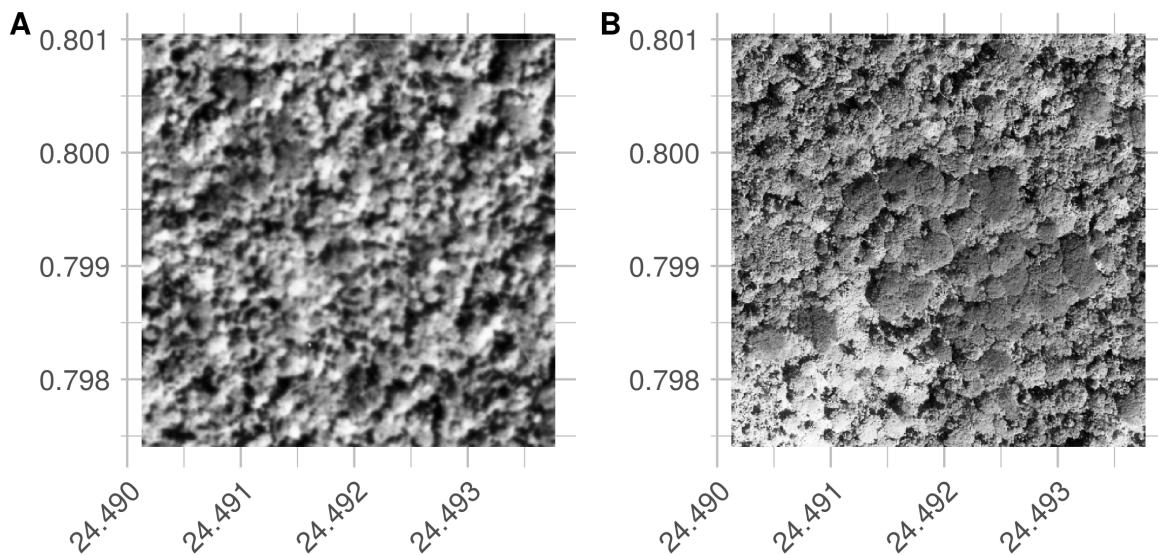


Figure 8. Visual comparison between a historical (A) and contemporary (B) permanent sampling plot. The site is currently listed as a mono-dominant *Brachystegia laurentii* stand. Note the structural differences with a 'coarser' canopy structure in the historical image compared to the more closed contemporary stand.

Table 4. Landscape metrics for historical and contemporary deforestation patterns. We report patch perimeter and area, their ratio and fractal dimension. Values are reported as mean \pm standard deviation, across all deforestation patches.

	perimeter (m)	area (ha)	ratio (m^{-1})	fractal dimension
historical	1451.58 ± 943.27	5.25 ± 5.02	0.03 ± 0.01	1.09 ± 0.04
contemporary	921.74 ± 362.59	1.86 ± 0.75	0.05 ± 0.01	1.1 ± 0.05

272 3.4. Canopy structure & FOTO texture analysis

273 Visual interpretation of the scenes provide evidence that most locations do not change dramatically
 274 with respect to canopy composition, except for the large areas of disturbances in contemporary fallow
 275 or young-regrowth plots. One marked difference is noted in the mono-dominant plot 6 (Appendix
 276 Table 1). Here, the current mono-dominant *Brachystegia laurentii* is a recent development, changing
 277 the canopy structure visibly during the last half century (Figure 8). The previous varied canopy
 278 structure gave way to a more dense and uniform canopy. This is reflected in a change of the FOTO PC
 279 value from 0.19 historically to its current value of 0.54. This historical value is similar to the mean of
 280 contemporary mono-dominant stands of *Gilbertiodendron dewevrei* with PC averaging 0.34 ± 0.1 , and is
 281 only slightly higher than historical values for a mixed forest (0.18 ± 0.08). The reverse pattern is seen
 282 in the contemporary PC values. Here, the value of 0.54 exceeds those of most mono-dominant stands
 283 (0.35 ± 0.08), and is even further removed from the values noted for mixed forests (0.12 ± 0.03).

Using only small subsets around existing permanent sampling plots we show distinct differences between forest types, with PC values in both historical and contemporary imagery markedly higher for the mono-dominant forest types compared to all others (**Appendix Figure 4**). Provided that the young-regrowth and fallow permanent sampling plots have seen recent disturbance the Wilcoxon signed rank test on the mixed and mono-dominant plots between the historical and contemporary PC values did not show a significant difference ($p > 0.05$). Similarly, no significant different using PC values extracted from the whole scene analysis was noted ($p > 0.05$). Any relationships between contemporary Geo-Eye data and permanent sampling plot measurements of Above Ground Carbon, stem density and species richness were non-significant ($p > 0.05$, Appendix **Figures 5-7**).

Furthermore, visual inspection of the scene wide analysis suggests historical scenes do not show landscape wide canopy features (**Figure 9 A-B**), unlike the contemporary scene (**Figure 9 C-D**). Where the FOTO algorithm picks up landscape features such as changes in texture across the contemporary Geo-Eye scene (e.g. the river valley as a diagonal line in **Figure 9D**), however, no corresponding landscape patterns are found by the FOTO algorithm in the historical orthomosaic.

4. Discussion

Finely grained spatial data sources, such as remote sensing imagery, are rare before the satellite era (<1972). This lack of data limits our understanding of how forest structure has varied over longer time periods in remote areas. Long term assessment can be extended by using large inventories of historical aerial survey data [22,23,47]. Despite the difficulties in recovering this data and its limitations, such as invisible disturbances [48], remote sensing generally remains the best way to map and quantify LULCC [2]. In our study we used novel numerical remote sensing techniques to valorize, for the first time, historical remote sensing data in order to quantify (long term) land-use and land-cover change and canopy structural properties in the central Congo Basin. Despite these successes some methodological and research considerations remain.

