TITLE

Mapping peat thickness and carbon stocks of the central Congo Basin using field data

AUTHOR LIST

Bart Crezee[1], Greta C. Dargie[1], Corneille E.N. Ewango[2,3], Edward T.A. Mitchard[4], Ovide Emba B.[5], Joseph Kanyama T.[2], Pierre Bola[5], Jean-Bosco N. Ndjango[3], Nicholas T. Girkin[6], Yannick E. Bocko[7], Suspense A. Ifo[8], Wannes Hubau[9,10], Dirk Seidensticker[11], Rodrigue Batumike[12], Gérard Imani[13], Aida Cuní-Sanchez[14,15], Christopher A. Kiahtipes[16], Judicaël Lebamba[17,18], Hans-Peter Wotzka[17], Hollie Bean[19], Timothy R. Baker[1], Andy J. Baird[1], Arnoud Boom[19], Paul J. Morris[1], Susan E. Page[19], Ian T. Lawson[20], Simon L. Lewis[1,21]

AFFILIATIONS

[1] School of Geography, University of Leeds, Leeds, LS2 9JT, UK

[2] Faculté de Gestion des Ressources Naturelles Renouvelables, Université de Kisangani, Kisangani, Democratic Republic of the Congo

[3] Faculté des Sciences, Université de Kisangani, Kisangani, Democratic Republic of the Congo

[4] School of GeoSciences, University of Edinburgh, Edinburgh, EH9 3FF, UK

[5] Institut Supérieur Pédagogique de Mbandaka, Mbandaka, Democratic Republic of the Congo

[6] School of Water, Energy and Environment, Cranfield University, Cranfield, MK43 0AL, UK

[7] Laboratoire de Botanique et Ecologie, Faculté des Sciences et Techniques, Université Marien Ngouabi, Brazzaville, Republic of the Congo

[8] École Normale Supérieure, Département des Sciences et Vie de la Terre, Laboratoire de Géomatique et d'Ecologie Tropicale Appliquée, Université Marien Ngouabi, Brazzaville, Republic of the Congo

[9] Department of Environment, Laboratory of Wood Technology, Ghent University, 9000 Ghent, Belgium

[10] Service of Wood Biology, Royal Museum for Central Africa, 3080 Tervuren, Belgium

[11] Department of Archaeology, Ghent University, 9000 Ghent, Belgium

[12] Département des Sciences de l'Environnement, Université du Cinquantenaire de Lwiro, Kabare, Sud-Kivu, Democratic Republic of the Congo

[13] Département de Biologie, Université Officielle de Bukavu, Bukavu, Democratic Republic of the Congo

[14] Department of Environment and Geography, University of York, York, YO10 5NG, UK

[15] Department of International Environmental and Development Studies (NORAGRIC), Norwegian University of Life Sciences, 1433 Ås, Norway

[16] Institute for the Advanced Study of Culture and the Environment, University of South Florida, Tampa, FL 33620, USA

[17] Department of Prehistoric Archaeology, University of Cologne, 50931 Köln, Germany

[18] Département de Biologie, Université des Sciences et Techniques de Masuku, BP 943, Franceville, Gabon

[19] School of Geography, Geology & the Environment, University of Leicester, Leicester, LE1 7RH, UK

[20] School of Geography and Sustainable Development, University of St Andrews, St Andrews, KY16 9AL, UK

[21] Department of Geography, University College London, London, WC1E 6BT, UK

CORRESPONDING AUTHOR

Correspondence and requests for materials should be addressed to B.C. (crezeebart@gmail.com).

ABSTRACT

The world's largest tropical peatland complex is found in the central Congo Basin. However, there is a lack of *in situ* measurements needed to understand the peat's distribution and the amount of carbon stored in it. So far, peat in this region has only been sampled in largely rain-fed interfluvial basins in the north of the Republic of the Congo. Here we present the first extensive field surveys of peat in the Democratic Republic of the Congo, which covers two-thirds of the estimated peatland area, including from previously undocumented river-influenced settings. We use field data from both countries to compute the first spatial models of peat thickness (mean 1.7 ± 0.9 m; maximum 5.6 m) and peat carbon density (mean 1,712 ± 634 Mg C ha^{-1}; maximum 3,970 Mg C ha^{-1}) for the basin. We show that the peatland complex covers 167,600 km^2, 15% more than previously estimated, and that 29.0 Pg C is stored belowground in peat across the region (95% confidence interval, 26.3-32.2 Pg C). Our measurement-based constraints give high confidence of globally significant peat carbon stocks in the central Congo Basin, totalling approximately one-third of the world's tropical peat carbon. Only 8% of this peat carbon lies within nationally protected areas, suggesting its vulnerability to future land-use change.

MAIN TEXT

Peatlands cover just 3% of Earth's land surface[1], yet store an estimated 600 Pg of carbon (C)[2,3], approximately one-third of Earth's soil carbon[4]. While most peatlands are located in the temperate and boreal zones[1], recent research is revealing the existence of tropical peatlands with high carbon densities[1,2,5,6]. Tropical peatlands are vulnerable to drainage and drying, with subsequent fires resulting in large carbon emissions from degraded peatlands, particularly in Southeast Asia[3,6–8].

In the central depression of the Congo basin (the 'Cuvette Centrale') the only field-verified peatland map to date reported that peat underlies 145,500 km^2 of swamp forests, making this the world's largest tropical peatland complex[9]. The field data used in this estimate are from northern Republic of the Congo (ROC), yet two-thirds of the central Congo Basin peatlands are predicted to be found in neighbouring Democratic Republic of the Congo (DRC)[9], sometimes hundreds of kilometres from existing field data (Figure 1a). Similarly, peat carbon stocks are estimated to be 30.6 Pg C, but the lower confidence interval is just 6 Pg C (ref. [9]). Thus, it is unclear if the central Congo peatlands are truly as extensive or deep as suggested, and it is unclear whether they store globally significant quantities of carbon.

Uncertainties are further compounded by a limited understanding of the processes that determine peat formation in central Congo, particularly hydrology[9,10]. Peat has only been systematically documented in interfluvial basins in ROC[9,11], where an absence of annual flood waves[9], modest domes[12], and remotely-sensed water-table depths[13] all suggest peatlands are largely rain-fed and receive little river water input. However, peat is also predicted in other hydro-geomorphological settings[9], including what appear to be river-influenced regions close to the Congo River mainstem and dendritic-patterned valley-floors along some of its left-bank tributaries[9] (Figure 1a). These areas of swamp forest are likely seasonally inundated[14] to depths up to 1.5 m during the main wet season[15], suggesting seasonal river flooding and/or upland runoff as key sources of water. Whether peat accumulates under these river-influenced conditions is currently unknown.

Here, we present the first *in situ* data on peat presence, thickness, and carbon density (mass per unit area) from the central Congo Basin in DRC. We specifically investigated the river-influenced swamp forests along the Congo River and its Ruki, Busira and Ikelemba tributaries that contrast with previous data collection from interfluvial basins[9] (Figure 1a). Every 250 m along 18 transects, we recorded vegetation characteristics, peat presence and thickness. We targeted a first group of ten transects in locations highly likely to contain peat, to help test hypotheses (detailed in Supplementary Table 1) about the role of vegetation, surface wetness, nutrient status, and topography in peat accumulation. To improve mapping capabilities, we sampled a second group of eight transects specifically to test preliminary maps that gave conflicting results or suspected false predictions of peat presence (detailed in Supplementary Table 1). We combine these new field measurements from DRC with previous transect records in ROC using the same protocols[9] and other ground-truth data (Supplementary Table 2) to produce (i) a second-generation map of peatland extent, (ii) a first-generation map of peat thickness, and (iii) a first-generation map of belowground peat carbon density for the central Congo Basin. These maps enable us to compute the first well-constrained estimate of total belowground peat carbon stocks in the world's largest tropical peatland complex.

Mapping peatland extent

We found peat along all ten hypothesis–testing transects in DRC that were predicted to be peatlands[9]. Our new field data shows that extensive carbon-rich peatlands are present in the forested wetlands of the DRC's Cuvette Centrale, including in geomorphologically distinct river-influenced regions predicted as peatlands by Dargie et al.[9].

The best-performing algorithm (Maximum Likelihood classifier, based on its ability to most accurately predict in regions with no training data; see Methods) was run 1,000 times on nine remotely-sensed datasets, using a random two-thirds of 1,736 ground-truth datapoints each time (Extended Data Figure 1), giving a median total peatland area for the central Congo Basin of 167,600 km^2 (95% CI, 159,400-175,100 km^2). This is 15% higher than the previous estimate[9]. We found that 90% of all pixels that are predicted as peat in the median map result were predicted as peat in at least 950 out of 1,000 runs (i.e., with ≥ 95% probability, either as hardwood- or palm-dominated peat swamp forest; Figure 1b), showing that peat predictions are consistent across model runs and thus are robust. Overall model performance, using the Matthews correlation coefficient, is 78.0% (95% CI, 74.2-81.6%).

Comparing our field results with the original first-generation map[9] shows that of the 382 locations assessed across DRC, 77.7% were correctly classified as either being peat swamp or not by the first-generation map[9]. Comparing our new map with the first-generation map[9] shows large areas of agreement (white in Figure 1c). However, we predict areas of peat which were previously not mapped[9], particularly around Lake

Mai-Ndombe and the Ngiri and upper Congo/Lulonga Rivers in DRC (red in Figure 1c). In addition, small areas of previously predicted peat deposits[9] are no longer predicted by our new model, particularly along the Sangha and Likouala-Mossaka Rivers in ROC (blue in Figure 1c). These areas of difference are likely areas of high uncertainty and should therefore be priorities for future fieldwork.

More formally, we compare our new second-generation map with the original map[9] using balanced accuracy (BA), which is similar to Matthews correlation coefficient but better suited for comparison across different datasets[16]. For our new map, median BA is 91.9% (95% CI, 90.2-93.6%), compared with 89.8% (86.0-93.4%) for the first-generation map[9]. The substantially smaller BA interval indicates improved confidence in our new peatland map, despite only a small increase in median BA. This is likely due to the effect of our larger sample size being partly offset by an increase in its spatial extent and ecological diversity, particularly data from the Congo River region, where all algorithms that we tested are underperforming (Supplementary Table 3). Overall, our *in situ* data from DRC, including from river-influenced settings that are being reported for the first time, confirm the central Congo Basin peatlands as the world's largest tropical peatland complex, and that DRC and ROC are the second and third most important countries in the tropics for peatland area after Indonesia[5], respectively (Extended Data Figure 2).

Mapping peat thickness and carbon density

We measured peat thickness at 238 locations in DRC (including 59 laboratory-verified measurements; Extended Data Figure 3), finding a mean (± s.d.) thickness of 2.4 (± 1.6) m and a maximum of 6.4 m. This shows that river-influenced peatlands can attain similar peat thickness as rain-fed interfluvial basins reported in ROC[9] (Table 1). There is no uniform increase in peat thickness with distance from the peatland margin (Extended Data Figure 4), with linear regression being only a modest fit (R^2 = 41.0%; RMSE = 1.21 m). Thus, we developed a Random Forest (RF) regression to estimate peat thickness, using 463 thickness measurements across both countries. Our final RF model includes four predictors after variable selection (see Methods): distance from the peatland margin, precipitation seasonality, climatic water balance (precipitation minus potential evapotranspiration), and distance from the nearest drainage point (R^2 = 93.4%; RMSE = 0.42 m). The RF model outperforms multiple linear regression with interactions using the same four variables (adj-R^2 = 73.6%, RMSE = 0.80 m; Extended Data Figure 5).

