

Article

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

Koen Hufkens<sup>1,2,\*</sup>, Thalès de Haulleville<sup>3</sup>, Elizabeth Kearsley<sup>1</sup>, Kim Jacobsen<sup>1,3</sup>, Hans Beeckman<sup>3</sup>, Piet Stoffelen<sup>4</sup>, Filip Vandelook<sup>4</sup>, Sofie Meeus<sup>4</sup>, Michael Amara<sup>5</sup>, Leen Van Hirtum<sup>5</sup>, Jan Van den Bulcke<sup>1</sup>, Hans Verbeeck<sup>1</sup>, Lisa Wingate<sup>2</sup>

<sup>1</sup> Ghent University, Ghent, Belgium;

<sup>2</sup> INRA, UMR ISPA, Villeneuve d'Ornon, France;

<sup>3</sup> Royal Museum for Central Africa, Tervuren, Belgium;

<sup>4</sup> Botanic Garden Meise, Meise, Belgium;

<sup>5</sup> National Archives of Belgium, Brussels, Belgium;

\* Correspondence: koen.hufkens@gmail.com

Version January 30, 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 **decrease** the number of large trees along forest edges [8,9],  
22 while species composition and biodiversity are equally negatively affected [10–12]. Estimates show  
23 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 Germany [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 1930s.

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 (1958) and contemporary remote  
48 sensing data (2000–2018) to map long-term changes in the extent and structure of the tropical forest  
49 surrounding Yangambi (DR Congo), in the central Congo Basin, effectively linking the start of the  
50 **anthropocene** [31] with current assessments. Yangambi was, and remains, a focal center of forest and

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

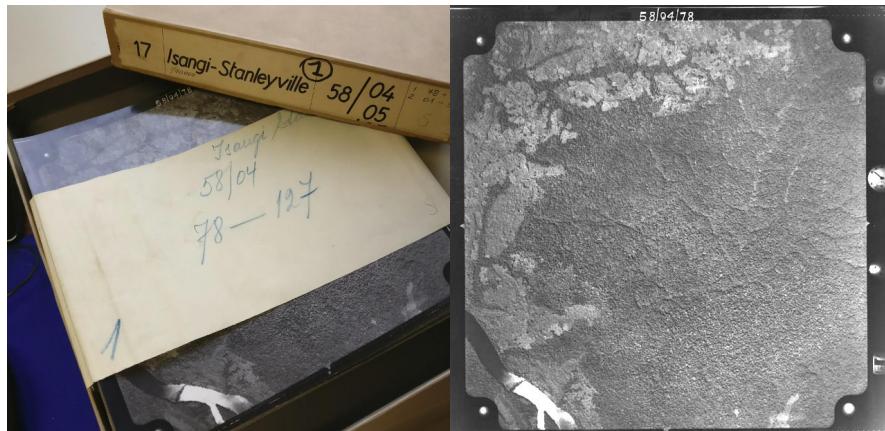
## 63 2. Methods

### 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 **panchromatic** images (0.4 - 0.9  $\mu\text{m}$ ) were gathered along flight paths running mostly  
70 from west to east, between 09:00 and 11:00 local time. Along a flight path continuous images were  
71 taken using a Wild Heerburg RC5a (currently Leica Geosystems) with an Aviogon lens assembly  
72 (114.83mm / f 5.6, with a 90°view angle) resulting in square photo negatives of 180 by 180 mm. Flights  
73 were flown at an average absolute altitude of ~5200 m above sea level, covering roughly 18 530 km<sup>2</sup>  
74 at an approximate scale of 1/40 000. The use of the integrated autograph system ensured timely  
75 acquisition of pictures with a precise overlap (~1/3) between images. This large overlap between  
76 images together with flight parameters would allow post-processing, using stereographs, to create  
77 accurate topographic maps. Original data from this campaign are stored in the [Royal Museum for](#)  
78 [Central Africa](#) in Tervuren, 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



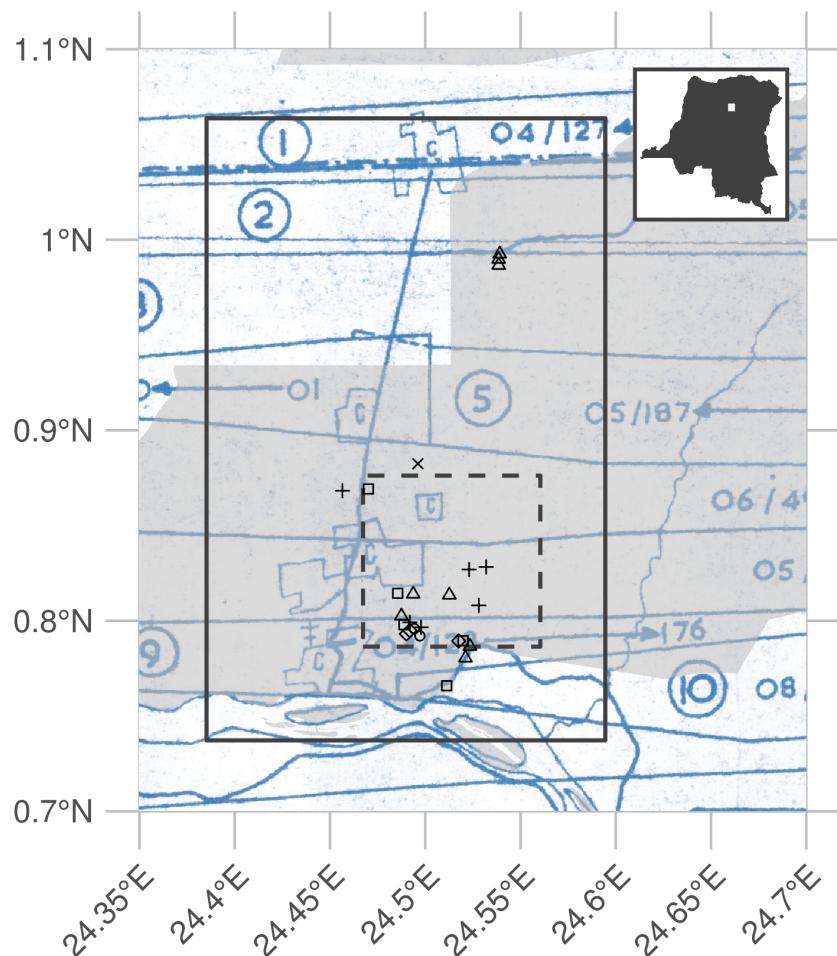
**Figure 1.** A box of historical aerial photographs (left) and a single aerial photograph (right) showing part of the Congo river. Note the meta-data provided in the right side margin of the image, such as acquisition time and flight height.

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.

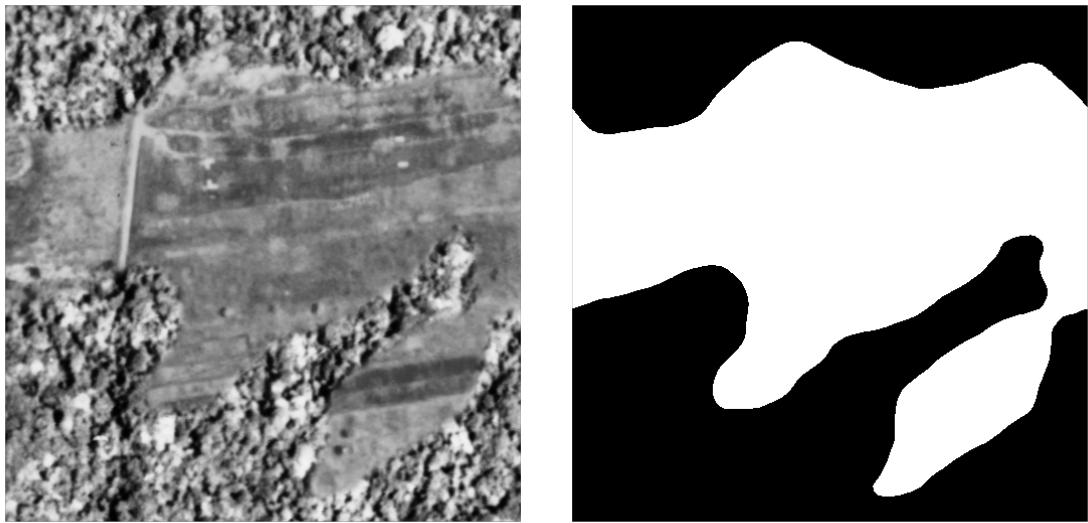
108 We calculated the orthomosaic using a low resolution point cloud and digital elevation map  
109 (DEM). Additional ground control points were provided to assist in the referencing of image and  
110 constrain the optimization routine used in the SfM algorithm. Ground control points consisted of  
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 pool  
113 etc.). **Although most clouds were removed during the SfM routine some were retained to provide**  
114 **sufficient SfM tie points to maximize continuous forest coverage in the final orthomosaic.** The  
115 final scene was cropped, to provide consistent wall to wall coverage of the reconstructed scene. The  
116 orthomosaic was exported as a geotiff for further georeferencing in QGIS [35] using the georeferencer  
117 plugin (version 3.1.9) and additional ESRI World Imagery high resolution reference data. We used 3<sup>rd</sup>  
118 degree polynomial and 16 ground control points to correct the final image. Ground control points, raw  
119 image data and final processed image are provided in addition to measures of uncertainty such as  
120 mean and median error across all ground control points. All subsequent analysis are executed on the  
121 final geo-referenced orthomosaic or subsets 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 and  
133 validation polygons were sampled at 100 random locations. Tiles extracted from locations close to the  
134 polygon border at times contained mixed cover types. Tiles with mixed cover types were removed  
135 from the list of source tiles (Table 1). Homogeneous source tiles were combined in synthetic landscapes  
136 using a random gaussian field based binary mask (Figure 3). We generated 5000 synthetic landscapes