4.1. Methodological considerations

4.1.1. Data recovery challenges

In our study the archive data recovered was limited to contact prints and therefore did not represent the true resolution of the original negative. In addition, analogue photography clearly produces a distinct softness compared to digital imagery (**Figure 8**). Despite favourable nadir image acquisitions [49] image softness combined with illumination effects between flight paths, and the self-similar nature of vast canopy expanses [50–52], limited our ability to provide wall-to-wall coverage of the entire dataset containing 334 images. Few man made features in the scenes also

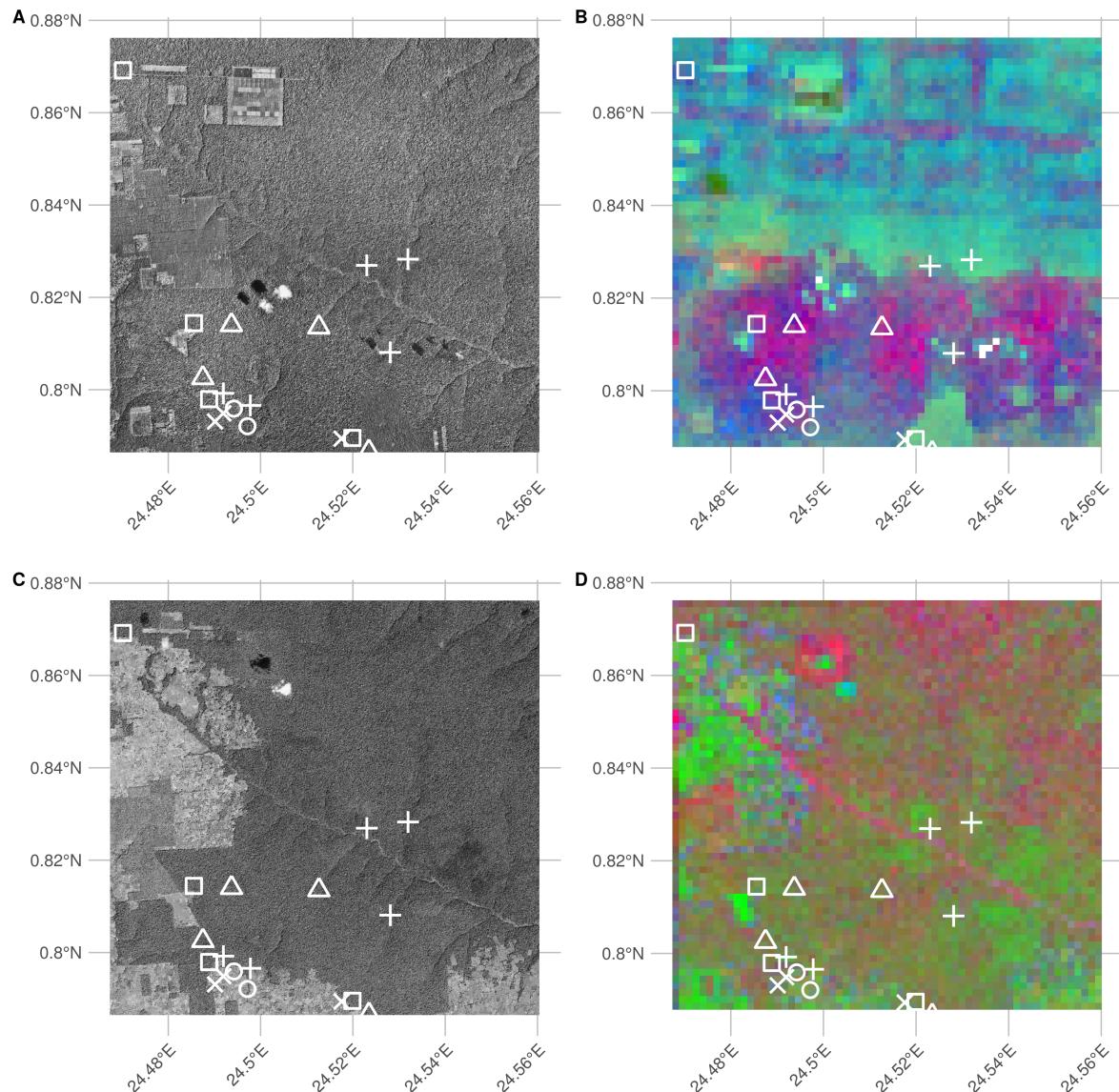


Figure 9. RGB visualizations of the first three principal components of scene wide FOTO texture analysis of historical and current (Geo-Eye) imagery. Current permanent sampling plots of mono-dominant, mixed, fallow and young (edge) forest plots are marked with open triangles, open circles, open squares and crosses, respectively.

316 made georeferencing challenging. Although the village of Yangambi provided a range of buildings
317 as (hard-edge) references, other areas within the central Congo Basin might have fewer permanent
318 structures and would require the use of soft-edged landscape features (e.g. trees, river outflows).
319 Research has shown that soft-edged features can help georeference scenes even when containing few
320 man-made features [53]. Our two step georeferencing approach resulted in a referencing accuracy of
321 $\sim 4.7 \pm 4.3$ m across reference points. However, it shoud be noted that referencing accuracy of the final
322 scene is less constrained toward the edges of the scene.