Spatially, we predict thick peat deposits in the centres of the largest interfluvial basins (far from peatland margins), and in smaller, river-influenced valley-floor peatlands along the Ruki/Busira Rivers (Figure 2a). The river valley's thick deposits are most likely driven by greater climatic water balance and lower precipitation seasonality in the eastern part of the Cuvette Centrale region (Extended Data Figure 6), plus potentially greater water inputs from nearby higher ground, which offsets the shorter distances from peatland margins. Our modelled results are consistent with our field data, as the two deepest peat cores are from the interfluvial Centre transect in ROC (5.9 m), and the river-influenced Bondamba transect on the Busira River in DRC (6.4 m). Overall, mean (± s.d.) modelled peat thickness (1.7 ± 0.9 m) is lower than our field measurements (2.4 ± 1.5 m; Table 1), as expected given our linear transects, which oversample deeper peat at the centre relative to the periphery in approximately ovoid peatlands. Areas of high uncertainty in peat thickness occur where distance from the margin is uncertain (Figure 2b). Our results contrast strongly with an “expert system approach” that assigned peat thickness values based on hydrological terrain relief alone and estimated a thickness of 6.5 ± 3.5 m for the central Congo Basin peatlands[17], compared to our field-derived estimate of 1.7 ± 0.9 m (Figure 2a).

After distance from the margin, precipitation seasonality and climatic water balance are the most important predictors of peat thickness in the RF model, reflecting the relative importance of rainfall inputs in peat accumulation in central Congo. This appears to differ from smaller-scale assessments in temperate[18] or other tropical peatlands[19], where surface topography (elevation and slope) are primary predictors of peat thickness. However, this is potentially merely an artefact of the spatial scale of the studies, as climate only varies over large scales. Alternatively, the relatively low rainfall in the central Congo Basin (~1700 mm yr^{-1}), compared to other tropical peatland regions (e.g., ~2,500-3,000 mm yr^{-1} in Northwest Amazonia and Southeast Asia)[9,20], may mean that peat thickness is more strongly related to climate in central Congo, as it implies greater exposure to (seasonal) drought conditions that may cross thresholds that negatively impact peat accumulation rates.

Peat bulk density measured across the central Congo Basin is 0.17 ± 0.06 g cm^{-3} (mean ± s.d.; n = 80 cores), and mean carbon concentration is 55.7 ± 3.2 % (n = 80; 56.6 [± 4.5] % for the 22 well-sampled cores). While peat bulk density is significantly lower in largely river-influenced sites than in rain-fed interfluvial basins ($P < 0.01$), no significant difference between these peatland types is found for either peat carbon concentration or carbon density (mass per unit area; Table 1).

We used the peat thickness, bulk density, and carbon concentration measurements to construct a linear peat thickness-carbon density regression (Extended Data Figure 7). We applied this regression model to our peat thickness map to spatially model carbon stocks per unit area (Figure 3a). Modelled belowground peat carbon density for the central Congo Basin is 1,712 ± 634 Mg C ha^{-1}, similar to the field-measured mean of 1,741 ± 1,186 Mg C ha^{-1} (mean ± s.d., n = 80; Table 1). This carbon density is approximately nine times the mean carbon stored in aboveground live tree biomass of African tropical moist forests (~198 Mg C ha^{-1})[21]. Compared with recently mapped peatlands in the lowland Peruvian Amazon (mean 867 Mg C ha^{-1})[22], the central Congo peatlands store almost twice as much carbon per hectare. Spatial patterns of peat carbon density (Figure 3a) and uncertainty (Figure 3b) follow similar patterns as peat thickness (Figures 2a and 2b).

Estimating basin-wide peat carbon stocks

Median estimated total peat carbon stock in the central Congo Basin is 29.0 Pg (95% CI, 26.3-32.2; Extended Data Figure 8a), based on bootstrapping the area estimate and peat thickness-carbon density regression. This is similar to the median 30.6 Pg C reported by Dargie et al.[9], but their lower 95% confidence interval was 6.3 Pg, which our study increases to 26.3 Pg. This constraint on the carbon stock estimate is possible because our larger field-based dataset allows a spatial modelling approach, so that we can sum carbon density across all peat pixels. Therefore, the possibility of low values of carbon storage in the central Congo peatlands can now confidently be discarded.

Our new results show that the central Congo Basin peatlands are a globally important carbon stock, harbouring approximately one-third of all the carbon stored in the world's tropical peatlands[5,9]. About two-thirds of this peat carbon is in DRC (19.6 Pg C; 95% CI, 17.9-21.9), and one-third in ROC (9.3 Pg C; 95% CI, 8.4-10.2; Extended Data Figure 2), which is equivalent to approximately 82% and 238% of each country's aboveground forest carbon stock, respectively[23]. The high peat carbon stocks are found across several administrative regions in both countries, with the largest stocks in DRC's Équateur province (Extended Data Figure 2). Sensitivity analysis shows that uncertainty in total peat carbon stock is now mostly driven by uncertainty in peatland area (Extended Data Figure 8b).

Because the central Congo peatlands are relatively undisturbed[24,25], our new maps of peatland extent, thickness and carbon density form a baseline description for the decade 2000-2010, given the remotely-sensed data used. Today, the peatlands of the central Congo Basin are threatened by hydrocarbon exploration, logging, palm oil plantations, hydroelectric dams and climate change[24,26]. While the peatlands are largely within a UN Ramsar Convention transboundary wetland designation, we estimate that only 2.4 Pg C in peat, just 8% of total stocks, currently lies within formal national-level protected areas (Extended Data Figures 9 and 10). Meanwhile, logging, mining, or palm oil concessions together overlie 7.4 Pg C in peat, or 26% of total stocks (Extended Data Figures 9 and 10), while hydrocarbon concessions cover almost the entire peatland complex[24,26].

Keeping the central Congo Basin peatlands wet is vital to prevent peat carbon being released to the atmosphere. The identification of extensive river-influenced peatlands suggests that there is more than one geomorphological setting where peat is found in the central Congo Basin. Further work is required to understand both the sources and flows of water in these river-influenced peatlands, specifically the relative contributions of water from precipitation, riverbank overflow, and run-off from higher ground to peat formation and maintenance. Given the current areas of formal protection of peatlands are largely centred around interfluvial basins, we suggest that additional protective measures will be needed to safeguard the newly identified river-influenced peatlands of the central Congo Basin. Keeping the central Congo peatlands free from disturbance would also help protect the rich biodiversity, including forest elephants, lowland gorillas, chimpanzees and bonobos[24,27,28], that form part of this globally important, but threatened ecosystem.

METHODS

Field data collection

Fieldwork was conducted in DRC between January 2018 and March 2020. Ten transects (4-11 km long) were installed, identical to Dargie et al.'s approach[9], in locations that were highly likely to be peatland. These were selected to help test hypotheses about the role of vegetation, surface wetness, nutrient status, and topography in peat accumulation (Figure 1a; Supplementary Table 1). A further eight transects (0.5-3 km long) were installed to assess our peat mapping capabilities (Figure 1a; Supplementary Table 1).

Every 250 m along each transect, landcover was classified as one of six classes: water, savanna, *terra firme* forest, non-peat forming seasonally inundated forest, hardwood-dominated peat swamp forests, or palm-dominated peat swamp forests. Peat swamp forest was classified as palm-dominated when > 50% of the canopy, estimated by eye, were palms (commonly *Raphia laurentii* or *Raphia sese*). In addition, several ground-truth points were collected at locations in the vicinity of each transect from the clearly identifiable landcover classes water, savanna, or *terra firme* forest.

Peat presence/absence was recorded every 250 m along all transects, and peat thickness (if present) was measured by inserting metal poles into the ground until the poles were prevented from going any further by the underlying mineral layer, identical to Dargie et al.'s pole-method[9]. Additionally, a core of the full peat profile was extracted every kilometre along the ten hypothesis-testing transects, if peat was present, with a

Russian-type corer (52-mm stainless steel Eijkelkamp model); these 63 cores were sealed in plastic for laboratory analysis.

Peat thickness laboratory measurements

Peat was defined as having an organic matter (OM) content of ≥ 65% and a thickness of ≥ 0.3 m (*sensu* Dargie et al.[9]). Therefore, down-core OM content of all 63 cores was analysed to measure peat thickness. The organic matter content of each 0.1-m thick peat sample was estimated via Loss-On-Ignition (LOI), whereby samples were heated at 550°C for 4h. The mass fraction lost after heating was used as an estimate of total OM content (% of mass). Peat thickness was defined as the deepest 0.1-m with OM ≥ 65%, after which there is a transition to mineral soil. Samples below this depth were excluded from further analysis. Rare mineral intrusions into the peat layer above this depth, where OM < 65% for a sample within the peat column, were retained for further analysis. In total, 59 out of 63 collected cores had LOI-verified peat thickness ≥ 0.3 m.

The pole-method used to estimate peat thickness in the field was calibrated against LOI-verified measurements, by fitting a linear regression model between all LOI-verified and pole-method peat thickness measurements sampled at the same location (93 sites across ROC and DRC, including 37 from ref. [9]). Three measurements from DRC with a Cook's distance > 4x the mean Cook's distance were excluded as influential outliers. Mean pole-method offset was significantly higher along the DRC transects (0.94 m) than along those in ROC (0.48 m; $P < 0.001$), due to the presence of softer alluvium substrate in river-influenced sites in DRC. We therefore added this grouping as a categorical variable to the regression. The resulting model (adj-R^2 = 0.95, $P < 0.001$; Extended Data Figure 3) was used to correct all pole-method measurements in each group for which no LOI-verified thickness was available: corrected peat thickness = -0.1760 + 0.8626 x (pole-method thickness) – 0.3284 x (country), with country dummy coded as: ROC (0) and DRC (1).

Carbon density estimates

To calculate carbon density (mass per unit area), estimates of carbon storage in each 0.1-m thick peat sample (thickness × bulk density × carbon concentration) were summed to provide an estimate of total carbon density per core (in Mg C ha^{-1}), identical to Dargie et al.[9]. We estimated carbon density for 80 peat cores (OM ≥ 65%, thickness ≥ 0.3 m), located every other kilometre along 18 transects, including 37 cores from the ten transects used for hypothesis testing in DRC, and 43 cores from transects in ROC[9].

Peat thickness of the 80 cores was obtained by laboratory LOI. To estimate peat bulk density, every other 0.1-m down-core, samples of a known peat volume were weighed after being dried for 24h at 105°C (n = 906). Bulk density (in g cm^{-3}) was then calculated by dividing the dry sample mass (in g) by the volume of the sample taken from the peat corer dimensions (in cm^3). Within each core, linear interpolation was used to estimate bulk density for the alternate 0.1m-thick samples of the core that were not measured.