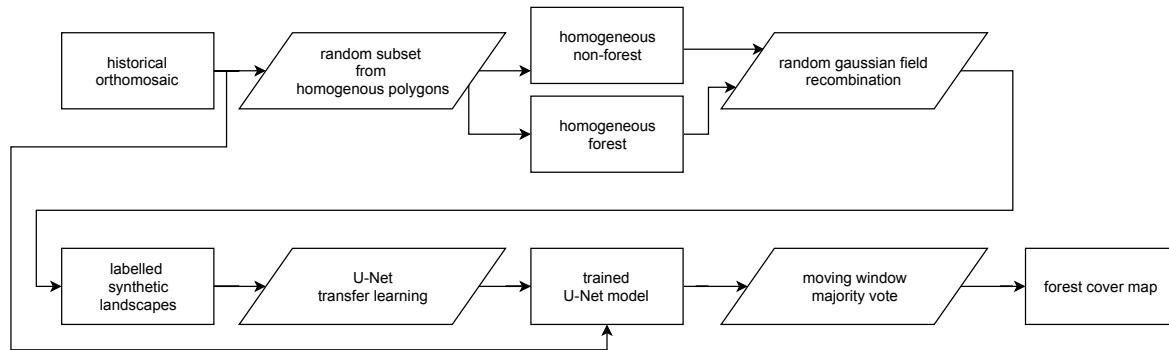
**Figure 3.** An example synthetic landscape, combining homogeneous forest and non-forest images into a patchy landscape using random gaussian field based masks. The left panel shows a combined synthetic landscape, while the right panel shows the corresponding forest (black) and non-forest (white) labels.

137 (balancing forest and non-forest classes) for training, while 500 landscapes were generated for both  
 138 the validation and the testing datasets for a total of 6000 synthetic landscapes. In order to limit stitch  
 139 line misclassifications, along the seams of mosaicked images, we created synthetic landscapes with  
 140 different forest tiles to mimick forest texture transitions. We applied this technique to 10% of the  
 141 generated synthetic 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 of  
 153     synthetic landscapes. In addition, we report confusion matrices for all pixels across the homogeneous  
 154     validation polygons, as well as the training and testing polygons (see Figure 5). Furthermore, we used  
 155     the first acquisition of a recent pan-chromatic Geo-Eye 1 stereo pair (Geo-Eye, Thornton, Colorado,  
 156     U.S.A., order 737537, 2011-11-11 08:55 GMT or 09:55 local time) to classify and assess the robustness  
 157     of the CNN algorithm on contemporary remote sensing imagery with similar spectral and spatial  
 158     characteristics. We used the Global Forest Change version 1.6 (GFC, tile 10N-020E) [1] map data,  
 159     capturing **deforestation** up to 2011, to quantify accuracy on downsampled CNN Geo-Eye classification  
 160     results. Once more, we report the confusion matrix of between the GFC and CNN derived forest cover  
 161     maps, masking clouds and cloud shadows. To summarize confusion matrices we report accuracy as:

$$162 \quad Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)}$$

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

165     2.4.2. Characterizing long term change

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

175 statistics of historical and contemporary **deforestation and reforestation**. We map permanent forest  
176 loss after 1958, **reforestation** after 1958, **recent deforestation** and long term (stable) forest cover. All  
177 references to changes over time in the context of our analysis explicitly compare the historical and  
178 contemporary periods from hereon forth.

179 2.4.3. Landscape fragmentation & Above Ground Carbon estimates

180 To quantify changes in the structure of forest cover and its disturbances we used spatial pattern  
181 analysis landscape (fragmentation) metrics [42]. Landscape metrics provide a mathematical framework  
182 for the analysis of discrete land-cover classes and allows us to capture their composition and  
183 configuration. These metrics are therefore commonly used to compare how landscapes change over  
184 time [43]. In particular, fractals provide a way to quantify complex natural landscapes, including their  
185 self-similarity, across scales [44,45] We report the ratio of edge to area and the fractal dimension to  
186 quantify landscape complexity of forest disturbances. A fractal dimension closer to 2 suggest a more  
187 complex (fragmented) landscape.

188 Statistics were calculated for all forest disturbance patches larger than 1 ha and smaller than the  
189 95<sup>th</sup> percentile of the patch size distribution using the R package *landscapemetrics* [43]. We provide  
190 mean and standard deviation on edge, area, their ratio and fractal dimension for both the historical  
191 and contemporary Hansen *et al.* [1] forest cover maps.

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

205 Surveys of old plantations show a large variation in AGC, depending on age and the crop type.  
206 For example, the AGC values varied from 168.67 to 86.55 Mg C ha<sup>-1</sup>, for *Hevea brasiliensis* (rubber  
207 tree) and *Elaeis guineensis* (oil palm) plots respectively (Bustillo *et al.* [46], personal communications).

208 These higher values are in line with the mixed AGC estimates in the area, while the palm plantations  
209 resemble old-regrowth values ( $81.8 \text{ Mg C ha}^{-1}$ , see Appendix Table 1). We therefore use both the  
210 estimates of old-regrowth and mixed forest to estimate AGC for regrowth. We did not have sufficient  
211 data to account for individual changes in AGC across plantations.

212 *2.5. Canopy structure & FOTO texture analysis*

213 We compared the structure of the canopy both visually and using Fourier Transform Textural  
214 Ordination (FOTO, Couteron [47]). FOTO uses a principal component analysis (PCA) on radially  
215 averaged 2D Fourier spectra to characterize canopy (image) texture. The FOTO technique was first  
216 described by Couteron [47] to quantify canopy stucture in relation to biomass and biodiversity, and  
217 can be used across multiple scenes using normalization [16].

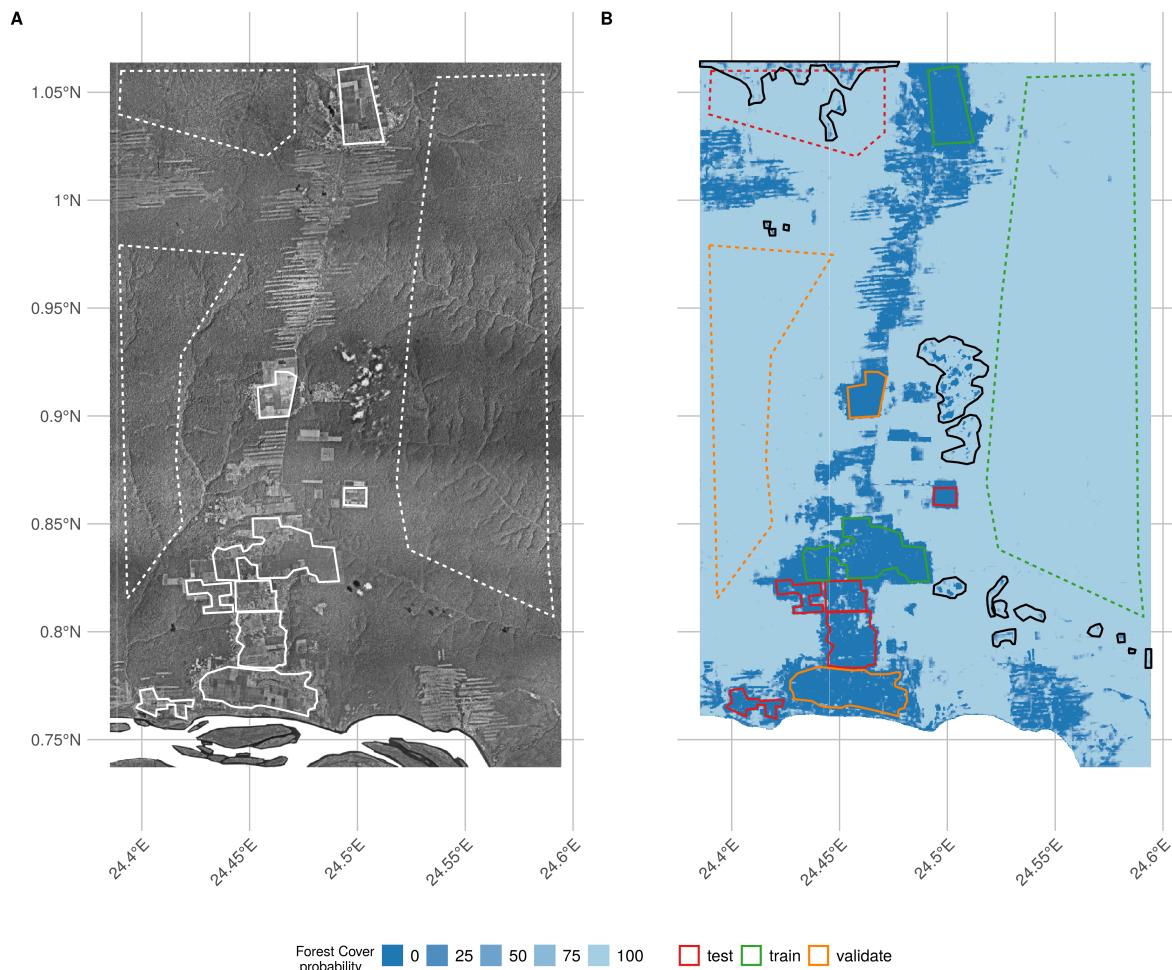
218 We used the standard FOTO methodology with fixed zones, instead of the moving window  
219 approach. The window size was set to the same size (187 pixels or  $\sim 150 \text{ m}$ ) as used in the moving  
220 window analysis above. To account for illumination differences between the two scenes we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