323 4.1.2. LULC classification & validation

324 When classifying the orthomosaic into forest and non-forest states we favoured a deep
325 learning supervised classification using a CNN over manual segmentation to guarantee an
326 “apples-to-apples” comparison between the historical and the contemporary GFC forest cover map
327 used in our analysis. We acknowledge that both the CNN and GFC land-use and land-cover maps
328 use different underlying features, i.e. spatial or spectral data, yet attain a similarly high accuracy
329 of up to 99% [1]. More so, when validating our CNN classifier against GFC data (Figure 7) for a
330 contemporary high resolution Geo-Eye panchromatic image we reach an accuracy of $\sim 87\%$, despite
331 a time difference of almost 60 years. Visual inspection of the classification data in Figure 7 suggests
332 that the GFC map more often than not classifies non-forest areas as forest. Actual classification
333 accuracy of our algorithm might therefore be higher than our reported value.

334 4.1.3. Scaling opportunities

335 Our approach uses broadly defined homogeneous polygons to construct a balanced dataset of
336 synthetic landscapes. The methodology is analoguous to the use of sparse labelling as used by
337 Buscombe and Ritchie [54] and contrasts with the standard methodologies which generally extract
338 pixel (windows) [22] or delineate land cover classes [23] to drive a classifier or analysis. More
339 so, the use of heavy image augmentation during model training sidesteps texture representation
340 issues which affect classification of image scenes with inconsistent illumination or sharpness [25]
341 or ad-hoc feature engineering [22]. The use of synthetic landscapes allowed us to account for most,
342 but not all, of the variability within our orthomosaic. Our analysis has shown that despite being
343 trained on historical data our model could map contemporary forest cover in remote sensing data
344 with similar spatial and spectral characteristics (Figure 7), suggesting that the classifier consistently
345 works across both space and time. We acknowledge that the use of synthetic landscapes is limited
346 by the available homogeneous areas to sample from and the number of classes. Yet, the latter
347 should not be a constrained as previous research efforts have focussed on simple forest loss maps
348 [1].

349 4.2. Research context

350 4.2.1. Long term changes in LULC and Above Ground Carbon

351 Our analysis shows that the majority of deforestation around Yangambi happened toward the late
352 1950's (~16200 ha). Considerable regrowth has occurred since the aerial survey was executed (~9918
353 ha), and socio-economic instability prevented further large scale forest exploitation. In particular, many
354 plantations have reached maturity and forest has re-established in previously cleared or disturbed
355 areas. The majority of this forest recovery takes the form of isolated patches of forest but is offset
356 by further deforestation of previously untouched forest. Generally, the function and structure of
357 forests can be influenced by forest edges that are located up to 1km away however most effects are
358 pronounced within the first 300 m from the edge [55]. Our analysis of edge effects on AGC has shown
359 that the influence is negligible 200 m away from the edge. Phillips *et al.* [56] have shown similar weak
360 responses to edge effects in the Amazon forest. Due to a lack of data on the extent (depth) of edge
361 effects and their influence on AGC beyond 200 m we did not include any estimates of carbon loss
362 or gain within these zones. However, it must be stated that the influence of edges throughout the
363 landscape was not marginal as these areas would account for 13151 ha. Thus edges could have a
364 substantial negative [6] or positive [57] influence on AGC. Similarly, uncertainties in how to explicitly
365 correct for plantations in the landscape present a further challenge. Thus, although our estimates
366 are only indicative they do underscore the important influence of landscape structure when carbon
367 accounting. However, our findings do not indicate that deforestation in Congo basin is declining, on
368 the contrary.

369 Over the past half century there has been a clear shift in land use in Yangambi (Figure 5). Land
370 use has shifted away from for example a regular fishbone deforestation pattern that emerges when
371 (large scale) agricultural interests dominate the landscape [58], to a more fragmented landscape
372 (Figure 5D). The latter former? is consistent with historical land management at the time of the aerial
373 survey [44]. These regular patterns have since been reversed because of a decrease in large scale
374 intensive agriculture, replaced by ad-hoc small scale subsistence farming with large perimeter to area
375 relationships (i.e. ragged edges). Consequently, edge effects in the current landscape are far more
376 pronounced than in the historical scene.

377 Visual inspection of the images also suggests that reforestation within the historically cleared areas
378 and experimental plots is not necessarily limited to areas far removed from more densely populated
379 areas. For example, large reforested areas exist close to the Congo stream and Yangambi village itself
380 (Fig. 5). Here, regional political components, such as land leases and large scale ownership could
381 have played a role in safeguarding some of these areas for rewinding or sustainable management

[59,60]. Despite widespread anthropogenic influences throughout the tropics [31] the retention of these forested areas show the potential of explicit or implicit protective policy measures (e.g. INERA concessions, Bustillo *et al.* [44]) on a multi-decadal time scale. Forest regrowth in non-continuous areas within Yangambi could increase landscape connectivity and help increase biodiversity [12].

Our analysis therefore provides an opportunity to highlight and study those regions that have previously suffered confirmed long-term disturbances, and those that have been restored since. Assessments of old plantations and recovering clear-cut forests can serve as a guide to help estimate carbon storage capacity and forest recovery rates in managed and unmanaged conditions [18,20,61] over the mid- to long-term, in support of rewilling and general forest restoration [12,59,60]. In addition, mapping long-term edge effects can further support research into issues such as receding forest edges [55].