For total carbon concentration (%), only the deepest core per transect, plus additional deep cores from the Lokolama transect (1) in DRC and Ekolongouma transect (3) in ROC (22 in total, 11 from DRC and 11 from ROC[9]) were sampled down-core. Every other 0.1-m thick sample was measured using an elemental analyser (Elementar Vario

MICRO Cube with thermal conductivity detection for all cores, except those from Boboka, Lobaka and Ipombo transects, which were analysed using Sercon ANCA GSL with isotope-ratio mass spectrometer detection, due to COVID-19 disruption). All samples (n = 422) were pre-dried for 48h at 40°C and ground to < 100 μm using a MM301 mixer mill. Again, linear interpolation was used within each core for the alternate samples that were not measured.

The remaining 58 cores had less-intensive carbon concentration sampling. We therefore interpolated the carbon concentration for each 0.1-m thick sample, because well-sampled cores show a consistent pattern with depth: an increase to a depth of about 0.5 m, followed by a long, very weak decline, and finally a strong decline over the deepest approximately 0.5 m of the core[9]. We used segmented regression on the 22 well-sampled cores (*segmented* package in R, version 1.3-1) to parameterize the three sections of the core, using the means of these relationships to interpolate carbon concentrations for the remaining 58 cores, following Dargie et al.[9].

To estimate carbon density from modelled peat thickness across the basin, we developed a regression model between peat thickness and per-unit-area carbon density using the 80 sampled cores. We compared linear regressions for normal, logarithmic-, and square root-transformed peat thickness, selecting the model with lowest AICc and highest R^2. A linear model with square root-transformed peat thickness was found to provide the best fit ($R^2 = 0.86$; $P < 0.001$; Extended Data Figure 7). Bootstrapping was applied (*boot* package in R, version 1.3-25) to assess uncertainty around the regression.

Modelling peatland extent

Satellites cannot detect peat directly. We therefore mapped vegetation and used field-based associations between peat and vegetation to infer peat presence[9,29]. Five landcover classes were used for the purpose of peatland mapping: water, savanna, palm-dominated peat swamp forest, hardwood-dominated peat swamp forest, and non-peat forming forest. In this classification, field recordings of non-peat forming seasonally inundated forest (< 30 cm thickness of ≥ 65% OM) were grouped together with field recordings of *terra firme* forest, which also does not form peat, to form the non-peat forming forest class. Our field recordings of hardwood- or palm-dominated peat swamp forest, by definition, consist of all forest sites that form peat, including any seasonally inundated forest that forms peat (≥ 30 cm of ≥ 65% OM).

A total of 1,736 ground-truth datapoints was used: 172 in water, 476 in savanna, 632 in non-peat forming forest (97 non-peat forming seasonally inundated forest, and 535 *terra firme* forest), 188 in palm-dominated peat swamp forest, and 268 in hardwood-dominated peat swamp forest (Extended Data Figure 1). This data comes from eight sources (Supplementary Table 2). First, ground-truth locations collected for this study using a GPS (Garmin GPSMAP 64s) at all transect sites in DRC for which a landcover class was determined (382 points). Second, published ground-truth data from nine transects in ROC (292 points)[9]. Third, 299 GPS locations of known savanna and *terra firme* forest landcover classes from archaeological research databases across the basin[30,31]. Fourth, 191 GPS locations from permanent long-term forest inventory plots of the African Tropical Rainforest Observation Network (AfriTRON), mostly from *terra firme* forest[32], retrieved from the ForestPlots database[33,34]. Fifth, 229 GPS datapoints from *terra firme* forest or savanna locations in and around Lomami National Park (*pers.*

comm., R.B., G.I. and A. C-S.). Sixth, 24 published savanna datapoints in and around Lomami NP[35]. Seventh, 23 published locations of savanna, *terra firme* forest, palm- or hardwood-dominated peat swamp forest in DRC[11]. Eighth, 296 datapoints from Google Earth for unambiguous savanna and water sites (middle of lakes or rivers), distributed across the region.

We used nine remote sensing products to map peat-associated vegetation (Supplementary Figure 1). Eight of these are identical to those used by Dargie et al.[9]: three optical products (Landsat 7 ETM+ bands 5 [SWIR 1], 4 [NIR], and 3 [Red]); three L-band Synthetic Aperture Radar products (ALOS PALSAR HV, HH, and HV/HH); and two topographic products (SRTM DEM [Digital Elevation Model] void-filled with ASTER GDEM v2 data, and slope; acquisition date 2000). To this, we added a HAND-index (Height Above Nearest Drainage point), which significantly improved model performances (median Matthews correlation coefficient [MCC]: 79.7%, compared with 77.8% or 75.6% for just DEM or HAND alone, respectively; $P < 0.001$).

HAND was derived from the SRTM DEM with Clubb et al.'s algorigthm[36], using the HydroSHEDS global river network at 15s resolution as reference product[37]. Alternative NASADEM- or MERIT DEM-derived[38–40] combinations of DEM, HAND and slope were tested with an initial subset of data in R, while keeping all other remote sensing products the same (median MCC: 79.0% and 75.1%, respectively), but did not significantly improve model performance compared with SRTM-derived products (80.9% median MCC; $P < 0.001$).

The Landsat bands are pre-processed, seamless cloud-free mosaics for ROC (composite of three years, 2000, 2005, 2010) and DRC (composite of six years, 2005-2010)[41]. These mosaics performed better than more recent basin-wide automated cloud-free Sentinel-2 mosaics that we developed (bands 5, 8A, 11; composite of five years, 2016-2020), likely because they contain less directional reflectance artefacts (the median MCC of 80.9% for the pre-processed Landsat mosaics is significantly higher than the 78.1% for our Sentinel-2 mosaics, $P < 0.005$).

The ALOS PALSAR radar bands are mosaics of mean values of annual JAXA composites for the years 2007-2010 (ref. [9]). More recent radar data (ALOS 2-PALSAR 2 HV, HH, HV/HH; 2015-2017) did not significantly improve model performances (median MCC 80.9% and 80.6%, respectively; $P < 0.01$). All remote sensing products were resized to a common 50 m grid, using a cubic convolution resampling method.

We then tested which classification algorithm to use, as more sophisticated algorithms might improve overall accuracy against our training dataset, but might also reduce regional accuracy of the map in areas far from test data, critical in this case given large areas of the central Congo peatland region remain unsampled.

Three supervised classification algorithms were tested in order of increasing complexity: Maximum Likelihood (ML), Support Vector Machine (SVM) and Random Forest (RF). We assessed each classifier using both a random and spatial cross-validation (CV) approach[42–44]. Random CV was implemented using stratified two-thirds Monte Carlo selection, whereby we 1,000 times randomly selected two-thirds of all datapoints per class as training data, to be evaluated against the remaining one-third per class as testing data.

Spatial CV was implemented by grouping all transects datapoints in four distinct hydro-geomorphological regions: (i) transects perpendicular to the blackwater Likouala-aux-Herbes River (n = 179 datapoints); (ii) transects perpendicular to the white-water Ubangi River (n = 113); (iii) transects perpendicular to the Congo River, intermediate between black and white-water (n = 123); and (iv) transects perpendicular to the blackwater Ruki, Busira and Ikelemba Rivers, plus other nearby transects (collectively named the Ruki group; n = 258). To each group we added ground-truth datapoints from other non-transect data sources (Supplementary Table 2) that belonged to the same map regions (n = 82, 27, 20, 113, respectively). We then tested 1,000 times how well each classifier performs in each of the four regions, when trained only on a stratified two-thirds Monte Carlo selection of the remaining datapoints (i.e., datapoints from the three other regional transect groups), plus ground-truth datapoints not associated with or near any transect group (n = 821; for example, the savanna and *terra firme* forest datapoints in Lomami National Park in DRC which are far [> 300 km] from any transect group).

Model performance was based on Matthews correlation coefficient (MCC) for binary peat/non-peat predictions (hardwood- and palm-dominated peat swamp forest classes combined into one peat class; water, savanna and non-peat forming forest combined into one non-peat class). We compared MCC, rather than popular metrics such as Cohen's kappa, F1-score or accuracy, because it is thought to be the most reliable evaluation metric for binary classifications[45,46]. We also computed balanced accuracy

(BA) from random cross-validation to compare with the first-generation map. While less robust than MCC, BA is independent of imbalances in the prevalence of positives/negatives in the data, thus allowing better comparison between classifiers trained on different datasets[16]. The best estimate of each accuracy metric or area estimate per model or region is the median value of 1,000 runs, alongside a 95% confidence interval.

In the case of SVM and RF, random CV models were implemented in Google Earth Engine (GEE)[47] using all nine remote sensing products. However, because ML is currently not supported by GEE, random CV with this algorithm was implemented in IDL-ENVI software (version 8.7-5.5), using a principal component analysis (PCA) to reduce the nine remote sensing products to six uncorrelated principal components to reduce computation time. All spatial CV models were implemented in R (*superClass* function from the *RStoolbox* package, version 0.2.6), with PCA also applied in the case of ML only. All RF models were trained using 500 trees, with three input products used at each split in the forest (the default, the square root of the number of variables). All SVM model were implemented with a radial basis function kernel, with all other parameters set to default values.

Comparison of the ML, SVM and RF models with Dargie et al.'s model performance[9], using balanced accuracy from random cross-validation, shows improved results only in the case of the ML classifier (Supplementary Table 3). Comparing MCC using the spatial CV approach, we found that the ML algorithm is also most transferable to regions for which we lack training data. While RF gives slightly better MCC with random CV, when no regions are omitted, spatial CV shows particularly poor predictive performance of this algorithm for the Congo and Ruki regions, when trained on data from the other regions. SVM has lowest MCC of all three classifiers with random CV, and also performs worst of all three in the Congo region with spatial CV.

Additionally, applying spatial CV to the largely interfluvial basin region (ROC transects; n = 401), and the largely river-influenced region (DRC transects; n = 540), also shows RF performs poorly (Supplementary Table 3). This further supports selecting the ML algorithm to produce our second-generation peat extent map of the central Congo peatlands. The final peatland extent estimate is then obtained as the median value (alongside 95% confidence interval) out of the combined hardwood- and palm-dominated peat swamp forest extent from 1,000 ML runs, each time trained with two-thirds of the ground-truth data.

Modelling peat thickness

A map of distance from the peatland margins was developed in GEE using the median ML peat probability map, i.e. the ML map with a 50% peat probability threshold (> 500 hardwood- or palm-dominated peat swamp predictions out of 1,000 runs). For each peat pixel in this binary classification, a cost function was used to calculate the Euclidean distance to the nearest non-peat pixel, after speckle and noise were removed using a 5x5 squared-kernel majority filter. Using this distance map, transects were found to have markedly different relationships between peat thickness and distance from the peatland margin, i.e. different slopes (n = 18, $P < 0.001$, Extended Data Figure 4). The modest linear fit (R^2 = 41.0%; RMSE = 1.21 m) cautions against a uniform regression between peat thickness and distance from the margin across the basin.