224 For both site based and scene analysis we correlated PC values with permanent sample plot  
225 inventory data such as stem density, above ground biomass and tree species richness. Additional  
226 comparisons are made between contemporary Geo-Eye data and the historical orthomosaic derived PC  
227 values. Due to the few available permanent sampling plots in both scenes we used a non-parametric  
228 paired signed rank (Wilcoxon) test [48] to determine differences between the PC values of the Geo-Eye  
229 and historical orthomosaic image scenes across mono-dominant and mixed forest types. In all analysis,  
230 mono-dominant site 4 was removed from the analysis due to cloud contamination.

231 **3. Results**

232 *3.1. Orthomosaic construction*

233 Our analysis provides a first spatially explicity historical composite of aerial survey images in  
234 support of mapping land-use and land-cover within the Congo Basin. The use of high resolution  
235 historical images combined with SfM image processing techniques allowed us to mosaic old imagery  
236 across a large extent. The final orthomosaic composition of the Yangambi region provided an image  
237 scene covering approximately 733 million pixels across  $\sim 82800 \text{ ha}$  ( $\sim 23 \times 36 \text{ km}$ , Figure 2). The overall  
238 accuracy of the SfM orthomosaic composition was 0.88 m/pixel using the sparse cloud DEM for  
239 corrections at 45.8 m/pixel. The resulting georeferenced scene had a spatial accuracy of approximately



**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).

240 23 m. Further georeferencing outside the SfM workflow reduced the mean error at the ground control  
 241 points to  $5.3 \pm 4.9$  px ( $\sim 4.7 \pm 4.3$  m), with a median error of 2.9 px (2.6 m). The orthomosaic served  
 242 as input for all subsequent LULCC analysis with all derived maps provided with the manuscript  
 243 repository (see data & code availability statements below).

244 *3.2. Land-use and Land-Cover Classification*

245 *3.2.1. CNN model validation*

246 The CNN deep learning classifier reached an Intersection-over-Union accuracy of 97% on the  
 247 detection of disturbed forest in the out-of-sample (validation) synthetic landscape data. Using all pixels

248 within the **validation** polygons (Figure 5) showed a similar accuracy value of ~98%. Using all polygons  
 249 across the scene, including those used in the generation of testing and training synthetic landscapes,  
 250 increased the accuracy to ~99% (Table 2). A comparison with recent pan-chromatic Geo-Eye data  
 251 shows good agreement, with an accuracy of ~87% across all pixels, between the landsat based GFC  
 252 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. In addition, overall accuracy is reported for each confusion matrix.

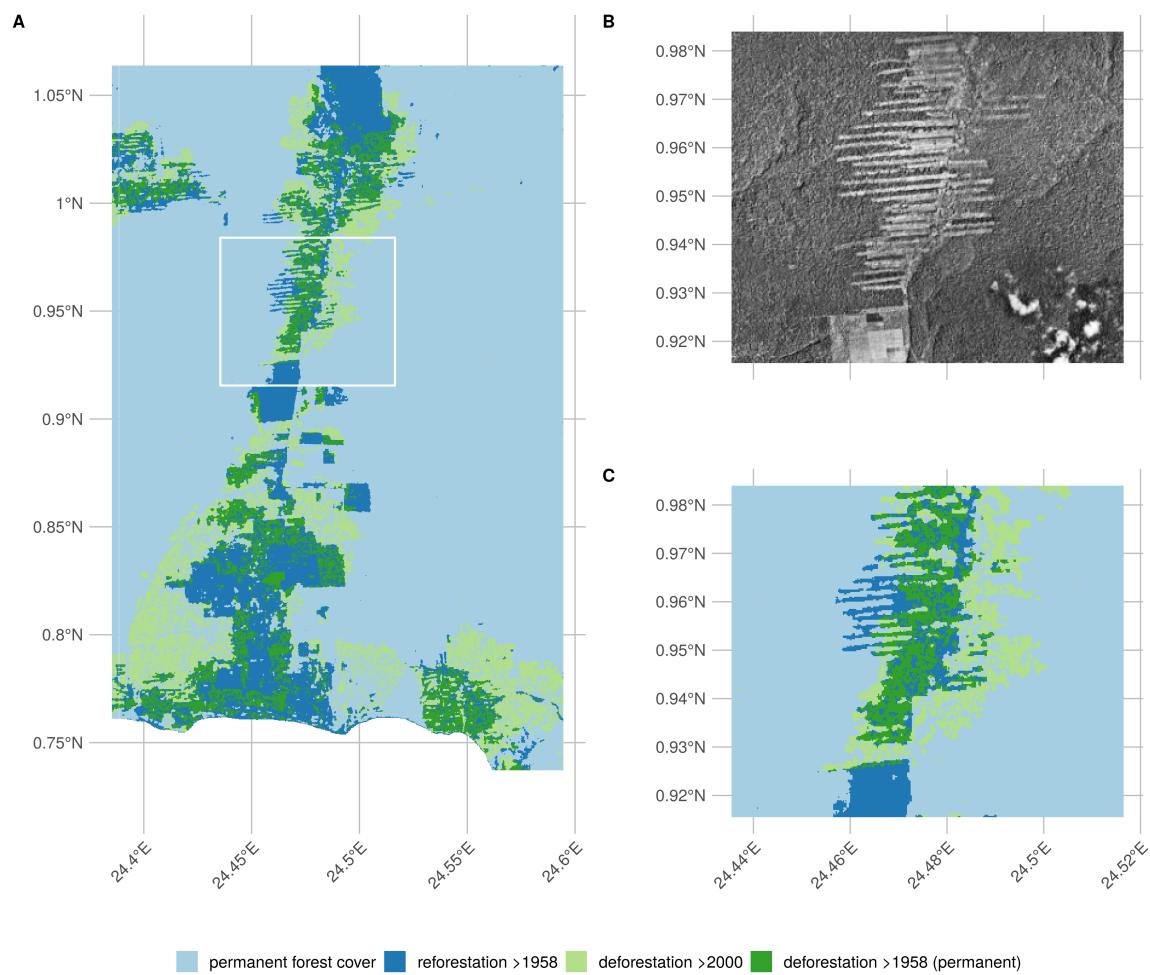
CNN	Geo-Eye (full scene)		Aerial Survey (all polygons)		Aerial Survey (validation polygons)	
	non-forest	forest	non-forest	forest	non-forest	forest
non-forest	10.17	8.55	97.4	0.19	98.59	0.0
forest	3.75	77.52	2.6	99.81	1.41	100.0
<b>Accuracy</b>		87.70		98.61		99.3