4.2.2. Canopy structure & FOTO texture analysis

Finally, the FOTO technique used to quantify relationships between canopy structure and forest characteristics rendered no valuable insights of either the historical orthomosaic or recent Geo-Eye scene. Similarly weak correlations were found previously by Solórzano *et al.* [62]. In contrast, site based texture metric statistics did show correspondence between historical and contemporary satellite imagery. None of them were either consistent or significant. Although visual interpretation shows distinctly different canopy structures (Figure 5) the differences in how resolution is defined and issues related to image quality prevented us from quantifying these further. Unlike large scale studies by Ploton *et al.* [14] we could not scale this technique to historical data. **The successful use of our CNN classification model on a contemporary remote sensing data does suggest that texture can be used to consistently capture canopy properties 60 years apart. This suggests inflexibility on part of the FOTO technique in dealing with non-standardized data.** We advise that future valorisation efforts should preferentially work from foto negatives (if available) to ensure optimal data quality in resolution, contrast and sharpness.

5. Conclusion

Given the impact of tropical forest disturbances on atmospheric carbon emissions, biodiversity and ecosystem productivity accurate long term reporting of LULCC is an imperative. Our analysis of historical aerial survey images (1958) of the Central Congo Basin provides a window into the state of the forest at the start of the anthropocene. The use of a CNN based LULC classifier, using synthetic landscapes based image augmentation, provides a robust semi-supervised solution which scales across space and time, even for image scenes with inconsistent illumination or sharpness. Combined with contemporary remote sensing data we have shown that historical aerial survey data can be used to

415 quantify long-term changes in LULC and AGC. We showed a shift from previously highly structured
416 industrial deforestation of large areas **for plantation purposes**, to discrete smallholder clearing for
417 farming, increasing landscape fragmentation but also opportunities for substantial regrowth. Efforts to
418 quantify canopy texture features and their link to AGC had limited to no success. Our analysis provides
419 insights into the rate at which deforestation and reforestation has taken place over a multi-decadal
420 scale in the central Congo basin. As such, it provides a useful historical context while interpreting past
421 and ongoing forest research in the area.

422 6. Additional Information and Declarations

423 6.1. Data availability

424 Hufkens et al. (2019). A curated dataset of aerial survey images over the central Congo Basin, 1958.
425 Zenodo: <https://doi.org/10.5281/zenodo.3547767>. All data not included in the latter repository can be
426 found bundled with the analysis code as listed below. Proprietary datasets (i.e. Geo-Eye data) are not
427 shared, but purchase order numbers allow for acquisition of these datasets to ensure reproducibility.

428 6.2. Code availability

429 All analysis code is available as an R / python [63] projects (<https://khufkens.github.io/orthodrc>
430 & https://khufkens.github.io/orthodrc_cnn/). The analysis relied heavily on the 'raster' [64],
431 'RStoolbox' [65], and 'landscapemetrics' [43] packages, while post-processing and plotting was
432 facilitated by the 'tidyverse' ecosystem [66], 'ggthemes' [67], 'scales' [68] and 'cowplot' [69]. Additional
433 plotting elements were formatted or provided by 'sf' [70] and 'rnaturalearth' [71] packages, respectively.
434 I'm grateful for the contributions to the scientific community by the developers of these packages.

435 **Acknowledgments:** This research was supported through the Belgian Science Policy office COBECORE
436 project (BELSPO; grant BR/175/A3/COBECORE). KH acknowledges funding from the European Union Marie
437 Skłodowska-Curie Action (project number 797668). I'm grateful for the support of Thales de Haulleville
438 volunteering his time in scanning the images.

439 **Author Contributions:** K.H. conceived and designed the study, analyzed the data, prepared figures, tables and
440 authored the final draft of the manuscript. T.d.H. scanned all image data. E.K. and T.d.H. provided plot based
441 AGC estimates. T.D., K.J., E.K, H.B., P.S., F.V.S.M., M.A., J.V.D.B., H.V and L.W. reviewed the final manuscript.

442 **Conflicts of Interest:** The authors declare no conflict of interest. The founding sponsors had no role in the design
443 of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, an in the
444 decision to publish the results.

445 References

- 446 1. Hansen, M.C.; Potapov, P.V.; Moore, R.; Hancher, M.; Turubanova, S.A.; Tyukavina, A.; Thau, D.; Stehman,
447 S.V.; Goetz, S.J.; Loveland, T.R.; Kommareddy, A.; Egorov, A.; Chini, L.; Justice, C.O.; Townshend,