Instead, we developed a spatially-explicit Random Forest regression model to predict peat thickness, derived from 14 remotely-sensed potential covariates that may explain variation in peat thickness. These 14 variables included the nine optical, radar and topographic products used in the peatland extent analysis, as well as distance from the peatland margin, distance from the nearest drainage point (same reference network as for HAND)[37], precipitation seasonality[48], climatic water balance (mean annual precipitation[48] minus mean annual potential evapotranspiration[49]), and live woody aboveground biomass[50]. Ten of these variables were found to be significantly correlated with peat thickness (Kendall's τ, $P < 0.01$): all three optical bands, all three radar bands, distance from the peatland margin, distance from the nearest drainage point, precipitation seasonality, and climatic water balance. Applying stepwise backward selection, we tested combinations of these ten predictors by each time dropping one predictor out of the model in order from low to high variable importance, selecting as the best model the one with highest median R^2 and lowest median root mean square error (RMSE) obtained from 100 random (two-thirds) cross-validations. The importance of each variable was assessed by calculating Mean Decrease Impurity (MDI), the total decrease in the residual sum of squares of the regression after splitting on that variable, averaged over all decision trees in the random forest. Median MDI was calculated for each variable based on 100 random (two-thirds) cross-validations of the overall model containing all ten significant predictors.

The best model contained four predictors: distance from the peatland margin, distance to the nearest drainage point, climatic water balance (all positively correlated with peat thickness; Kendall's τ coefficient = 0.49, 0.15 and 0.13, respectively; $P < 0.001$ for all), and precipitation seasonality (negatively correlated with thickness; Kendall's $\tau = -0.11$, $P < 0.01$); see Extended Data Figure 6 for their spatial variability.

The RF regression was implemented in GEE with 500 trees and all other parameters set to default values. Predictor variables were resampled to 50 m resolution. As training data, we included all LOI-verified and corrected pole-method thickness measurements that fell within the masked map of > 50% peat probability (n = 463), including thickness > 0 and < 0.3 m from non-peat sites that could improve predictions of shallow peat deposits near the margins (n = 12).

Our final RF model (R^2 = 93.4%, RMSE = 0.42 m) had consistently smaller residuals compared to a multiple linear regression model containing the same four predictors with interaction effects (adj-R^2 = 73.6%, RMSE = 0.80 m; Extended Data Figure 5). It also performed better when testing out-of-sample performance, using 100 random two-thirds cross-validations of training data (median R^2 = 82.2%, RMSE = 0.68 m; and median adj-R^2 = 73.6%, RMSE = 0.85 m; for RF model and multiple linear regression, respectively).

For uncertainty on our thickness predictions, we first estimated area uncertainty by creating 100 different maps of distance from the peat margin, by randomly selecting (with replacement) a minimum peat probability threshold > 0% and < 100%, removing speckle and noise, and re-calculating the closest distance to the nearest non-peat pixel. We then combined the 100 distance maps each time with the three other selected predictors (precipitation seasonality, climatic water balance, distance from nearest drainage point) as input in a RF model to develop 100 different peat thickness maps. For these model runs, we included all available thickness measurements (> 0 m) that fell within each specific distance map. Each output map was masked to an area ≥ 0.3 m thickness, consistent with our peat definition. A map of median peat thickness (Figure 3a) and relative uncertainty (± half the width of the 95% CI as percentage of the median; Figure 3b) was then calculated for each pixel based on the 100 available thickness estimates.

Carbon stock estimates

We mapped carbon density across the central Congo Basin in GEE, by applying 20 bootstrapped thickness-carbon regressions that were normally distributed around the best fit (Extended Data Figure 7 6) to the 100 peat thickness maps from the RF regression model, generating a map of median carbon density out of 2,000 estimates (Figure 3a), together with relative uncertainty (± half the width of the 95% CI as percentage of the median; Figure 3b).

Total peat carbon stocks were computed in GEE by summing carbon density (in Mg ha^{-1}) over all 50 m grid squares defined as peat. To assess uncertainty around this estimate, we again combined the 100 peat thickness maps (i.e., uncertainty from area and thickness), with 20 bootstrapped thickness-carbon regressions (i.e., uncertainty from carbon density, including bulk density and carbon concentration). We thus obtained 2,000 peat carbon stock estimates for the total central Congo Basin peatland complex, which were used to estimate the mean, median and 95% CI (Extended Data Figure 8a).

Regional carbon stock estimates were similarly obtained for each sub-national administrative region (departments in ROC and provinces in DRC; Extended Data Figure 2), as well as national-level protected areas (national parks and nature/biosphere/community reserves)[51] and logging[52,53], mining[54,55] and palm oil[56–58] concessions (Extended Data Figures 9 and 10). As hydrocarbon concessions cover almost the whole peatlands area[24,26], they cover almost 100% of the central Congo peat carbon stocks.

Sensitivity analysis was performed by bootstrapping either the area, thickness, or carbon density component, whilst keeping the others constant (Extended Data Figure 8b). For area, we bootstrapped 100 randomly selected peatland area estimates; for thickness, 100 randomly selected two-thirds subsets of all thickness measurements; for carbon density, 20 normally distributed regression equations from the bootstrapped thickness-carbon relationship.

DATA AVAILABILITY

All map results from this study are available for download as raster files from https://congopeat.net/maps/. The supporting ground-truth data, peat thickness measurements, and carbon density measurements are available from https://github.com/CongoPeat/Peatland-mapping.git . The remote sensing datasets used are available for download from https://www.eorc.jaxa.jp/ALOS/en/dataset/fnf_e.htm (ALOS PALSAR and ALOS-2 PALSAR-2 25 m HV and HH data), http://osfac.net/ (OSFAC ROC and DRC 60 m Landsat ETM+ bands 5, 4 and 3 mosaics), and http://earthexplorer.usgs.gov/ (SRTM DEM 1-arc second and ASTER GDEM v2 1-arc second data).

CODE AVAILABILITY

The IDL-ENVI script to run the Maximum Likelihood peatland extent model is available from https://github.com/CongoPeat/Peatland-mapping.git. The scripts to run the peat thickness model and carbon stock calculations are available on Google Earth Engine: https://code.earthengine.google.com/?accept_repo=users/gybjc/Central_Congo_Peatlands_2022. All R code is available from the corresponding author upon request.

ACKNOWLEDGEMENTS

We sincerely thank the communities that hosted and assisted with our fieldwork in DRC: Lokolama, Bosukela, Mpama, Befale, Bonsole, Mweko, Mpeka, Bondamba, Bolengo, Boleke, Pombi, Boboka, Ipombo, Lobaka, Bolombo, and Bonzembo. We thank the Groupe d'action pour sauver l'homme et son environnement (GASHE), especially Julien Mathe, and Greenpeace Africa, especially Raoul Monsembula, for essential logistical support. We thank the government of the Democratic Republic of the Congo, the Province of Équateur, and the Ministry of Environment and Sustainable Development for assistance with our fieldwork. We thank Bolivard Bongwemisa, Juress Sando, Jean-Pierre Lokila, Papy Bosange, Felly Mongonga and Roger Kendewa for essential field support. David Milodowski and Adam Hastie provided modelling advice, Rachel Gasior, David Ashley, Martin Gilpin and Dave Wilson provided laboratory assistance, and Donna Hawthorne, George Biddulph, Shona Jenkins, Sofie Sjögersten, Guy Ziv and the CongoPeat network provided invaluable discussions and feedback. The work was funded by a NERC Large Grant to S.L.L. (“CongoPeat”, NE/R016860/1), a NERC Doctoral Training Partnership award to B.C. (“SPHERES DTP”, NE/L002574/1), and a Greenpeace Fund award to S.L.L.. JAXA, NASA, METI, USGS, ESA, OSFAC, and WWF are acknowledged for collecting and/or processing remote sensing data.

AUTHOR CONTRIBUTIONS

S.L.L., E.T.A.M., I.T.L., G.C.D., and S.E.P. conceived the study; B.C., G.C.D., S.L.L., E.T.A.M., I.T.L., S.E.P., S.A.I., C.E.N.E. and T.R.B. developed the study; B.C., G.C.D., S.L.L. C.E.N.E., O.E.B., P.B., J.K.T., N.T.G., and J-B.N.N. organised and conducted the fieldwork; Y.E.B., S.A.I., W.H., D.S., R.B., G.I., A.C-S., C.A.K., J.L. and H-P.W. provided additional data; B.C., G.C.D., A.B. and H.B. performed laboratory analyses; B.C. and E.T.A.M. analysed the remote sensing data and developed the models; B.C., S.L.L., E.T.A.M., G.C.D., A.J.B., T.R.B., P.J.M. and C.A.K. evaluated the results. B.C. and S.L.L. wrote the paper, with input from all co-authors.

COMPETING INTERESTS

The authors declare no competing interests.

TABLES

Table 1 | Field-measured and spatially modelled estimates of peat thickness, bulk density, carbon concentration, and carbon density in the central Congo Basin peatland complex.

	Field measurements *														**Spatial model †**					
	Peat thickness (m) #			**Peat bulk density ($g\ cm^{-3}$) §**				**Peat carbon concentration (%) ‡**				**Peat carbon density ($Mg\ C\ ha^{-1}$) ‡**			**Peat thickness (m) ¶**			**Peat carbon density ($Mg\ C\ ha^{-1}$) $**		
	Mean ± s.d.	**Median**	**Max**	**Mean ± s.d.**	**Median**	**Min**	**Max**	**Mean ± s.d.**	**Median**	**Min**	**Max**	**Mean ± s.d.**	**Median**	**Max**	**Mean ± s.d.**	**Median**	**Max**	**Mean ± s.d.**	**Median**	**Max**
Interfluvial basin peatlands (ROC)	2.4 (1.5)	2.1	5.9	0.19 (0.06)	0.19	0.10	0.31	56.2 (2.7)	56.5	49.6	61.8	1,619 (810)	1,640	3,183	1.7 (0.9)	1.3	5.4	1,653 (687)	1,402	3,852
River-influenced peatlands (DRC)	2.4 (1.6)	2.0	6.4	0.15 (0.07)	0.15	0.02	0.33	55.0 (3.6)	55.8	42.0	59.2	1,883 (1,511)	1,762	5,162	1.8 (0.8)	1.6	5.6	1,740 (604)	1,697	3,970
Central Congo Basin peatlands (ROC + DRC)	2.4 (1.5)	2.0	6.4	0.17 (0.06)	0.17	0.02	0.33	55.7 (3.2)	56.3	42.0	61.8	1,741 (1,186)	1,700	5,162	1.7 (0.9)	1.6	5.6	1,712 (634)	1,661	3,970

** Field measurement statistics include either the Likouala-aux-Herbes and Ubangi River groups of transects only ('Interfluvial basin peatlands'), or the Congo and Ruki River groups of transects only ('River-influenced peatlands'), or all groups ('Central Congo Basin peatlands').*

† Spatial model statistics include all 50 m resolution pixels mapped in either Republic of the Congo only (ROC), Democratic Republic of the Congo only (DRC), or both countries (ROC + DRC).

In situ *measurements (laboratory and corrected pole-methods) from 213, 238 and 451 locations in ROC (ref. [9]), DRC (this study) and combined, respectively. Peat is ≥ 0.3 m thickness and ≥ 65% organic matter.*

§ n = 43, 37, and 80 well-sampled cores in ROC (ref. [9]), DRC (this study) and combined, respectively, based on 0.1-m thick samples.