### 253 3.3. Long term changes in LULC and Above Ground Carbon

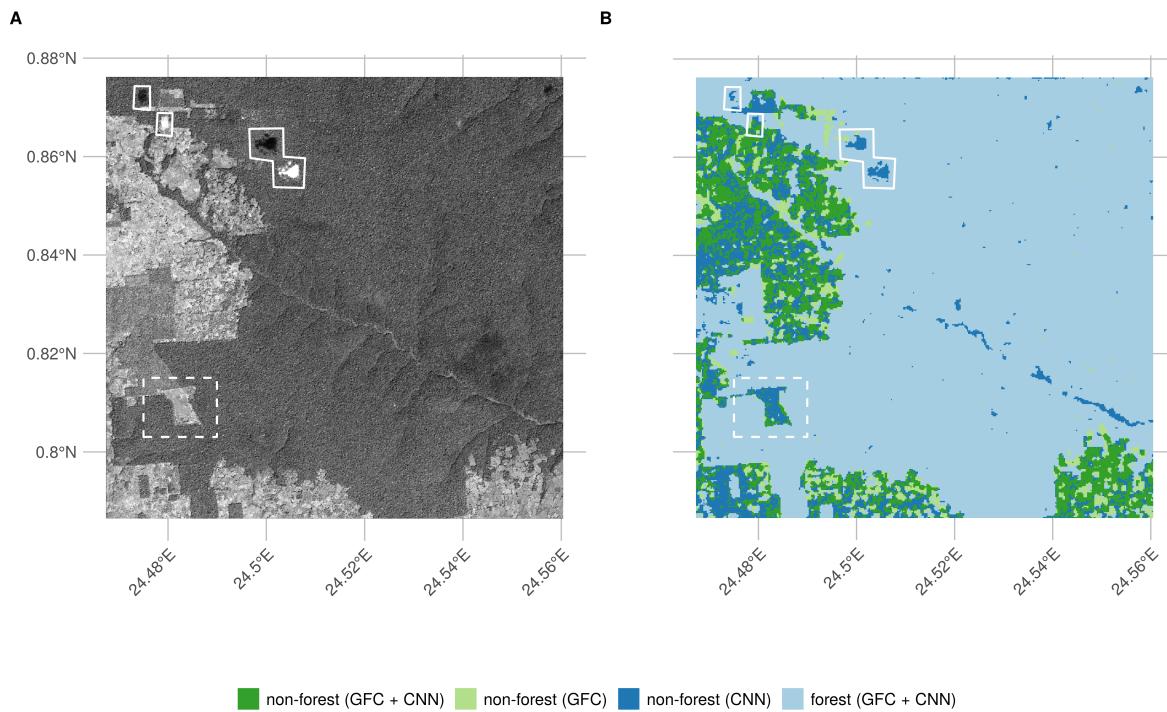
254 Scaling our classifier to the whole historical orthomosaic we detected ~16200 ha (or ~20% of the  
 255 scene) of disturbed forests. A large fraction of the disturbed area was restored in the period between  
 256 the two acquisition periods. In total, 9918 ha, or little over half of the affected forest was restored  
 257 (Figure 5C-D, dark blue). Recent deforested areas, as registered through satellite remote sensing (>  
 258 2000), approximate 8776 ha (Table 3, Figure 6 - light green).

259 Recent deforestation follows a distinctly different pattern compared to historical patterns.  
 260 Historical deforestation showed a classical fishbone pattern for forest clearing with very sharp edges,  
 261 while current patterns are patchy and ad-hoc (figure 6C, dark blue and green colours respectively).  
 262 These differences are reflected in the analysis of landscape metrics of **deforestation**. Between the  
 263 historical and contemporary LULCC maps we see an increase in small disturbances, as indicated by  
 264 the 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.  
 265 Perimeter lengths were longer historically, at  $1451 \pm 943$  m, compared to contemporary landscapes  
 266  $\sim 921 \pm 362$  m (Table 4). This shift in perimeter area ratio led to a similar change in the fractal index,  
 267 slightly increasing in value to  $1.1 \pm 0.05$  from  $1.09 \pm 0.04$  over time. Values closer to a fractal index of  
 268 2 suggest a more complex (fragmented) landscape.

269 A comparison of forest edge plots with mixed forest plots showed no significant difference in  
 270 AGC, or other reported values such as species richness, basal area or stem density (Mann Whitney U  
 271 test,  $p < 0.05$ ). Edge influence did not extend beyond 200 m from a forest edge, but still represented an  
 272 area of 13151 ha (Table 3).



**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).

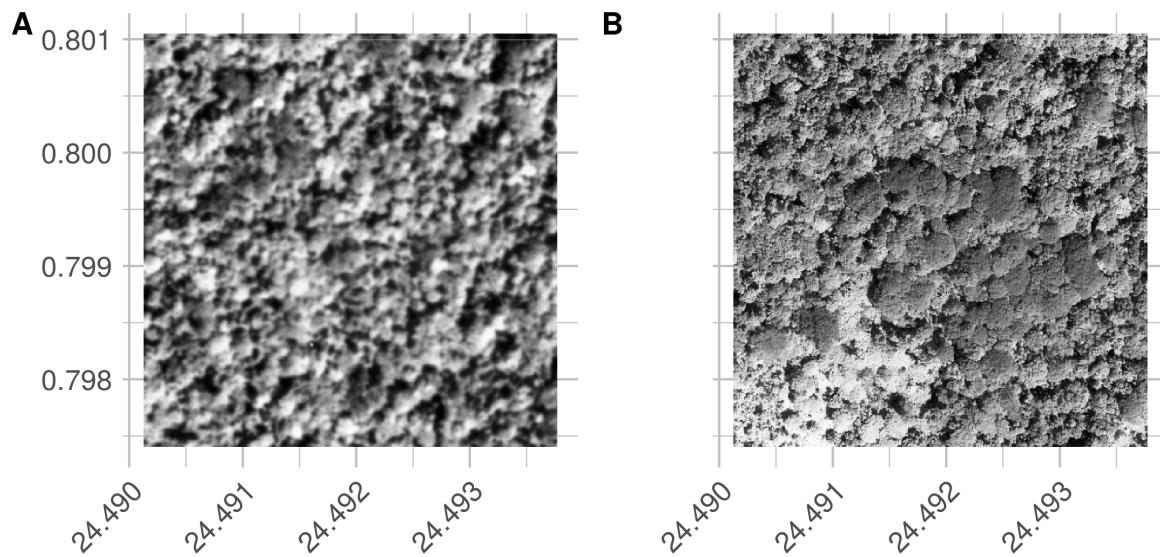


**Figure 7.** Convolutional Neural Network (CNN) based forest cover classification results (B) as run on a recent (2011) Geo-Eye panchromatic image (A). We show the difference between the Convolutional Neural Network based classification and a recent Landsat based forest cover map by Hansen et al. 2013.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

Changes in both land-use and land-cover led to concomitant changes in AGC stocks. Recovery throughout the region was characterized for patches of forest and plantations. Assuming high density stands, based on previous work, this could amount to a carbon gains of 1592 Gg C across our study area, offsetting more recent losses of approximately 1408 Gg C. On the other hand, at the low end, if 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 C. Using our approach results indicate that overall deforestation around Yangambi has resulted in a 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	$\text{Gg C ha}^{-1}$
Non-forest cover (1958)	16200	
Forest cover	68455	
- of which forest edges	13151	
reforestation >1958	9918	811 - 1592
deforestation >2000	8776	1408
deforestation >1958 (permanent)	6282	1008



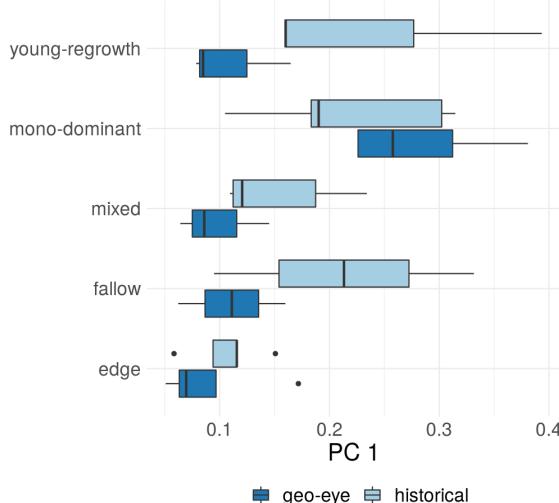
**Figure 8.** Visual comparison between a historical (A) and contemporary (B) permanent sampling plot. The site is currently listed as a mono-dominant *Brachystegia laurentii* stand. Note the structural differences with a 'coarser' canopy structure in the historical image compared to the more closed 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	$1451.58 \pm 943.27$	$5.25 \pm 5.02$	$0.03 \pm 0.01$	$1.09 \pm 0.04$
contemporary	$921.74 \pm 362.59$	$1.86 \pm 0.75$	$0.05 \pm 0.01$	$1.1 \pm 0.05$

### 280 3.4. Canopy structure & FOTO texture analysis

281 Visual interpretation of the scenes provide evidence that most locations do not change dramatically  
 282 with respect to canopy composition, except for the large areas of disturbances in contemporary fallow  
 283 or young-regrowth plots. One marked difference is noted in the mono-dominant plot 6 (Appendix  
 284 Table 1). Here, the current mono-dominant *Brachystegia laurentii* is a recent development, changing the  
 285 canopy structure visibly during the last half century (Figure 8). The previous varied canopy structure  
 286 gave way to a more dense and uniform canopy. This is reflected in a change of the FOTO PC value  
 287 from 0.19 historically to its current value of 0.54 (Figure 9). This historical value is similar to the  
 288 mean of contemporary mono-dominant stands of *Gilbertiodendron dewevrei* with PC averaging  $0.34 \pm$   
 289 0.1, and is only slightly higher than historical values for a mixed forest ( $0.18 \pm 0.08$ , Figure 9). The  
 290 reverse pattern is seen in the contemporary PC values. Here, the value of 0.54 exceeds those of most  
 291 mono-dominant stands ( $0.35 \pm 0.08$ ), and is even further removed from the values noted for mixed  
 292 forests ( $0.12 \pm 0.03$ , Figure 9).