- 448 J.R.G. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* **2013**, *342*, 850–853.
449 doi:10.1126/science.1244693.
- 450 2. Houghton, R.A.; House, J.I.; Pongratz, J.; van der Werf, G.R.; DeFries, R.S.; Hansen, M.C.; Le Quéré, C.;
451 Ramankutty, N. Carbon emissions from land use and land-cover change. *Biogeosciences* **2012**, *9*, 5125–5142.
452 doi:10.5194/bg-9-5125-2012.
- 453 3. Tyukavina, A.; Baccini, A.; Hansen, M.C.; Potapov, P.V.; Stehman, S.V.; Houghton, R.A.; Krylov, A.M.;
454 Turubanova, S.; Goetz, S.J. Aboveground carbon loss in natural and managed tropical forests from 2000 to
455 2012. *Environmental Research Letters* **2015**, *10*, 074002. doi:10.1088/1748-9326/10/7/074002.
- 456 4. van der Werf, G.R.; Morton, D.C.; DeFries, R.S.; Olivier, J.G.J.; Kasibhatla, P.S.; Jackson, R.B.; Collatz, G.J.;
457 Randerson, J.T. CO₂ emissions from forest loss. *Nature Geoscience* **2009**, *2*, 737–738. doi:10.1038/ngeo671.
- 458 5. Fauset, S.; Gloor, M.U.; Aidar, M.P.M.; Freitas, H.C.; Fyllas, N.M.; Marabesi, M.A.; Rochelle, A.L.C.;
459 Shenkin, A.; Vieira, S.A.; Joly, C.A. Tropical forest light regimes in a human-modified landscape. *Ecosphere*
460 **2017**, *8*, e02002. doi:10.1002/ecs2.2002.
- 461 6. Brinck, K.; Fischer, R.; Groeneveld, J.; Lehmann, S.; Dantas De Paula, M.; Pütz, S.; Sexton, J.O.; Song, D.;
462 Huth, A. High resolution analysis of tropical forest fragmentation and its impact on the global carbon
463 cycle. *Nature Communications* **2017**, *8*. doi:10.1038/ncomms14855.
- 464 7. Didham, R.K. Edge Structure Determines the Magnitude of Changes in Microclimate and Vegetation
465 Structure in Tropical Forest Fragments. *Biotropica* **1999**, *31*, 17–30.
- 466 8. Laurance, W.F.; Delamônica, P.; Laurance, S.G.; Vasconcelos, H.L.; Lovejoy, T.E. Rainforest fragmentation
467 kills big trees. *Nature* **2000**, *404*, 836–836. doi:10.1038/35009032.
- 468 9. Magnago, L.F.S.; Magrach, A.; Laurance, W.F.; Martins, S.V.; Meira-Neto, J.A.A.; Simonelli, M.; Edwards,
469 D.P. Would protecting tropical forest fragments provide carbon and biodiversity cobenefits under REDD+?
470 *Global Change Biology* **2015**, *21*, 3455–3468. doi:10.1111/gcb.12937.
- 471 10. Poorter, L.; Bongers, F. Leaf traits are good predictors of plant performance across 53 rain forest species.
472 *Ecology* **2006**, *87*, 1733–1743. doi:10.1890/0012-9658(2006)87[1733:LTAGPO]2.0.CO;2.
- 473 11. Barlow, J.; Lennox, G.D.; Ferreira, J.; Berenguer, E.; Lees, A.C.; Nally, R.M.; Thomson, J.R.; Ferraz, S.F.d.B.;
474 Louzada, J.; Oliveira, V.H.F.; Parry, L.; Ribeiro de Castro Solar, R.; Vieira, I.C.G.; Aragão, L.E.O.C.; Begotti,
475 R.A.; Braga, R.F.; Cardoso, T.M.; Jr, R.C.d.O.; Souza Jr, C.M.; Moura, N.G.; Nunes, S.S.; Siqueira, J.V.;
476 Pardini, R.; Silveira, J.M.; Vaz-de Mello, F.Z.; Veiga, R.C.S.; Venturieri, A.; Gardner, T.A. Anthropogenic
477 disturbance in tropical forests can double biodiversity loss from deforestation. *Nature* **2016**, *535*, 144–147.
478 doi:10.1038/nature18326.
- 479 12. Van de Perre, F.; Willig, M.R.; Presley, S.J.; Bapeamoni Andemwana, F.; Beeckman, H.; Boeckx, P.; Cooleman,
480 S.; de Haan, M.; De Kesel, A.; Dessein, S.; Grootaert, P.; Huygens, D.; Janssens, S.B.; Kearsley, E.; Kabeya,
481 P.M.; Leponce, M.; Van den Broeck, D.; Verbeeck, H.; Würsten, B.; Leirs, H.; Verheyen, E. Reconciling
482 biodiversity and carbon stock conservation in an Afrotropical forest landscape. *Science Advances* **2018**,
483 *4*, eaar6603. doi:10.1126/sciadv.aar6603.
- 484 13. Mitchard, E.T.A. The tropical forest carbon cycle and climate change. *Nature* **2018**, *559*, 527–534.
485 doi:10.1038/s41586-018-0300-2.
- 486 14. Ploton, P.; Pélassier, R.; Proisy, C.; Flavenot, T.; Barbier, N.; Rai, S.N.; Couturon, P. Assessing aboveground
487 tropical forest biomass using Google Earth canopy images. *Ecological Applications* **2012**, *22*, 993–1003.
- 488 15. Couturon, P.; Pelissier, R.; Nicolini, E.a.; Paget, D. Predicting tropical forest stand structure parameters
489 from Fourier transform of very high-resolution remotely sensed canopy images. *Journal of Applied Ecology*
490 **2005**, *42*, 1121–1128. doi:10.1111/j.1365-2664.2005.01097.x.
- 491 16. Barbier, N.; Couturon, P.; Proisy, C.; Malhi, Y.; Gastellu-Etchegorry, J.P. The variation of apparent crown
492 size and canopy heterogeneity across lowland Amazonian forests. *Global Ecology and Biogeography* **2010**,
493 *19*, 72–84. doi:10.1111/j.1466-8238.2009.00493.x.
- 494 17. DeFries, R.S.; Houghton, R.A.; Hansen, M.C.; Field, C.B.; Skole, D.; Townshend, J. Carbon emissions from
495 tropical deforestation and regrowth based on satellite observations for the 1980s and 1990s. *Proceedings of
496 the National Academy of Sciences* **2002**, *99*, 14256–14261. doi:10.1073/pnas.182560099.
- 497 18. Achard, F.; Beuchle, R.; Mayaux, P.; Stibig, H.J.; Bodart, C.; Brink, A.; Carboni, S.; Desclée, B.; Donnay, F.;
498 Eva, H.D.; Lupi, A.; Raši, R.; Seliger, R.; Simonetti, D. Determination of tropical deforestation rates and
499 related carbon losses from 1990 to 2010. *Global Change Biology* **2014**, *20*, 2540–2554. doi:10.1111/gcb.12605.