‡ n = 43, 37, and 80 well-sampled and interpolated cores in ROC (ref. [9]), DRC (this study) and combined, respectively, based on 0.1-m thick samples.

¶ Median estimate from 100 thickness estimates per 50 m resolution pixel across the median extent map, with thickness estimated from 100 RF regression models trained with four predictor variables, each with a randomly selected Maximum Likelihood peat probability threshold to derive distance from the peatland margin.

$ Median estimate from 2,000 carbon density estimates per 50 m resolution pixel across the median peat area map, with carbon density estimates derived from 20 normally distributed thickness-carbon regressions (Extended Data Figure 7) applied to 100 peat thickness estimates.

FIGURE LEGENDS/CAPTIONS

Figure 1: Maps of field sampling locations (a), peat swamp forest predictions from this study (b), and a comparison of our predictions with a previous map[9] (c). **a**, Points indicate transects, coloured by region. The Congo and Ruki River regional groups appear to be in largely river-influenced peatlands, predominating in DRC, sampled for this study. The Likouala-aux-Herbes and Ubangi River regional groups are in largely rain-fed interfluvial basins, predominating in ROC, from Ref. [9]. The base map, in green, shows the first-generation peat swamp forest map[9]. Inset: Location of central Congo Basin peatlands. **b**, Predicted landcover classes across the central Congo Basin as the most likely class per pixel (>50%), using a legend identical to Ref. [9] to facilitate comparison. **c**, Peat swamp forest predictions from this study and Ref. [9] using the most likely class per pixel. White indicates peat in both studies; red indicates peat in this study only; blue indicates peat only in Ref. [9]. Open water is dark grey. In all panels, national boundaries are black lines; sub-national boundaries are grey lines; non-peat forming forest includes both *terra firme* and non-peat forming seasonally inundated forests.

Figure 2: Maps of peat thickness and uncertainty across the central Congo Basin. **a**, Median prediction of peat thickness (m) from 100 Random Forest regression models with four predictors: distance from the peatland margin, precipitation seasonality, climatic water balance, and distance from the nearest drainage point. **b**, Relative uncertainty (%) of the peat thickness estimate, expressed as ± half the width of the 95% confidence interval as percentage of the median. Black lines represent national boundaries; grey lines represent sub-national administrative boundaries.

Figure 3: Maps of belowground peat carbon density and uncertainty across the central Congo Basin. a, Median prediction of belowground peat carbon density (Mg C ha^{-1}), obtained from applying 20 normally distributed thickness-carbon density regressions (Extended Data Figure 7) to 100 peat thickness estimates (Figure 2a), generating 2,000 carbon density estimates. **b**, Relative uncertainty (%) of the carbon density estimate, expressed as ± half the width of the 95% confidence interval as percentage of the median. Black lines represent national boundaries; grey lines represent sub-national administrative boundaries.

REFERENCES

1. Xu, J., Morris, P. J., Liu, J. & Holden, J. PEATMAP: Refining estimates of global peatland distribution based on a meta-analysis. *Catena* **160**, 134–140 (2018).

2. Yu, Z., Loisel, J., Brosseau, D. P., Beilman, D. W. & Hunt, S. J. Global peatland dynamics since the Last Glacial Maximum. *Geophys. Res. Lett.* **37**, 1–5 (2010).

3. Leifeld, J. & Menichetti, L. The underappreciated potential of peatlands in global climate change mitigation strategies. *Nat. Commun.* **9**, 1–7 (2018).

4. Scharlemann, J. P. W., Tanner, E. V. J., Hiederer, R. & Kapos, V. Global soil carbon: understanding and managing the largest terrestrial carbon pool. *Carbon Manag.* **5**, 81–91 (2014).

5. Page, S. E., Rieley, J. O. & Banks, C. J. Global and regional importance of the tropical peatland carbon pool. *Glob. Chang. Biol.* **17**, 798–818 (2011).

6. Ribeiro, K. *et al.* Tropical peatlands and their contribution to the global carbon cycle and climate change. *Glob. Chang. Biol.* **00**, 1–17 (2020).

7. Leifeld, J., Wüst-Galley, C. & Page, S. Intact and managed peatland soils as a source and sink of GHGs from 1850 to 2100. *Nat. Clim. Chang.* **9**, 945–947 (2019).

8. Page, S. E. *et al.* The amount of carbon released from peat and forest fires in Indonesia during 1997. *Nature* **420**, 61–65 (2002).

9. Dargie, G. C. *et al.* Age, extent and carbon storage of the central Congo Basin peatland complex. *Nature* **542**, 86–90 (2017).

10. Alsdorf, D. *et al.* Opportunities for hydrologic research in the Congo Basin. *Rev. Geophys.* **54**, 378–409 (2016).

11. Kiahtipes, C. A. & Schefuß, E. Congo Basin peatlands as a baseline record for past hydrology and climate [W-68]. *Earth Sp. Sci. Open Arch.* (2019) doi:https://www.essoar.org/doi/10.1002/essoar.10500726.1.

12. Davenport, I. J. *et al.* First Evidence of Peat Domes in the Congo Basin using LiDAR from a Fixed-Wing Drone. *Remote Sens.* **12**, 1–13 (2020).

13. Lee, H. *et al.* Characterization of terrestrial water dynamics in the Congo Basin using GRACE and satellite radar altimetry. *Remote Sens. Environ.* **115**, 3530–3538 (2011).

14. Rosenqvist, A. Mapping of seasonal inundation in the Congo River basin - Prototype study using ALOS PALSAR. *Proc. 33rd Int. Symp. Remote Sens. Environ.* **ISRSE 33**, 709–712 (2009).

15. Lee, H., Yuan, T., Jung, H. C. & Beighley, E. Mapping wetland water depths over the central Congo Basin using PALSAR ScanSAR, Envisat altimetry, and MODIS VCF data. *Remote Sens. Environ.* **159**, 70–79 (2015).

16. Chicco, D., Tötsch, N. & Jurman, G. The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation. *BioData Min.* **14**, 1–22 (2021).

17. Gumbricht, T. *et al.* An expert system model for mapping tropical wetlands and peatlands reveals South America as the largest contributor. *Glob. Chang. Biol.* **23**, 3581–3599 (2017).

18. Young, D. M., Parry, L. E., Lee, D. & Ray, S. Spatial models with covariates improve estimates of peat depth in blanket peatlands. *PLoS One* **13**, 1–19 (2018).

19. Rudiyanto *et al.* Digital mapping for cost-effective and accurate prediction of the depth and carbon stocks in Indonesian peatlands. *Geoderma* **272**, 20–31 (2016).

20. Malhi, Y. & Wright, J. Spatial patterns and recent trends in the climate of tropical rainforest regions. *Philos. Trans. R. Soc. B Biol. Sci.* **359**, 311–329 (2004).

21. Lewis, S. L. *et al.* Above-ground biomass and structure of 260 African tropical forests. *Philos. Trans. R. Soc. B Biol. Sci.* **368**, 1–14 (2013).

22. Hastie, A. *et al.* Risks to carbon storage from land-use change revealed by peat thickness maps of Peru. *Nat. Geosci.* (2022) doi:https://doi.org/10.1038/s41561-022-00923-4.

23. Verhegghen, A., Mayaux, P., De Wasseige, C. & Defourny, P. Mapping Congo Basin vegetation types from 300 m and 1 km multi-sensor time series for carbon stocks and forest areas estimation. *Biogeosciences* **9**, 5061–5079 (2012).

24. Miles, L. *et al.* Carbon, biodiversity and land-use in the Central Congo Basin Peatlands. *UN Environment Programme* (2017).

25. Vancutsem, C. *et al.* Long-term (1990–2019) monitoring of forest cover changes in the humid tropics. *Sci. Adv.* **7**, 1–21 (2021).

26. Dargie, G. C. *et al.* Congo Basin peatlands: threats and conservation priorities.

Mitig. Adapt. Strateg. Glob. Chang. **24**, 669–686 (2018).

27. Maisels, F. *et al.* Devastating Decline of Forest Elephants in Central Africa. *PLoS One* **8**, 1–13 (2013).

28. Strindberg, S. *et al.* Guns, germs, and trees determine density and distribution of gorillas and chimpanzees in Western Equatorial Africa. *Sci. Adv.* **4**, 1–14 (2018).

29. Lawson, I. T. *et al.* Improving estimates of tropical peatland area, carbon storage, and greenhouse gas fluxes. *Wetl. Ecol. Manag.* **23**, 327–346 (2015).

30. Seidensticker, D. *et al.* Population collapse in Congo rainforest from 400 CE urges reassessment of the Bantu Expansion. *Sci. Adv.* **7**, 1–13 (2021).

31. Seidensticker, D. dirkseidensticker/HumActCentralAfrica_Paper: Codebase (Version v1.0). *Zenodo* (2020) doi:10.5281/ZENODO.4394894.

32. Hubau, W. *et al.* Asynchronous carbon sink saturation in African and Amazonian tropical forests. *Nature* **579**, 80–87 (2020).

33. Lopez-Gonzalez, G., Lewis, S. L., Burkitt, M., Baker, T. R. & Phillips, O. L. ForestPlots.net Database. www.forestplots.net (2009).

34. Lopez-Gonzalez, G., Lewis, S. L., Burkitt, M. & Phillips, O. L. ForestPlots.net: A web application and research tool to manage and analyse tropical forest plot data. *J. Veg. Sci.* **22**, 610–613 (2011).

35. Batumike, R., Imani, G., Urom, C. & Cuni-Sanchez, A. Bushmeat hunting around Lomami National Park, Democratic Republic of the Congo. *Oryx* **55**, 1–11 (2020).

36. Clubb, F. J. *et al.* Geomorphometric delineation of floodplains and terraces from objectively defined topographic thresholds. *Earth Surf. Dyn.* **5**, 369–385 (2017).

37. Lehner, B., Verdin, K. & Jarvis, A. New Global Hydrography Derived From Spaceborne Elevation Data. *Eos, Trans. Am. Geophys. Union* **89**, 93–94 (2008).

38. NASA Jet Propulsion Laboratory. NASADEM Merged DEM Global 1 arc second V001. (2020) doi:https://doi.org/10.5067/MEaSUREs/NASADEM/NASADEM_HGT.001.

39. Yamazaki, D. *et al.* A high-accuracy map of global terrain elevations. *Geophys. Res. Lett.* **44**, 5844–5853 (2017).

40. Yamazaki, D. *et al.* MERIT Hydro : A High-Resolution Global Hydrography Map Based on Latest Topography Dataset. *Water Resour. Res.* **55**, 5053–5073 (2019).

41. Observatoire Satellital des Forêts d'Afrique Centrale (OSFAC). Forêts d'Afrique Centrale Evaluées par Télédétection. https://osfac.net/data-products/facet/ (2014).

42. Ploton, P. *et al.* Spatial validation reveals poor predictive performance of large-scale ecological mapping models. *Nat. Commun.* **11**, 1–11 (2020).

43. Roberts, D. R. *et al.* Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography (Cop.).* **40**, 913–929 (2017).

44. Meyer, H., Reudenbach, C., Wöllauer, S. & Nauss, T. Importance of spatial predictor variable selection in machine learning applications – Moving from

data reproduction to spatial prediction. *Ecol. Modell.* **411**, 11 (2019).