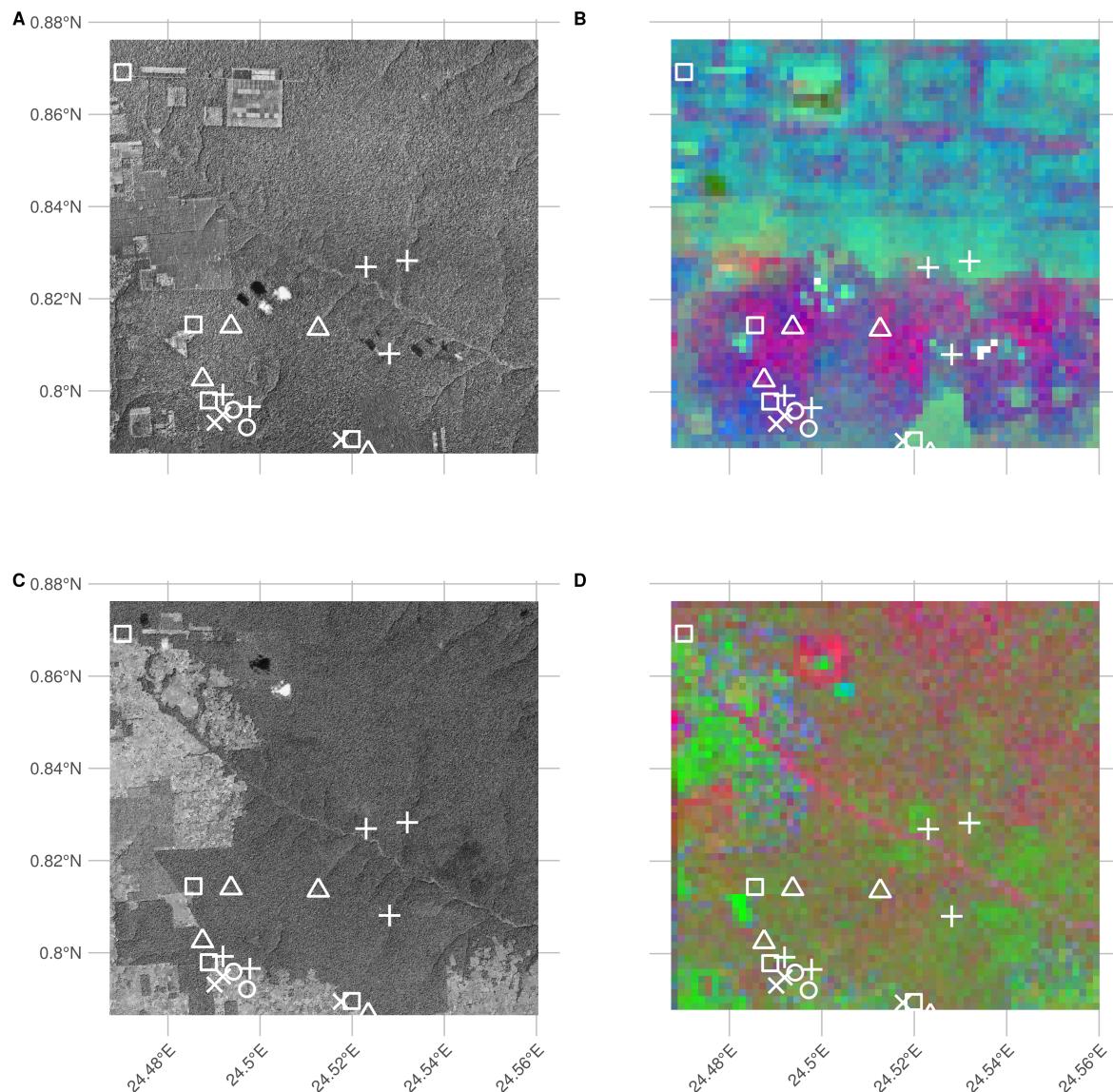
**Figure 9.** Boxplots comparing the first principal component (PC1) of a site based FOTO analysis across different forest types for both contemporary (Geo-Eye) and historical orthomosaic data.

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 difference 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 4-6**).

Furthermore, visual inspection of the scene wide analysis suggests historical scenes do not show landscape wide canopy features (Figure 10 A-B), unlike the contemporary scene (Figure 10 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 10D), 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 [22,23,49]. Despite the difficulties in recovering this data and its limitations, such as invisible disturbances [50], 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



**Figure 10.** 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.

314 first time, historical remote sensing data in order to quantify (long term) land-use and land-cover  
315 change and canopy structural properties in the central Congo Basin. Despite these successes some  
316 methodological and research considerations remain.

317 *4.1. Methodological considerations*

318 *4.1.1. Data recovery challenges*

319 In our study the archive data recovered was limited to contact prints and therefore did not  
320 represent the true resolution of the original negative. In addition, analogue photography clearly  
321 produces a distinct softness compared to digital imagery (Figure 8). Despite favourable nadir image  
322 acquisitions [51] image softness combined with illumination effects between flight paths, and the  
323 self-similar nature of vast canopy expanses [52–54], limited our ability to provide wall-to-wall  
324 coverage of the entire dataset containing 334 images. Few man made features in the scenes also  
325 made georeferencing challenging. Although the village of Yangambi provided a range of buildings  
326 as (hard-edge) references, other areas within the central Congo Basin might have fewer permanent  
327 structures and would require the use of soft-edged landscape features (e.g. trees, river outflows).  
328 Research has shown that soft-edged features can help georeference scenes even when containing few  
329 man-made features [55]. Our two step georeferencing approach resulted in a referencing accuracy of  
330  $\sim 4.7 \pm 4.3$  m across reference points. However, it shoud be noted that referencing accuracy of the final  
331 scene is less constrained toward the edges of the scene.

332 *4.1.2. LULC classification & validation*

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

343 4.1.3. Scaling opportunities

344 Our approach uses broadly defined homogeneous polygons to construct a balanced dataset of  
345 synthetic landscapes. The methodology is analogous to the use of sparse labelling as used by  
346 Buscombe and Ritchie [56] and contrasts with the standard methodologies which generally extract  
347 pixel (windows) [22] or delineate land cover classes [23] to drive a classifier or analysis. More so, the  
348 use of heavy image augmentation during model training sidesteps texture representation issues which  
349 affect classification of image scenes with inconsistent illumination or sharpness [25] or ad-hoc feature  
350 engineering [22]. The use of synthetic landscapes allowed us to account for most, but not all, of the  
351 variability within our orthomosaic. Our analysis has shown that despite being trained on historical  
352 data our model could map contemporary forest cover in remote sensing data with similar spatial and  
353 spectral characteristics (Figure 7), suggesting that the classifier consistently works across both space  
354 and time. We acknowledge that the use of synthetic landscapes is limited by the available homogeneous  
355 areas to sample from and the number of classes. Yet, the latter should not be a constraint as previous  
356 research efforts have focussed on simple forest cover maps [1].

357 4.2. Research context

358 4.2.1. Long term changes in LULC and Above Ground Carbon

359 Our analysis shows that the majority of deforestation around Yangambi happened toward the late  
360 1950's (~16200 ha). Considerable **reforestation** has occurred since the aerial survey was executed (~9918  
361 ha), and socio-economic instability prevented further large scale forest exploitation. In particular, many  
362 plantations have reached maturity and forest has re-established in previously cleared or disturbed  
363 areas. The majority of this **reforestation** takes the form of isolated patches of forest but is offset by  
364 further deforestation of previously untouched forest. Generally, the function and structure of forests  
365 can be influenced by forest edges that are located up to 1km away however most effects are pronounced  
366 within the first 300 m from the edge [57]. Our analysis of edge effects on AGC has shown that the  
367 influence is negligible 200 m away from the edge. Phillips *et al.* [58] have shown similar weak responses  
368 to edge effects in the Amazon forest. Due to a lack of data on the extent (depth) of edge effects and  
369 their influence on AGC beyond 200 m we did not include any estimates of carbon loss or gain within  
370 these zones. However, it must be stated that the influence of edges throughout the landscape was not  
371 marginal as these areas would account for 13151 ha. Thus edges could have a substantial negative [6]  
372 or positive [59] influence on AGC. Similarly, uncertainties in how to explicitly correct for plantations  
373 in the landscape present a further challenge. Thus, although our estimates are only indicative they do

374 underscore the important influence of landscape structure when carbon accounting. However, our  
375 findings do not indicate that deforestation in Congo basin is declining, on the contrary.

376 Over the past half century there has been a clear shift in land use in Yangambi (Figure 5). Land  
377 use has shifted away from for example a regular fishbone deforestation pattern that emerges when  
378 (large scale) agricultural interests dominate the landscape [60], to a more fragmented landscape  
379 (Figure 5D). The former is consistent with historical land management at the time of the aerial  
380 survey [46]. These regular patterns have since been reversed because of a decrease in large scale  
381 intensive agriculture, replaced by ad-hoc small scale subsistence farming with large perimeter to area  
382 relationships (i.e. ragged edges). Consequently, edge effects in the current landscape are far more  
383 pronounced than in the historical scene.

384 Visual inspection of the images also suggests that reforestation within the historically cleared areas  
385 and experimental plots is not necessarily limited to areas far removed from more densely populated  
386 areas. For example, large reforested areas exist close to the Congo stream and Yangambi village itself  
387 (Fig. 5). Here, regional political components, such as land leases and large scale ownership could  
388 have played a role in safeguarding some of these areas for rewilling or sustainable management  
389 [61,62]. Despite widespread anthropogenic influences throughout the tropics [31] the retention of  
390 these forested areas show the potential of explicit or implicit protective policy measures (e.g. INERA  
391 concessions, Bustillo *et al.* [46]) on a multi-decadal time scale. **Reforestation** in non-continuous areas  
392 within Yangambi could increase landscape connectivity and help increase biodiversity [12].