- 500 19. Ramankutty, N.; Foley, J.A. Estimating historical changes in global land cover: Croplands from 1700 to
501 1992. *Global Biogeochemical Cycles* **1999**, *13*, 997–1027. doi:10.1029/1999GB900046.
- 502 20. Sader, S.A.; Joyce, A.T. Deforestation Rates and Trends in Costa Rica, 1940 to 1983. *Biotropica* **1988**, *20*, 11.
503 doi:10.2307/2388421.
- 504 21. Willcock, S.; Phillips, O.L.; Platts, P.J.; Swetnam, R.D.; Balmford, A.; Burgess, N.D.; Ahrends, A.; Bayliss,
505 J.; Doggart, N.; Doody, K.; Fanning, E.; Green, J.M.H.; Hall, J.; Howell, K.L.; Lovett, J.C.; Marchant, R.;
506 Marshall, A.R.; Mbilinyi, B.; Munishi, P.K.T.; Owen, N.; Topp-Jorgensen, E.J.; Lewis, S.L. Land cover
507 change and carbon emissions over 100 years in an African biodiversity hotspot. *Global Change Biology* **2016**,
508 *22*, 2787–2800. doi:10.1111/gcb.13218.
- 509 22. Song, D.X.; Huang, C.; Sexton, J.O.; Channan, S.; Feng, M.; Townshend, J.R. Use of Landsat and
510 Corona data for mapping forest cover change from the mid-1960s to 2000s: Case studies from the Eastern
511 United States and Central Brazil. *ISPRS Journal of Photogrammetry and Remote Sensing* **2015**, *103*, 81–92.
512 doi:10.1016/j.isprsjprs.2014.09.005.
- 513 23. Nita, M.D.; Munteanu, C.; Gutman, G.; Abrudan, I.V.; Radeloff, V.C. Widespread forest cutting in the
514 aftermath of World War II captured by broad-scale historical Corona spy satellite photography. *Remote
515 Sensing of Environment* **2018**, *204*, 322–332. doi:10.1016/j.rse.2017.10.021.
- 516 24. Buitenhof, R.; Bond, W.J.; Stevens, N.; Trollope, W.S.W. Increased tree densities in South African
517 savannas: >50 years of data suggests CO₂ as a driver. *Global Change Biology* **2012**, *18*, 675–684.
518 doi:10.1111/j.1365-2486.2011.02561.x.
- 519 25. Hudak, A.T.; Wessman, C.A. Textural Analysis of Historical Aerial Photography to Characterize Woody
520 Plant Encroachment in South African Savanna. *Remote Sensing of Environment* **1998**, *66*, 317–330.
521 doi:10.1016/S0034-4257(98)00078-9.
- 522 26. Frankl, A.; Seghers, V.; Stal, C.; De Maeyer, P.; Petrie, G.; Nyssen, J. Using image-based modelling
523 (SfM-MVS) to produce a 1935 ortho-mosaic of the Ethiopian highlands. *International Journal of Digital Earth*
524 **2015**, *8*, 421–430. doi:10.1080/17538947.2014.942715.
- 525 27. Nyssen, J.; Petrie, G.; Mohamed, S.; Gebremeskel, G.; Seghers, V.; Debever, M.; Hadgu, K.M.; Stal, C.; Billi,
526 P.; Demaeyer, P.; Haile, M.; Frankl, A. Recovery of the aerial photographs of Ethiopia in the 1930s. *Journal
527 of Cultural Heritage* **2016**, *17*, 170–178. doi:10.1016/j.culher.2015.07.010.
- 528 28. Lewis, S.L.; Lopez-Gonzalez, G.; Sonké, B.; Affum-Baffoe, K.; Baker, T.R.; Ojo, L.O.; Phillips, O.L.; Reitsma,
529 J.M.; White, L.; Comiskey, J.A.; Djuiouko, M.N.; Ewango, C.E.N.; Feldpausch, T.R.; Hamilton, A.C.;
530 Gloor, M.; Hart, T.; Hladik, A.; Lloyd, J.; Lovett, J.C.; Makana, J.R.; Malhi, Y.; Mbago, F.M.; Ndangalasi,
531 H.J.; Peacock, J.; Peh, K.S.H.; Sheil, D.; Sunderland, T.; Swaine, M.D.; Taplin, J.; Taylor, D.; Thomas, S.C.;
532 Votere, R.; Wöll, H. Increasing carbon storage in intact African tropical forests. *Nature* **2009**, *457*, 1003–1006.
533 Publisher: Earth and Biosphere Institute, School of Geography, University of Leeds, Leeds LS2 9JT, UK.
534 s.l.lewis@leeds.ac.uk.
- 535 29. Butsic, V.; Baumann, M.; Shortland, A.; Walker, S.; Kuemmerle, T. Conservation and conflict in the
536 Democratic Republic of Congo: The impacts of warfare, mining, and protected areas on deforestation.
537 *Biological Conservation* **2015**, *191*, 266–273. doi:10.1016/j.biocon.2015.06.037.
- 538 30. Kearsley, E.; de Haulleville, T.; Hufkens, K.; Kidimbu, A.; Toirambe, B.; Baert, G.; Huygens, D.; Kebede,
539 Y.; Defourny, P.; Bogaert, J.; Beeckman, H.; Steppe, K.; Boeckx, P.; Verbeeck, H. Conventional tree
540 height-diameter relationships significantly overestimate aboveground carbon stocks in the Central Congo
541 Basin. *Nature communications* **2013**, *4*, 2269. doi:10.1038/ncomms3269.
- 542 31. Lewis, S.L.; Maslin, M.A. Defining the Anthropocene. *Nature* **2015**, *519*, 171–180. doi:10.1038/nature14258.
- 543 32. Bauters, M.; Ampoorter, E.; Huygens, D.; Kearsley, E.; De Haulleville, T.; Sellan, G.; Verbeeck, H.; Boeckx,
544 P.; Verheyen, K. Functional identity explains carbon sequestration in a 77-year-old experimental tropical
545 plantation. *Ecosphere* **2015**, *6*, art198. doi:10.1890/ES15-00342.1.
- 546 33. Zuiderveld, K. Contrast Limited Adaptive Histogram Equalization. In *Graphics GEMs IV*; Academic Press:
547 San Diego, CA, USA, 1994; pp. 474–485.
- 548 34. Ullman, S. The Interpretation of Structure from Motion. *Proceedings of the Royal Society of London. Series B,
549 Biological Sciences* **1979**, *203*, 405–426.
- 550 35. QGIS Development team., QGIS Geographic Information System., 2018. Available at: <http://qgis.osgeo.org>.