45. Chicco, D. & Jurman, G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics* **21**, 1–13 (2020).

46. Powers, D. M. W. Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness and Correlation. *J. Mach. Learn. Technol.* **2**, 37–63 (2011).

47. Gorelick, N. *et al.* Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **202**, 18–27 (2017).

48. Fick, S. E. & Hijmans, R. J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315 (2017).

49. Trabucco, A. & Zomer, R. J. Global Aridity Index and Potential Evapotranspiration (ET0) Climate Database v2. *figshare* (2019) doi:https://doi.org/10.6084/m9.figshare.7504448.v3.

50. Baccini, A. *et al.* Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nat. Clim. Chang.* **2**, 182–185 (2012).

51. Protected Planet: The World Database on Protected Areas (WDPA) and World Database on Other Effective Area-based Conservation Measures (WD-OECM). *UNEP-WCMC/IUCN* https://www.protectedplanet.net (2021).

52. Republic of the Congo logging concessions. *Global Forest Watch* https://data.globalforestwatch.org/datasets/gfw::republic-of-the-congo-logging-concessions/ (2019).

53. Democratic Republic of the Congo forest titles. *Global Forest Watch* https://data.globalforestwatch.org/datasets/535eb1335c4841b0bff272b78e2cc2f4_6 (2019).

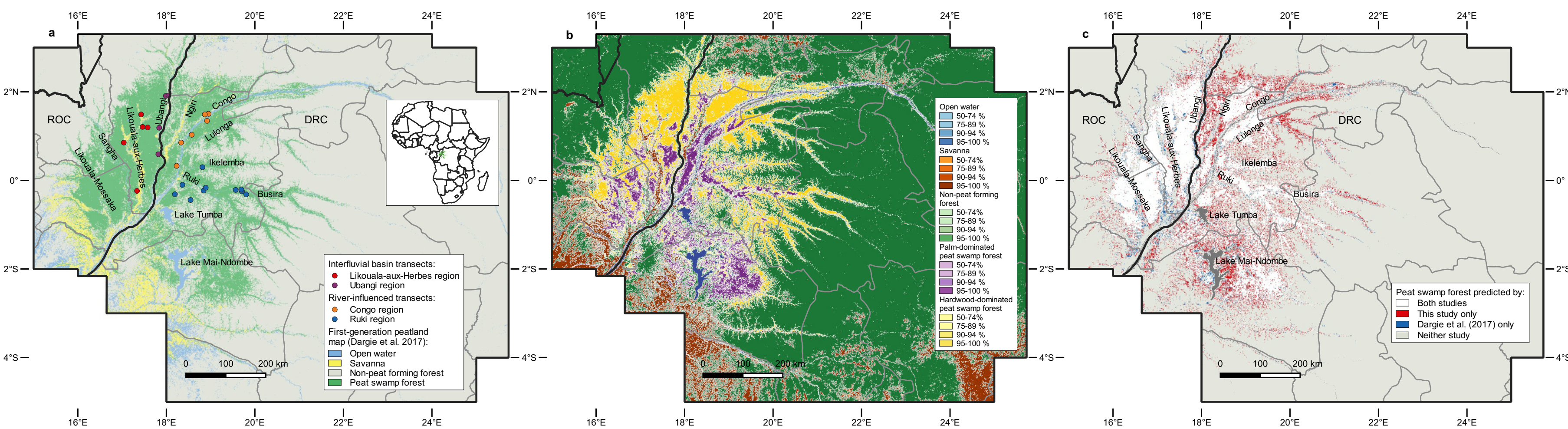
54. Republic of the Congo mining permits. *Global Forest Watch* https://data.globalforestwatch.org/datasets/84fbbcc10c9f47f890750dd42426cbd2_18/ (2019).

55. Democratic Republic of the Congo mining permits. *Global Forest Watch* https://data.globalforestwatch.org/datasets/3b4c0c91306c47abaec0c3fd46088242_5/ (2019).

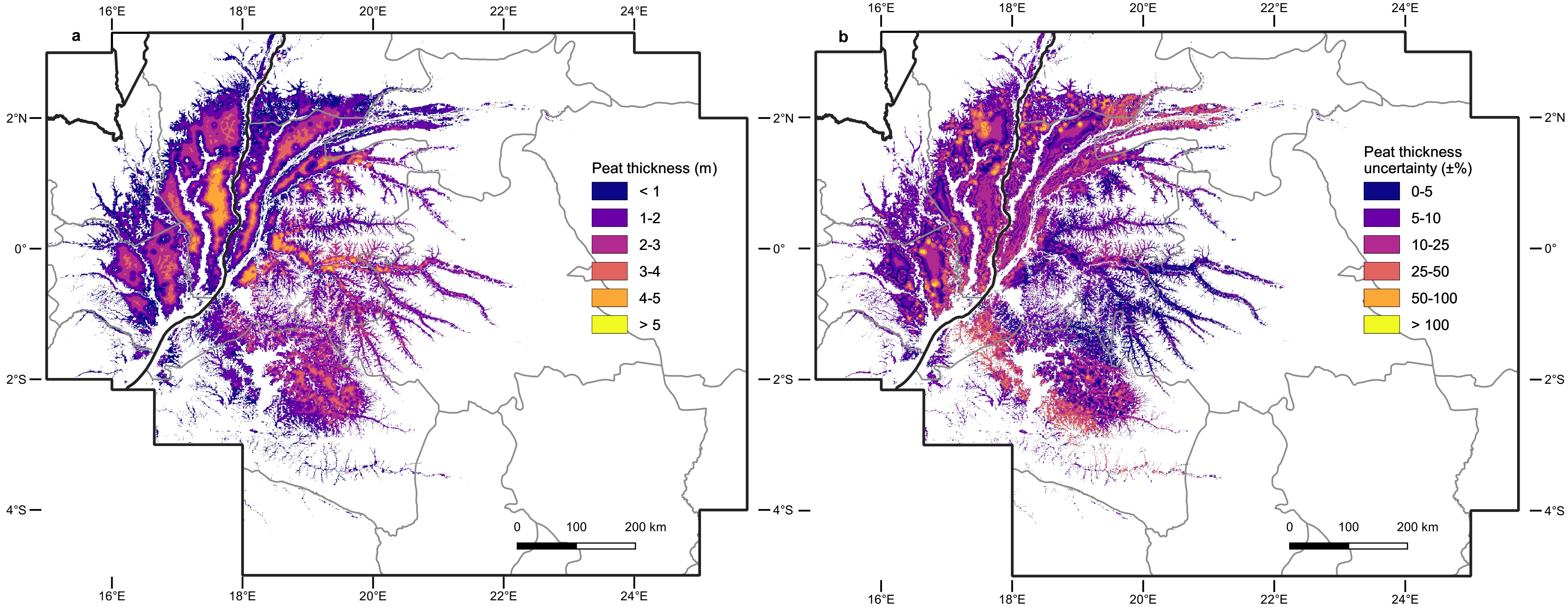
56. DRC Agriculture Plantations. *MapforEnvironment* https://mapforenvironment.org/layer/info/80/#5.24/-1.263/19.467 (2014).

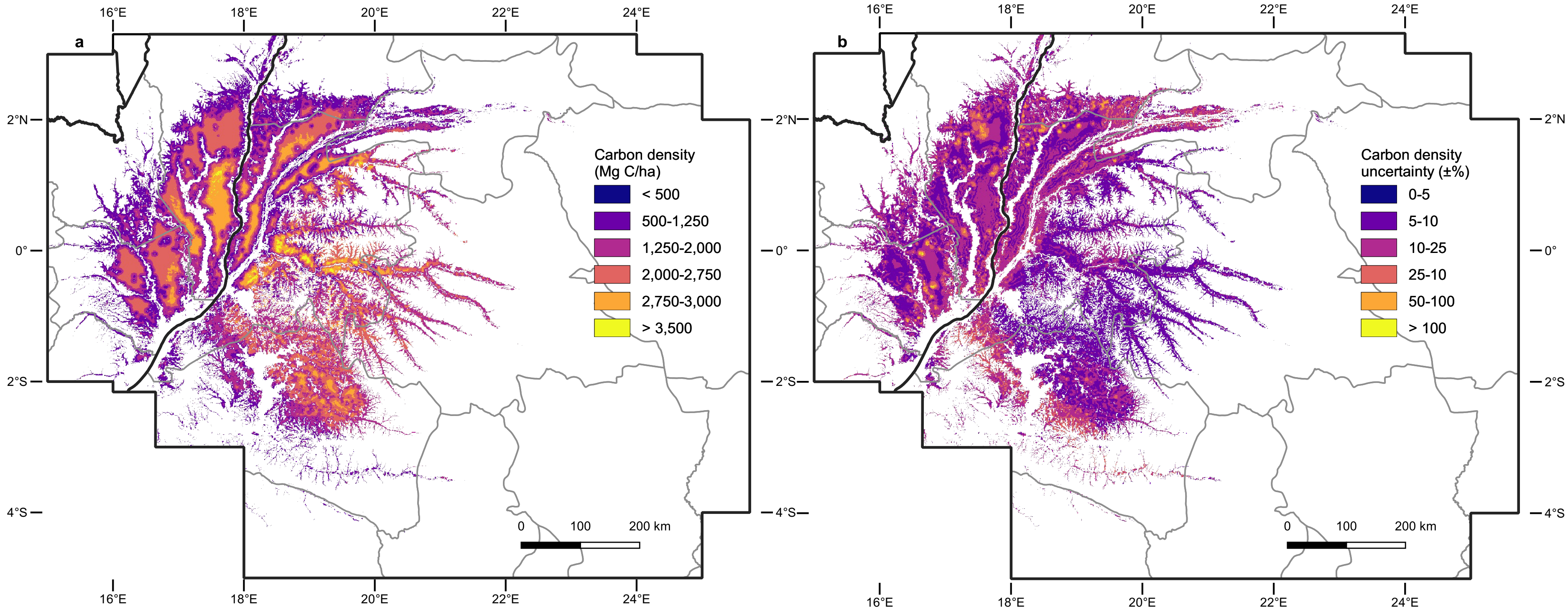
57. Republic of the Congo oil palm concessions. *Global Forest Watch* https://data.globalforestwatch.org/datasets/f1fb5773903244abbe8282cae189863e_17/ (2019).