393 Our analysis therefore provides an opportunity to highlight and study those regions that have  
394 previously suffered confirmed long-term disturbances, and those that have been restored since.  
395 Assessments of old plantations and recovering clear-cut forests can serve as a guide to help estimate  
396 carbon storage capacity and forest recovery rates in managed and unmanaged conditions [18,20,63]  
397 over the mid- to long-term, in support of rewilling and general forest restoration [12,61,62]. In addition,  
398 mapping long-term edge effects can further support research into issues such as receding forest edges  
399 [57].

#### 400 4.2.2. Canopy structure & FOTO texture analysis

401 Finally, the FOTO technique used to quantify relationships between canopy structure and forest  
402 characteristics rendered no valuable insights of either the historical orthomosaic or recent Geo-Eye  
403 scene. Similarly weak correlations were found previously by Solórzano *et al.* [64]. In contrast, site  
404 based texture metric statistics did show correspondence between historical and contemporary satellite  
405 imagery. None of them were either consistent or significant. Although visual interpretation shows  
406 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. **Differences in PC between forest types (e.g. mono-dominant vs. mixed, Figure 9) corroborate that texture can serve as a basis for LULC classification. However, inflexibility on part of the FOTO technique in dealing with non-standardized (historical) data, or scaling these results to AGC values, limits its use case.** 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 quantify long-term changes in LULC and AGC. We showed a shift from previously highly structured industrial deforestation of large areas for plantation purposes, to discrete smallholder clearing for farming, increasing landscape fragmentation but also opportunities for substantial regrowth. Efforts to quantify canopy texture features and their link to AGC had limited to no success. Our analysis provides insights into the rate at which deforestation and reforestation has taken place over a multi-decadal scale in the central Congo basin. As such, it provides a useful historical context while interpreting past and ongoing forest research in the area.

## 6. Additional Information and Declarations

### 6.1. Data availability

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

437 6.2. *Code availability*

438 All analysis code is available as an R / python [65] projects (<https://khufkens.github.io/orthodrc>  
439 & [https://khufkens.github.io/orthodrc\\_cnn/](https://khufkens.github.io/orthodrc_cnn/)). The analysis relied heavily on the ‘raster’ [66],  
440 ‘RStoolbox’ [67], and ‘landscapemetrics’ [43] packages, while post-processing and plotting was  
441 facilitated by the ‘tidyverse’ ecosystem [68], ‘ggthemes’ [69], ‘scales’ [70] and ‘cowplot’ [71]. Additional  
442 plotting elements were formatted or provided by ‘sf’ [72] and ‘rnatural-earth’ [73] packages, respectively.  
443 I’m grateful for the contributions to the scientific community by the developers of these packages.

444 **Acknowledgments:** This research was supported through the Belgian Science Policy office COBECORE project  
445 (BELSPO; grant BR/175/A3/COBECORE). K.H. acknowledges funding from the European Union Marie  
446 Skłodowska-Curie Action (project number 797668). K.H. grateful for the support of T.d.H. volunteering his  
447 time in scanning the images.

448 **Author Contributions:** K.H. conceived and designed the study, analyzed the data, prepared figures, tables and  
449 authored the final draft of the manuscript. T.d.H. scanned all image data. E.K. and T.d.H. provided plot based  
450 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.

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

454 **References**

- 455 1. Hansen, M.C.; Potapov, P.V.; Moore, R.; Hancher, M.; Turubanova, S.A.; Tyukavina, A.; Thau, D.; Stehman,  
456 S.V.; Goetz, S.J.; Loveland, T.R.; Kommareddy, A.; Egorov, A.; Chini, L.; Justice, C.O.; Townshend,  
457 J.R.G. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* **2013**, *342*, 850–853.  
458 doi:10.1126/science.1244693.
- 459 2. Houghton, R.A.; House, J.I.; Pongratz, J.; van der Werf, G.R.; DeFries, R.S.; Hansen, M.C.; Le Quéré, C.;  
460 Ramankutty, N. Carbon emissions from land use and land-cover change. *Biogeosciences* **2012**, *9*, 5125–5142.  
461 doi:10.5194/bg-9-5125-2012.
- 462 3. Tyukavina, A.; Baccini, A.; Hansen, M.C.; Potapov, P.V.; Stehman, S.V.; Houghton, R.A.; Krylov, A.M.;  
463 Turubanova, S.; Goetz, S.J. Aboveground carbon loss in natural and managed tropical forests from 2000 to  
464 2012. *Environmental Research Letters* **2015**, *10*, 074002. doi:10.1088/1748-9326/10/7/074002.
- 465 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.;  
466 Randerson, J.T. CO<sub>2</sub> emissions from forest loss. *Nature Geoscience* **2009**, *2*, 737–738. doi:10.1038/ngeo671.
- 467 5. Fauset, S.; Gloor, M.U.; Aidar, M.P.M.; Freitas, H.C.; Fyllas, N.M.; Marabesi, M.A.; Rochelle, A.L.C.;  
468 Shenkin, A.; Vieira, S.A.; Joly, C.A. Tropical forest light regimes in a human-modified landscape. *Ecosphere*  
469 **2017**, *8*, e02002. doi:10.1002/ecs2.2002.
- 470 6. Brinck, K.; Fischer, R.; Groeneveld, J.; Lehmann, S.; Dantas De Paula, M.; Pütz, S.; Sexton, J.O.; Song, D.;  
471 Huth, A. High resolution analysis of tropical forest fragmentation and its impact on the global carbon  
472 cycle. *Nature Communications* **2017**, *8*. doi:10.1038/ncomms14855.
- 473 7. Didham, R.K. Edge Structure Determines the Magnitude of Changes in Microclimate and Vegetation  
474 Structure in Tropical Forest Fragments. *Biotropica* **1999**, *31*, 17–30.
- 475 8. Laurance, W.F.; Delamônica, P.; Laurance, S.G.; Vasconcelos, H.L.; Lovejoy, T.E. Rainforest fragmentation  
476 kills big trees. *Nature* **2000**, *404*, 836–836. doi:10.1038/35009032.
- 477 9. Magnago, L.F.S.; Magrach, A.; Laurance, W.F.; Martins, S.V.; Meira-Neto, J.A.A.; Simonelli, M.; Edwards,  
478 D.P. Would protecting tropical forest fragments provide carbon and biodiversity cobenefits under REDD+?  
479 *Global Change Biology* **2015**, *21*, 3455–3468. doi:10.1111/gcb.12937.