- 552 36. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation.
553 In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*; Navab, N.; Hornegger, J.;
554 Wells, W.M.; Frangi, A.F., Eds.; Springer International Publishing: Cham, 2015; Vol. 9351, pp. 234–241.
555 doi:10.1007/978-3-319-24574-4_28.
- 556 37. Chollet, F. Keras, 2015. Available at: <https://github.com/fchollet/keras>.
- 557 38. Yakubovskiy, P. Segmentation Models, 2019. Available at: https://github.com/qubvel/segmentation_models.
- 559 39. Martín Abadi.; Ashish Agarwal.; Paul Barham.; Eugene Brevdo.; Zhifeng Chen.; Craig Citro.; Greg S.
560 Corrado.; Andy Davis.; Jeffrey Dean.; Matthieu Devin.; Sanjay Ghemawat.; Ian Goodfellow.; Andrew
561 Harp.; Geoffrey Irving.; Michael Isard.; Jia, Y.; Rafal Jozefowicz.; Lukasz Kaiser.; Manjunath Kudlur.; Josh
562 Levenberg.; Dandelion Mané.; Rajat Monga.; Sherry Moore.; Derek Murray.; Chris Olah.; Mike Schuster.;
563 Jonathon Shlens.; Benoit Steiner.; Ilya Sutskever.; Kunal Talwar.; Paul Tucker.; Vincent Vanhoucke.; Vijay
564 Vasudevan.; Fernanda Viégas.; Oriol Vinyals.; Pete Warden.; Martin Wattenberg.; Martin Wicke.; Yuan Yu.;
565 Xiaoqiang Zheng. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, 2015. Available
566 at: <https://www.tensorflow.org/>.
- 567 40. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv:1412.6980 [cs]* **2017**. arXiv:
568 1412.6980.
- 569 41. Dice, L.R. Measures of the Amount of Ecologic Association Between Species. *Ecology* **1945**, *26*, 297–302.
570 doi:10.2307/1932409.
- 571 42. Dale, M.R.T. *Spatial pattern analysis in plant ecology*; Cambridge studies in ecology, Cambridge University
572 Press: Cambridge ; New York, 1999.
- 573 43. Hesselbarth, M.H.K.; Sciaini, M.; With, K.A.; Wiegand, K.; Nowosad, J. landscapemetrics: an open-source
574 R tool to calculate landscape metrics. *Ecography* **2019**.
- 575 44. Bustillo, E.; Raets, L.; Beeckman, H.; Bourland, N.; Rousseau, M.; Hubau, W.; De Mil, T. Evaluation du
576 potentiel énergétique de la biomasse aérienne ligneuse des anciennes plantations de l'INERA Yangambi.
577 Technical report, CIFOR, 2018.
- 578 45. Couteron, P. Quantifying change in patterned semi-arid vegetation by Fourier analysis of digitized aerial
579 photographs. *International Journal of Remote Sensing* **2002**, *23*, 3407–3425. doi:10.1080/01431160110107699.
- 580 46. Wilcoxon, F. Individual comparisons by ranking methods. *Biometrics* **1945**, *1*, 80–83.
- 581 47. Kadmon, R.; Harari-Kremer, R. Studying Long-Term Vegetation Dynamics Using Digital
582 Processing of Historical Aerial Photographs. *Remote Sensing of Environment* **1999**, *68*, 164–176.
583 doi:10.1016/S0034-4257(98)00109-6.
- 584 48. Peres, C.A.; Barlow, J.; Laurance, W.F. Detecting anthropogenic disturbance in tropical forests. *Trends in
585 Ecology & Evolution* **2006**, *21*, 227–229. doi:10.1016/j.tree.2006.03.007.
- 586 49. Verhoeven, G. BRDF and its Impact on Aerial Archaeological Photography: BRDF and its impact on aerial
587 archaeological photography. *Archaeological Prospection* **2017**, *24*, 133–140. doi:10.1002/arp.1559.
- 588 50. Park, J.Y.; Muller-Landau, H.C.; Lichstein, J.W.; Rifai, S.W.; Dandois, J.P.; Bohlman, S.A. Quantifying Leaf
589 Phenology of Individual Trees and Species in a Tropical Forest Using Unmanned Aerial Vehicle (UAV)
590 Images. *Remote Sensing* **2019**, *11*, 1534. doi:10.3390/rs11131534.
- 591 51. Simini, F.; Anfodillo, T.; Carrer, M.; Banavar, J.R.; Maritan, A. Self-similarity and scaling
592 in forest communities. *Proceedings of the National Academy of Sciences* **2010**, *107*, 7658–7662.
593 doi:10.1073/pnas.1000137107.
- 594 52. Sole, R.V.; Manrubia, S.C. Self-similarity in rain forests: evidence for a critical state. *Physical Review E* **1995**,
595 *51*, 6250 – 6253.
- 596 53. Hughes, M.L.; McDowell, P.F.; Marcus, W.A. Accuracy assessment of georectified aerial photographs:
597 Implications for measuring lateral channel movement in a GIS. *Geomorphology* **2006**, *74*, 1–16.
598 doi:10.1016/j.geomorph.2005.07.001.
- 599 54. Buscombe, D.; Ritchie, A. Landscape Classification with Deep Neural Networks. *Geosciences* **2018**, *8*, 244.
600 doi:10.3390/geosciences8070244.
- 601 55. Gascon, C.; Williamson, G.B.; Fonseca, G.A.B.d. Receding Forest Edges and Vanishing Reserves. *Science, New Series* **2000**, *288*, 1356–1358.
- 602 56. Phillips, O.L.; Rose, S.; Mendoza, A.M.; Vargas, P.N. Resilience of Southwestern Amazon Forests to
603 Anthropogenic Edge Effects. *Conservation Biology* **2006**, *20*, 1698–1710. doi:10.1111/j.1523-1739.2006.00523.x.