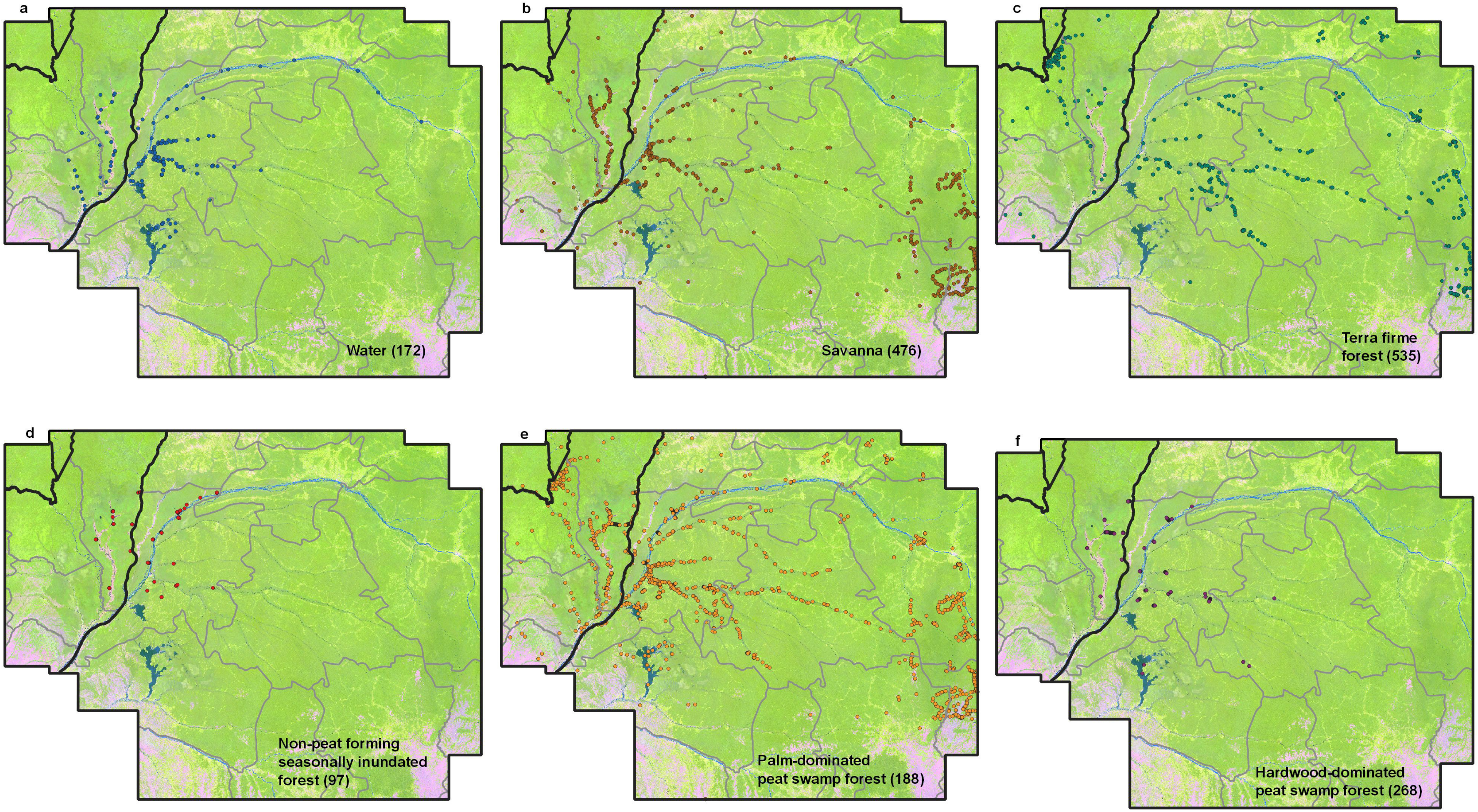
58. The Coming Storm: How Secrecy and Collusion in Industrial Agriculture Spell Disaster for the Congo Basin's Forests. *Earthsight* https://www.earthsight.org.uk/news/investigations/the-coming-storm (2018).



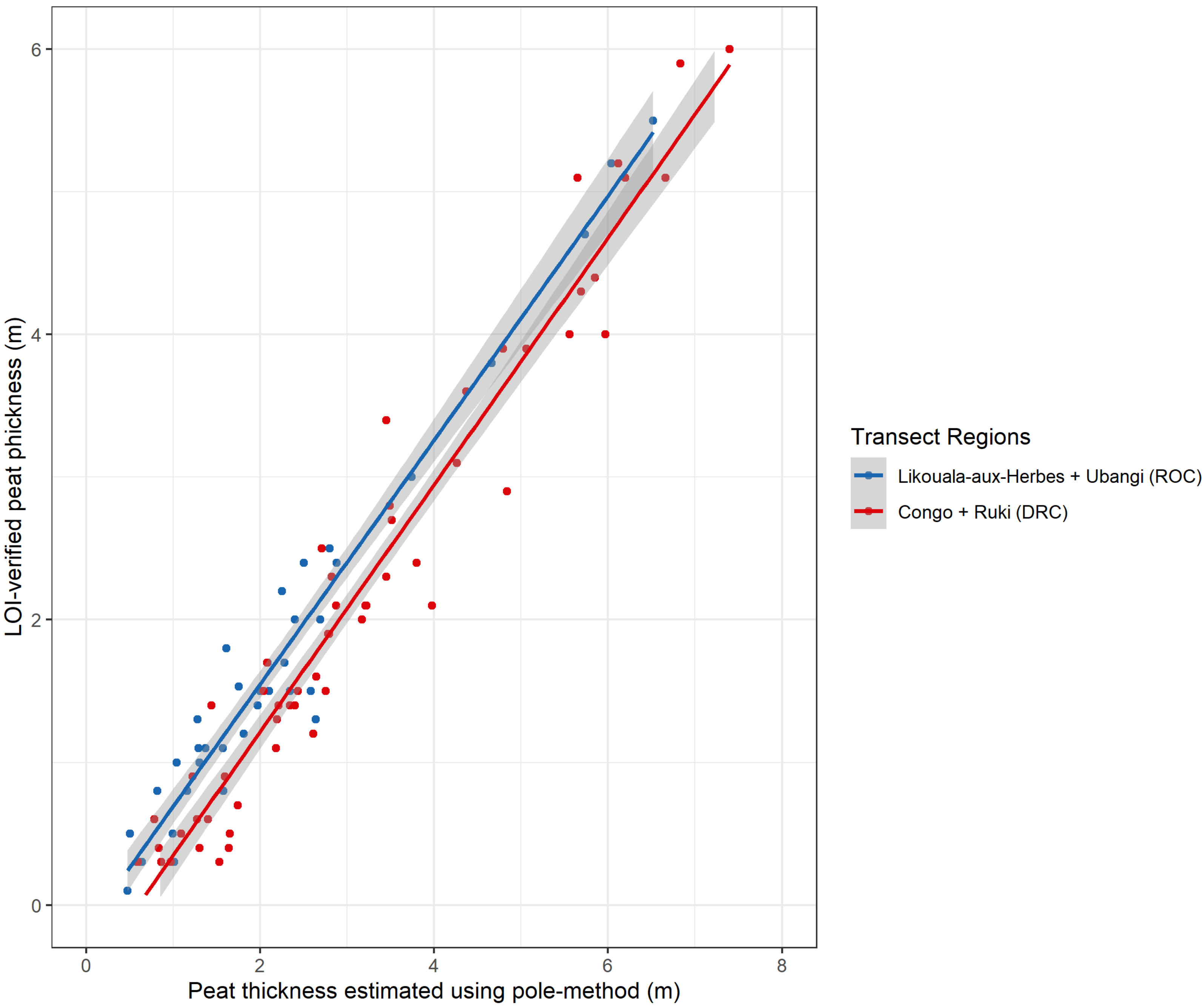
a
b
c
16°E
18°E
20°E
22°E
24°E
2°N
0°
2°S
4°S
ROC
DRC
Sangha
Likouala-Mossaka
Likouala-aux-Herbes
Ubangi
Ngiri
Congo
Lulonga
Ikelemba
Ruki
Busira
Lake Tumba
Lake Mai-Ndombe
Interfluvial basin transects:
Likouala-aux-Herbes region
Ubangi region
River-influenced transects:
Congo region
Ruki region
First-generation peatland
map (Dargie et al. 2017):
Open water
Savanna
Non-peat forming forest
Peat swamp forest
0 100 200 km
Open water
50-74 %
75-89 %
90-94 %
95-100 %
Savanna
50-74%
75-89 %
90-94 %
95-100 %
Non-peat forming
forest
50-74%
75-89 %
90-94 %
95-100 %
Palm-dominated
peat swamp forest
50-74%
75-89 %
90-94 %
95-100 %
Hardwood-dominated
peat swamp forest
50-74%
75-89 %
90-94 %
95-100 %
Peat swamp forest predicted by:
Both studies
This study only
Dargie et al. (2017) only
Neither study


a
b
Peat thickness (m)
< 1
1-2
2-3
3-4
4-5
> 5
Peat thickness uncertainty (±%)
0-5
5-10
10-25
25-50
50-100
> 100
0
100
200 km
16°E
18°E
20°E
22°E
24°E
2°N
0°
2°S
4°S

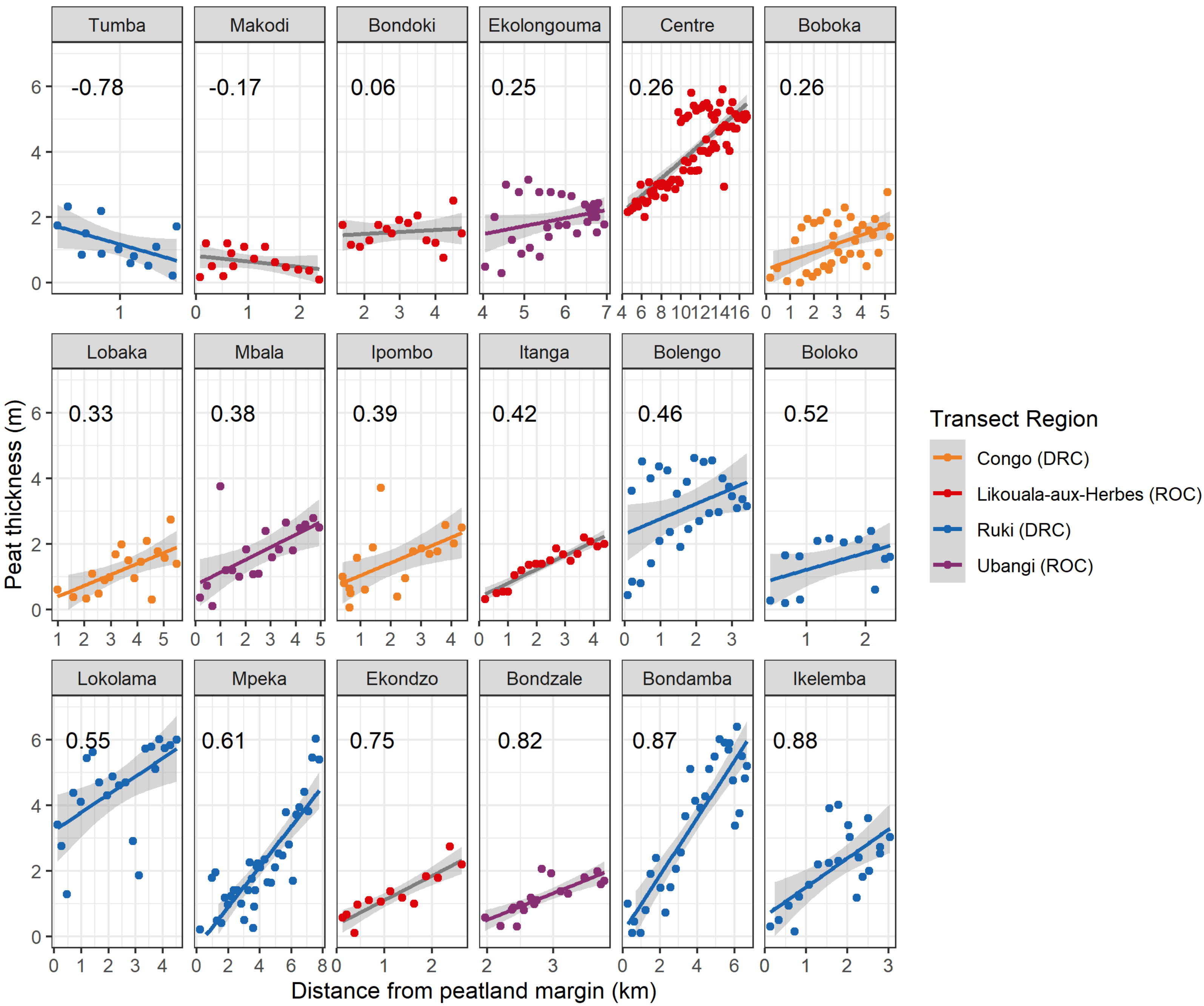

a
16°E
18°E
20°E
22°E
24°E
2°N
0°
2°S
4°S
Carbon density
(Mg C/ha)
< 500
500-1,250
1,250-2,000
2,000-2,750
2,750-3,000
> 3,500
0
100
200 km
b
Carbon density
uncertainty (±%)
0-5
5-10
10-25
25-10
50-100
> 100


a
Water (172)
b
Savanna (476)
c
Terra firme
forest (535)
d
Non-peat forming
seasonally inundated
forest (97)
e
Palm-dominated
peat swamp forest (188)
f
Hardwood-dominated
peat swamp forest (268)