- 480 10. Poorter, L.; Bongers, F. Leaf traits are good predictors of plant performance across 53 rain forest species.  
481 *Ecology* **2006**, *87*, 1733–1743. doi:10.1890/0012-9658(2006)87[1733:LTAGPO]2.0.CO;2.
- 482 11. Barlow, J.; Lennox, G.D.; Ferreira, J.; Berenguer, E.; Lees, A.C.; Nally, R.M.; Thomson, J.R.; Ferraz, S.F.d.B.;  
483 Louzada, J.; Oliveira, V.H.F.; Parry, L.; Ribeiro de Castro Solar, R.; Vieira, I.C.G.; Aragão, L.E.O.C.; Begotti,  
484 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.;  
485 Pardini, R.; Silveira, J.M.; Vaz-de Mello, F.Z.; Veiga, R.C.S.; Venturieri, A.; Gardner, T.A. Anthropogenic  
486 disturbance in tropical forests can double biodiversity loss from deforestation. *Nature* **2016**, *535*, 144–147.  
487 doi:10.1038/nature18326.
- 488 12. Van de Perre, F.; Willig, M.R.; Presley, S.J.; Bapeamoni Andemwana, F.; Beeckman, H.; Boeckx, P.; Cooleman,  
489 S.; de Haan, M.; De Kesel, A.; Dessein, S.; Grootaert, P.; Huygens, D.; Janssens, S.B.; Kearsley, E.; Kabeya,  
490 P.M.; Leponce, M.; Van den Broeck, D.; Verbeeck, H.; Würsten, B.; Leirs, H.; Verheyen, E. Reconciling  
491 biodiversity and carbon stock conservation in an Afrotropical forest landscape. *Science Advances* **2018**,  
492 *4*, eaar6603. doi:10.1126/sciadv.aar6603.
- 493 13. Mitchard, E.T.A. The tropical forest carbon cycle and climate change. *Nature* **2018**, *559*, 527–534.  
494 doi:10.1038/s41586-018-0300-2.
- 495 14. Ploton, P.; Pélassier, R.; Proisy, C.; Flavenot, T.; Barbier, N.; Rai, S.N.; Couteron, P. Assessing aboveground  
496 tropical forest biomass using Google Earth canopy images. *Ecological Applications* **2012**, *22*, 993–1003.
- 497 15. Couteron, P.; Pelissier, R.; Nicolini, E.a.; Paget, D. Predicting tropical forest stand structure parameters  
498 from Fourier transform of very high-resolution remotely sensed canopy images. *Journal of Applied Ecology*  
499 **2005**, *42*, 1121–1128. doi:10.1111/j.1365-2664.2005.01097.x.
- 500 16. Barbier, N.; Couteron, P.; Proisy, C.; Malhi, Y.; Gastellu-Etchegorry, J.P. The variation of apparent crown  
501 size and canopy heterogeneity across lowland Amazonian forests. *Global Ecology and Biogeography* **2010**,  
502 *19*, 72–84. doi:10.1111/j.1466-8238.2009.00493.x.
- 503 17. DeFries, R.S.; Houghton, R.A.; Hansen, M.C.; Field, C.B.; Skole, D.; Townshend, J. Carbon emissions from  
504 tropical deforestation and regrowth based on satellite observations for the 1980s and 1990s. *Proceedings of  
505 the National Academy of Sciences* **2002**, *99*, 14256–14261. doi:10.1073/pnas.182560099.
- 506 18. Achard, F.; Beuchle, R.; Mayaux, P.; Stibig, H.J.; Bodart, C.; Brink, A.; Carboni, S.; Desclée, B.; Donnay, F.;  
507 Eva, H.D.; Lupi, A.; Raši, R.; Seliger, R.; Simonetti, D. Determination of tropical deforestation rates and  
508 related carbon losses from 1990 to 2010. *Global Change Biology* **2014**, *20*, 2540–2554. doi:10.1111/gcb.12605.
- 509 19. Ramankutty, N.; Foley, J.A. Estimating historical changes in global land cover: Croplands from 1700 to  
510 1992. *Global Biogeochemical Cycles* **1999**, *13*, 997–1027. doi:10.1029/1999GB900046.
- 511 20. Sader, S.A.; Joyce, A.T. Deforestation Rates and Trends in Costa Rica, 1940 to 1983. *Biotropica* **1988**, *20*, 11.  
512 doi:10.2307/2388421.
- 513 21. Willcock, S.; Phillips, O.L.; Platts, P.J.; Swetnam, R.D.; Balmford, A.; Burgess, N.D.; Ahrends, A.; Bayliss,  
514 J.; Doggart, N.; Doody, K.; Fanning, E.; Green, J.M.H.; Hall, J.; Howell, K.L.; Lovett, J.C.; Marchant, R.;  
515 Marshall, A.R.; Mbilinyi, B.; Munishi, P.K.T.; Owen, N.; Topp-Jorgensen, E.J.; Lewis, S.L. Land cover  
516 change and carbon emissions over 100 years in an African biodiversity hotspot. *Global Change Biology* **2016**,  
517 *22*, 2787–2800. doi:10.1111/gcb.13218.
- 518 22. Song, D.X.; Huang, C.; Sexton, J.O.; Channan, S.; Feng, M.; Townshend, J.R. Use of Landsat and  
519 Corona data for mapping forest cover change from the mid-1960s to 2000s: Case studies from the Eastern  
520 United States and Central Brazil. *ISPRS Journal of Photogrammetry and Remote Sensing* **2015**, *103*, 81–92.  
521 doi:10.1016/j.isprsjprs.2014.09.005.
- 522 23. Nita, M.D.; Munteanu, C.; Gutman, G.; Abrudan, I.V.; Radloff, V.C. Widespread forest cutting in the  
523 aftermath of World War II captured by broad-scale historical Corona spy satellite photography. *Remote  
524 Sensing of Environment* **2018**, *204*, 322–332. doi:10.1016/j.rse.2017.10.021.
- 525 24. Buitewerf, R.; Bond, W.J.; Stevens, N.; Trollope, W.S.W. Increased tree densities in South African  
526 savannas: >50 years of data suggests CO<sub>2</sub> as a driver. *Global Change Biology* **2012**, *18*, 675–684.  
527 doi:10.1111/j.1365-2486.2011.02561.x.
- 528 25. Hudak, A.T.; Wessman, C.A. Textural Analysis of Historical Aerial Photography to Characterize Woody  
529 Plant Encroachment in South African Savanna. *Remote Sensing of Environment* **1998**, *66*, 317–330.  
530 doi:10.1016/S0034-4257(98)00078-9.

- 531 26. Frankl, A.; Seghers, V.; Stal, C.; De Maeyer, P.; Petrie, G.; Nyssen, J. Using image-based modelling  
532 (SfM-MVS) to produce a 1935 ortho-mosaic of the Ethiopian highlands. *International Journal of Digital Earth*  
533 **2015**, *8*, 421–430. doi:10.1080/17538947.2014.942715.
- 534 27. Nyssen, J.; Petrie, G.; Mohamed, S.; Gebremeskel, G.; Seghers, V.; Debever, M.; Hadgu, K.M.; Stal, C.; Billi,  
535 P.; Demaeyer, P.; Haile, M.; Frankl, A. Recovery of the aerial photographs of Ethiopia in the 1930s. *Journal*  
536 *of Cultural Heritage* **2016**, *17*, 170–178. doi:10.1016/j.culher.2015.07.010.
- 537 28. Lewis, S.L.; Lopez-Gonzalez, G.; Sonké, B.; Affum-Baffoe, K.; Baker, T.R.; Ojo, L.O.; Phillips, O.L.; Reitsma,  
538 J.M.; White, L.; Comiskey, J.A.; Djuikouo K, M.N.; Ewango, C.E.N.; Feldpausch, T.R.; Hamilton, A.C.;  
539 Gloor, M.; Hart, T.; Hladik, A.; Lloyd, J.; Lovett, J.C.; Makana, J.R.; Malhi, Y.; Mbago, F.M.; Ndangalasi,  
540 H.J.; Peacock, J.; Peh, K.S.H.; Sheil, D.; Sunderland, T.; Swaine, M.D.; Taplin, J.; Taylor, D.; Thomas, S.C.;  
541 Votere, R.; Wöll, H. Increasing carbon storage in intact African tropical forests. *Nature* **2009**, *457*, 1003–1006.  
542 Publisher: Earth and Biosphere Institute, School of Geography, University of Leeds, Leeds LS2 9JT, UK.  
543 s.l.lewis@leeds.ac.uk.
- 544 29. Butsic, V.; Baumann, M.; Shortland, A.; Walker, S.; Kuemmerle, T. Conservation and conflict in the  
545 Democratic Republic of Congo: The impacts of warfare, mining, and protected areas on deforestation.  
546 *Biological Conservation* **2015**, *191*, 266–273. doi:10.1016/j.biocon.2015.06.037.
- 547 30. Kearsley, E.; de Haulleville, T.; Hufkens, K.; Kidimbu, A.; Toirambe, B.; Baert, G.; Huygens, D.; Kebede,  
548 Y.; Defourny, P.; Bogaert, J.; Beeckman, H.; Steppe, K.; Boeckx, P.; Verbeeck, H. Conventional tree  
549 height-diameter relationships significantly overestimate aboveground carbon stocks in the Central Congo  
550 Basin. *Nature communications* **2013**, *4*, 2269. doi:10.1038/ncomms3269.
- 551 31. Lewis, S.L.; Maslin, M.A. Defining the Anthropocene. *Nature* **2015**, *519*, 171–180. doi:10.1038/nature14258.
- 552 32. Bauters, M.; Ampoorter, E.; Huygens, D.; Kearsley, E.; De Haulleville, T.; Sellan, G.; Verbeeck, H.; Boeckx,  
553 P.; Verheyen, K. Functional identity explains carbon sequestration in a 77-year-old experimental tropical  
554 plantation. *Ecosphere* **2015**, *6*, art198. doi:10.1890/ES15-00342.1.
- 555 33. Zuiderveld, K. Contrast Limited Adaptive Histogram Equalization. In *Graphics GEMs IV*; Academic Press:  
556 San Diego, CA, USA, 1994; pp. 474–485.
- 557 34. Ullman, S. The Interpretation of Structure from Motion. *Proceedings of the Royal Society of London. Series B,*  
558 *Biological Sciences* **1979**, *203*, 405–426.
- 559 35. QGIS Development team,. QGIS Geographic Information System., 2018. Available at: <http://qgis.osgeo.org>.
- 560 36. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation.  
561 In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*; Navab, N.; Hornegger, J.;  
562 Wells, W.M.; Frangi, A.F., Eds.; Springer International Publishing: Cham, 2015; Vol. 9351, pp. 234–241.  
563 doi:10.1007/978-3-319-24574-4\_28.
- 564 37. Chollet, F. Keras, 2015. Available at: <https://github.com/fchollet/keras>.
- 565 38. Yakubovskiy, P. Segmentation Models, 2019. Available at: [https://github.com/qubvel/segmentation\\_models](https://github.com/qubvel/segmentation_models).
- 566 39. Martín Abadi.; Ashish Agarwal.; Paul Barham.; Eugene Brevdo.; Zhifeng Chen.; Craig Citro.; Greg S.  
567 Corrado.; Andy Davis.; Jeffrey Dean.; Matthieu Devin.; Sanjay Ghemawat.; Ian Goodfellow.; Andrew  
568 Harp.; Geoffrey Irving.; Michael Isard.; Jia, Y.; Rafal Jozefowicz.; Lukasz Kaiser.; Manjunath Kudlur.; Josh  
569 Levenberg.; Dandelion Mané.; Rajat Monga.; Sherry Moore.; Derek Murray.; Chris Olah.; Mike Schuster.;  
570 Jonathon Shlens.; Benoit Steiner.; Ilya Sutskever.; Kunal Talwar.; Paul Tucker.; Vincent Vanhoucke.; Vijay  
571 Vasudevan.; Fernanda Viégas.; Oriol Vinyals.; Pete Warden.; Martin Wattenberg.; Martin Wicke.; Yuan Yu.;  
572 Xiaoqiang Zheng. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, 2015. Available  
573 at: <https://www.tensorflow.org/>.
- 574 40. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv:1412.6980 [cs]* **2017**. arXiv:  
575 1412.6980.
- 576 41. Dice, L.R. Measures of the Amount of Ecologic Association Between Species. *Ecology* **1945**, *26*, 297–302.  
577 doi:10.2307/1932409.
- 578 42. Dale, M.R.T. *Spatial pattern analysis in plant ecology*; Cambridge studies in ecology, Cambridge University  
579 Press: Cambridge ; New York, 1999.
- 580 43. Hesselbarth, M.H.K.; Sciaiani, M.; With, K.A.; Wiegand, K.; Nowosad, J. landscapemetrics: an open-source  
581 R tool to calculate landscape metrics. *Ecography* **2019**.
- 582
- 583