- 605 57. Reinmann, A.B.; Hutyra, L.R. Edge effects enhance carbon uptake and its vulnerability to climate
606 change in temperate broadleaf forests. *Proceedings of the National Academy of Sciences* **2017**, *114*, 107–112.
607 doi:10.1073/pnas.1612369114.
- 608 58. Arima, E.Y.; Walker, R.T.; Perz, S.; Souza, C. Explaining the fragmentation in the Brazilian Amazonian
609 forest. *Journal of Land Use Science* **2015**, pp. 1–21. doi:10.1080/1747423X.2015.1027797.
- 610 59. Arima, E.Y.; Barreto, P.; Araújo, E.; Soares-Filho, B. Public policies can reduce tropical deforestation: Lessons
611 and challenges from Brazil. *Land Use Policy* **2014**, *41*, 465–473. doi:10.1016/j.landusepol.2014.06.026.
- 612 60. Larson, A.M. Forest tenure reform in the age of climate change: Lessons for REDD+. *Global Environmental
613 Change* **2011**, *21*, 540–549. doi:10.1016/j.gloenvcha.2010.11.008.
- 614 61. Gourlet-Fleury, S.; Mortier, F.; Fayolle, A.; Baya, F.; Ouédraogo, D.; Bénédet, F.; Picard, N. Tropical forest
615 recovery from logging: a 24 year silvicultural experiment from Central Africa. *Philosophical Transactions of
616 the Royal Society B: Biological Sciences* **2013**, *368*, 20120302. doi:10.1098/rstb.2012.0302.
- 617 62. Solórzano, J.V.; Gallardo-cruz, J.A.; González, E.J.; Peralta-carreta, C.; Hernández-gómez, M.; Oca,
618 A.F.m.D.; Cervantes-jiménez, L.G.; Solórzano, J.V.; Gallardo-cruz, J.A.; González, E.J.; Peralta-carreta,
619 C.; Hernández-gómez, M.; Oca, A.F.m.D.; Cervantes-jiménez, L.G. Contrasting the potential of
620 Fourier transformed ordination and gray level co-occurrence matrix textures to model a tropical swamp
621 forest 's structural and diversity attributes. *Journal of Applied Remote Sensing* **2018**, *12*, 036006.
622 doi:10.1117/1.JRS.12.036006.
- 623 63. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical
624 Computing: Vienna, Austria, 2019.
- 625 64. Hijmans, R.J. raster: Geographic Data Analysis and Modeling, 2019. Available at: <https://CRAN.R-project.org/package=raster>.
- 626 65. Leutner, B.; Horning, N.; Schwalb-Willmann, J. RStoolbox: Tools for Remote Sensing Data Analysis, 2019.
627 Available at: <https://CRAN.R-project.org/package=RStoolbox>.
- 628 66. Wickham, H. tidyverse: Easily Install and Load the 'Tidyverse', 2017. Available at: <https://CRAN.R-project.org/package=tidyverse>.
- 629 67. Arnold, J.B. ggthemes: Extra Themes, Scales and Geoms for 'ggplot2', 2019. Available at: <https://CRAN.R-project.org/package=ggthemes>.
- 630 68. Wickham, H. scales: Scale Functions for Visualization, 2018. Available at: <https://CRAN.R-project.org/package=scales>.
- 631 69. Wilke, C.O. cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2', 2019. Available at:
632 <https://CRAN.R-project.org/package=cowplot>.
- 633 70. Pebesma, E. Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal* **2018**,
634 *10*, 439–446. doi:10.32614/RJ-2018-009.
- 635 71. South, A. rnaturalearth: World Map Data from Natural Earth, 2017. Available at: <https://CRAN.R-project.org/package=rnaturalearth>.

641 © 2020 by the authors. Submitted to *Remote Sens.* for possible open access publication
642 under the terms and conditions of the Creative Commons Attribution (CC BY) license
643 (<http://creativecommons.org/licenses/by/4.0/>).