- 584 44. Li, B.L. Fractal geometry applications in description and analysis of patch patterns and patch dynamics.  
585 *Ecological Modelling* **2000**, *132*, 33–50. ISBN: Ecol. Model.
- 586 45. Mandelbrot, B.B. *Fractals: Form, Chance and Dimension*; W.H. Freeman and co.: New York, US, 1977.
- 587 46. Bustillo, E.; Raets, L.; Beeckman, H.; Bourland, N.; Rousseau, M.; Hubau, W.; De Mil, T. Evaluation du  
588 potentiel énergétique de la biomasse aérienne ligneuse des anciennes plantations de l'INERA Yangambi.  
589 Technical report, CIFOR, 2018.
- 590 47. Couteron, P. Quantifying change in patterned semi-arid vegetation by Fourier analysis of digitized aerial  
591 photographs. *International Journal of Remote Sensing* **2002**, *23*, 3407–3425. doi:10.1080/01431160110107699.
- 592 48. Wilcoxon, F. Individual comparisons by ranking methods. *Biometrics* **1945**, *1*, 80–83.
- 593 49. Kadmon, R.; Harari-Kremer, R. Studying Long-Term Vegetation Dynamics Using Digital  
594 Processing of Historical Aerial Photographs. *Remote Sensing of Environment* **1999**, *68*, 164–176.  
595 doi:10.1016/S0034-4257(98)00109-6.
- 596 50. Peres, C.A.; Barlow, J.; Laurance, W.F. Detecting anthropogenic disturbance in tropical forests. *Trends in  
597 Ecology & Evolution* **2006**, *21*, 227–229. doi:10.1016/j.tree.2006.03.007.
- 598 51. Verhoeven, G. BRDF and its Impact on Aerial Archaeological Photography: BRDF and its impact on aerial  
599 archaeological photography. *Archaeological Prospection* **2017**, *24*, 133–140. doi:10.1002/arp.1559.
- 600 52. Park, J.Y.; Muller-Landau, H.C.; Lichstein, J.W.; Rifai, S.W.; Dandois, J.P.; Bohlman, S.A. Quantifying Leaf  
601 Phenology of Individual Trees and Species in a Tropical Forest Using Unmanned Aerial Vehicle (UAV)  
602 Images. *Remote Sensing* **2019**, *11*, 1534. doi:10.3390/rs11131534.
- 603 53. Simini, F.; Anfodillo, T.; Carrer, M.; Banavar, J.R.; Maritan, A. Self-similarity and scaling  
604 in forest communities. *Proceedings of the National Academy of Sciences* **2010**, *107*, 7658–7662.  
605 doi:10.1073/pnas.1000137107.
- 606 54. Sole, R.V.; Manrubia, S.C. Self-similarity in rain forests: evidence for a critical state. *Physical Review E* **1995**,  
607 *51*, 6250 – 6253.
- 608 55. Hughes, M.L.; McDowell, P.F.; Marcus, W.A. Accuracy assessment of georectified aerial photographs:  
609 Implications for measuring lateral channel movement in a GIS. *Geomorphology* **2006**, *74*, 1–16.  
610 doi:10.1016/j.geomorph.2005.07.001.
- 611 56. Buscombe, D.; Ritchie, A. Landscape Classification with Deep Neural Networks. *Geosciences* **2018**, *8*, 244.  
612 doi:10.3390/geosciences8070244.
- 613 57. Gascon, C.; Williamson, G.B.; Fonseca, G.A.B.d. Receding Forest Edges and Vanishing Reserves. *Science,  
614 New Series* **2000**, *288*, 1356–1358.
- 615 58. Phillips, O.L.; Rose, S.; Mendoza, A.M.; Vargas, P.N. Resilience of Southwestern Amazon Forests to  
616 Anthropogenic Edge Effects. *Conservation Biology* **2006**, *20*, 1698–1710. doi:10.1111/j.1523-1739.2006.00523.x.
- 617 59. Reinmann, A.B.; Hutyra, L.R. Edge effects enhance carbon uptake and its vulnerability to climate  
618 change in temperate broadleaf forests. *Proceedings of the National Academy of Sciences* **2017**, *114*, 107–112.  
619 doi:10.1073/pnas.1612369114.
- 620 60. Arima, E.Y.; Walker, R.T.; Perz, S.; Souza, C. Explaining the fragmentation in the Brazilian Amazonian  
621 forest. *Journal of Land Use Science* **2015**, pp. 1–21. doi:10.1080/1747423X.2015.1027797.
- 622 61. Arima, E.Y.; Barreto, P.; Araújo, E.; Soares-Filho, B. Public policies can reduce tropical deforestation: Lessons  
623 and challenges from Brazil. *Land Use Policy* **2014**, *41*, 465–473. doi:10.1016/j.landusepol.2014.06.026.
- 624 62. Larson, A.M. Forest tenure reform in the age of climate change: Lessons for REDD+. *Global Environmental  
625 Change* **2011**, *21*, 540–549. doi:10.1016/j.gloenvcha.2010.11.008.
- 626 63. Gourlet-Fleury, S.; Mortier, F.; Fayolle, A.; Baya, F.; Ouédraogo, D.; Bénédet, F.; Picard, N. Tropical forest  
627 recovery from logging: a 24 year silvicultural experiment from Central Africa. *Philosophical Transactions of  
628 the Royal Society B: Biological Sciences* **2013**, *368*, 20120302. doi:10.1098/rstb.2012.0302.
- 629 64. Solórzano, J.V.; Gallardo-cruz, J.A.; González, E.J.; Peralta-carreta, C.; Hernández-gómez, M.; Oca,  
630 A.F.m.D.; Cervantes-jiménez, L.G.; Solórzano, J.V.; Gallardo-cruz, J.A.; González, E.J.; Peralta-carreta,  
631 C.; Hernández-gómez, M.; Oca, A.F.m.D.; Cervantes-jiménez, L.G. Contrasting the potential of  
632 Fourier transformed ordination and gray level co-occurrence matrix textures to model a tropical swamp  
633 forest 's structural and diversity attributes. *Journal of Applied Remote Sensing* **2018**, *12*, 036006.  
634 doi:10.1117/1.JRS.12.036006.
- 635 65. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical  
636 Computing: Vienna, Austria, 2019.

- 637 66. Hijmans, R.J. raster: Geographic Data Analysis and Modeling, 2019. Available at: <https://CRAN.R-project.org/package=raster>.
- 638 67. Leutner, B.; Horning, N.; Schwalb-Willmann, J. RStoolbox: Tools for Remote Sensing Data Analysis, 2019.
- 639 Available at: <https://CRAN.R-project.org/package=RStoolbox>.
- 640 68. Wickham, H. tidyverse: Easily Install and Load the 'Tidyverse', 2017. Available at: <https://CRAN.R-project.org/package=tidyverse>.
- 641 69. Arnold, J.B. ggthemes: Extra Themes, Scales and Geoms for 'ggplot2', 2019. Available at: <https://CRAN.R-project.org/package=ggthemes>.
- 642 70. Wickham, H. scales: Scale Functions for Visualization, 2018. Available at: <https://CRAN.R-project.org/package=scales>.
- 643 71. Wilke, C.O. cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2', 2019. Available at: <https://CRAN.R-project.org/package=cowplot>.
- 644 72. Pebesma, E. Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal* 2018, 10, 439–446. doi:10.32614/RJ-2018-009.
- 645 73. South, A. rnaturalearth: World Map Data from Natural Earth, 2017. Available at: <https://CRAN.R-project.org/package=rnaturalearth>.
- 646
- 647
- 648
- 649
- 650
- 651
- 652

653 © 2020 by the authors. Submitted to *Remote Sens.* for possible open access publication  
654 under the terms and conditions of the Creative Commons Attribution (CC BY) license  
655 (<http://creativecommons.org/licenses/by/4.0/>).