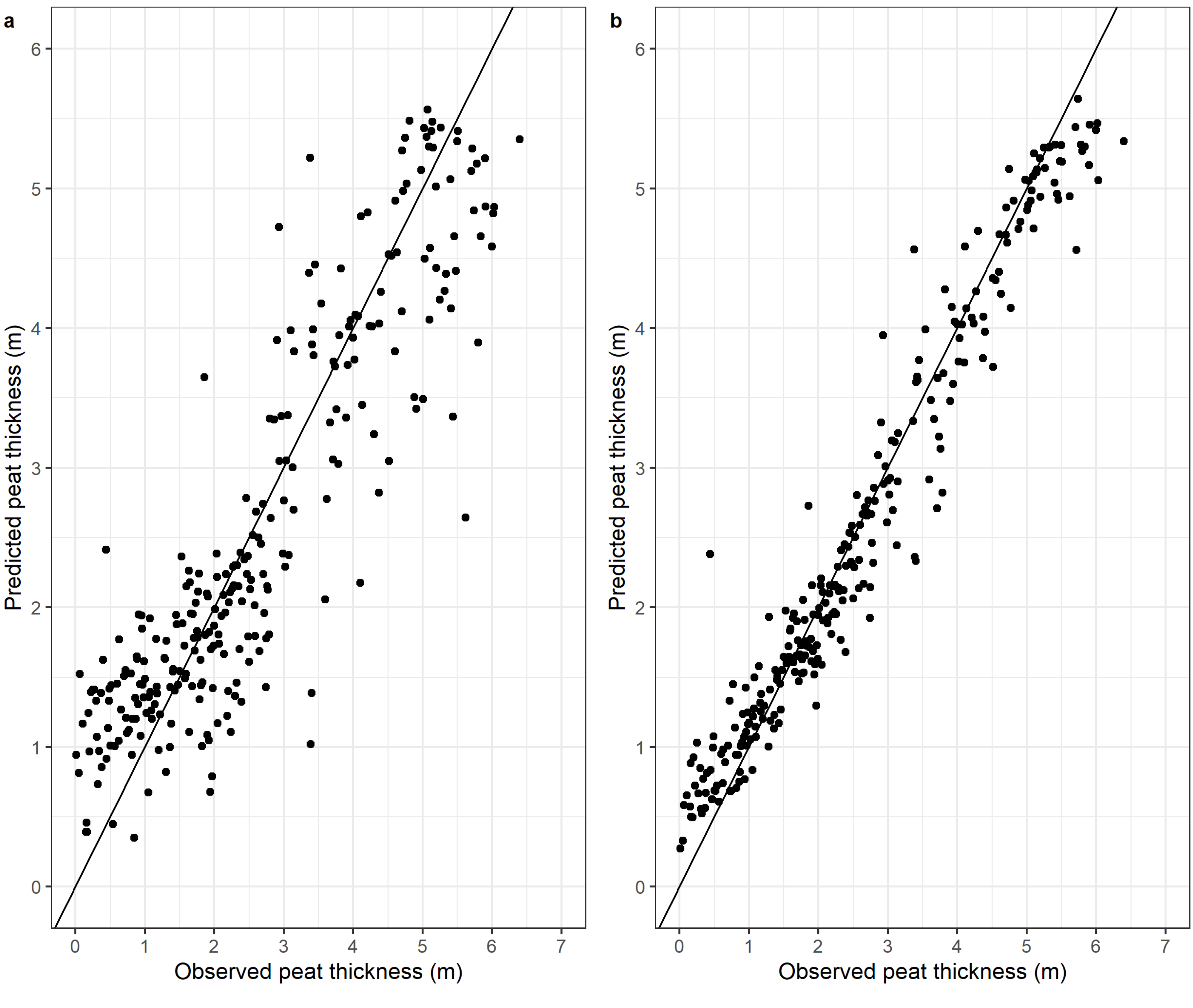
Country	Region	Peatland area (km^2)	Peat thickness (m)	Peat carbon density (Mg C ha^{-1})	Peat carbon stock (Pg C)
Republic of the Congo (ROC)	Likouala	28,636	1.9 ± 1.0	1,815 ± 740	5.4 (4.8 - 5.8)
	Cuvette	17,757	1.6 ± 0.8	1,626 ± 624	2.9 (2.7 - 3.2)
	Sangha	7,465	1.1 ± 0.4	1,218 ± 325	0.9 (0.8 - 1.0)
	Plateaux	1,183	0.9 ± 0.1	1,059 ± 162	0.1 (0.1 - 0.1)
	Total ROC	55,072	1.7 ± 0.9	1,653 ± 687	9.3 (8.4 - 10.2)
Democratic Republic of the Congo (DRC)	Équateur	58,276	1.9 ± 0.9	1,822 ± 658	10.7 (9.9 - 11.7)
	Mai-Ndombe	29,825	1.8 ± 0.7	1,752 ± 548	5.2 (4.8 - 5.7)
	Tshuapa	11,628	1.9 ± 0.5	1,917 ± 343	2.1 (1.8 - 2.6)
	Sud-Ubangi	7,557	1.1 ± 0.4	1,243 ± 370	1.0 (0.8 - 1.2)
	Mongala	5,329	1.2 ± 0.4	1,259 ± 360	0.6 (0.5 - 0.8)
	Total DRC	113,201	1.8 ± 0.8	1,740 ± 604	19.6 (17.9 - 21.9)
ROC and DRC combined	Total central Congo Basin peatlands	167,648 (159,378 - 175,079)	1.7 ± 0.9	1,712 ± 634	29.0 (26.3 - 32.2)

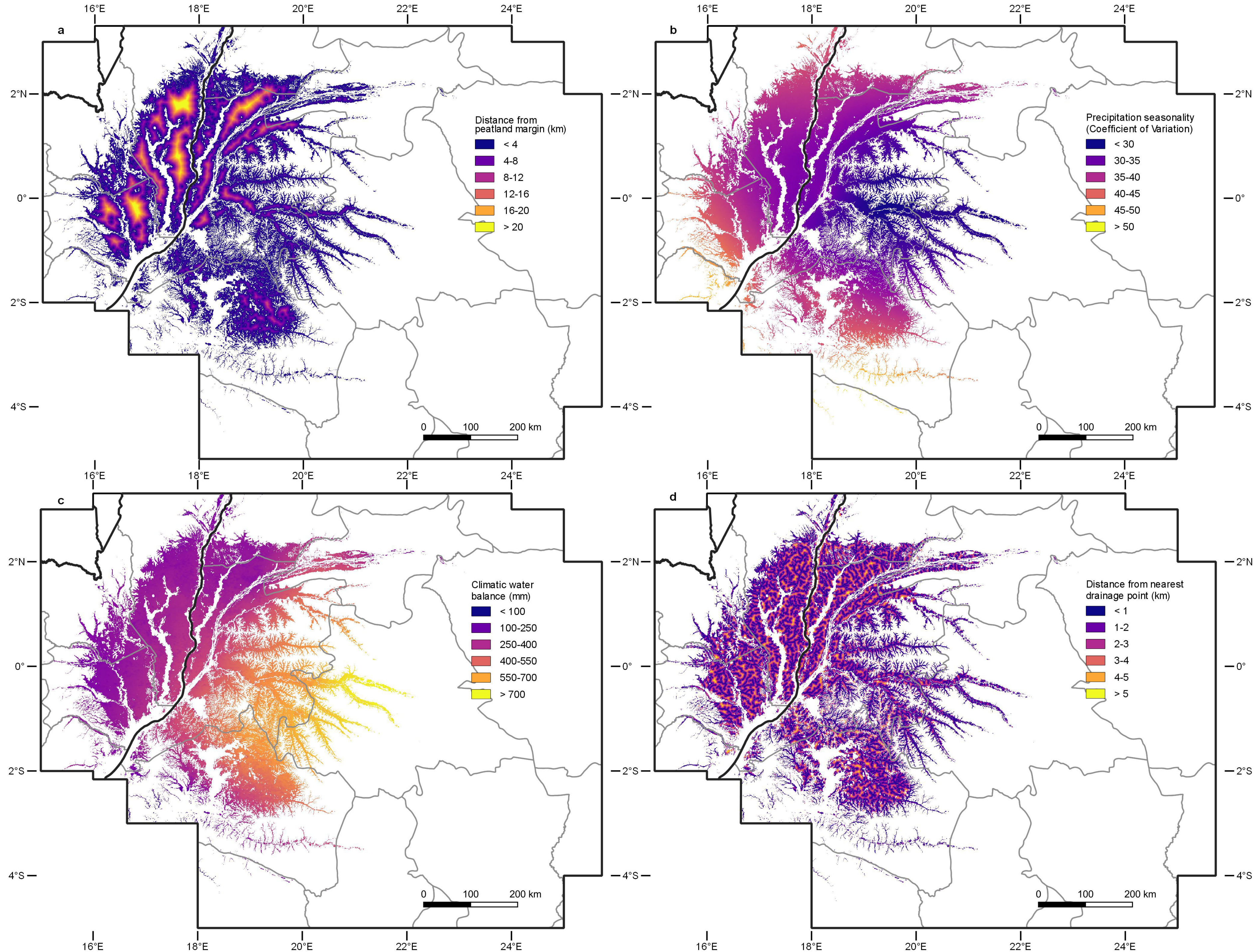


LOI-verified peat phickness (m)
Peat thickness estimated using pole-method (m)
0
2
4
6
8
Transect Regions
Likouala-aux-Herbes + Ubangi (ROC)
Congo + Ruki (DRC)

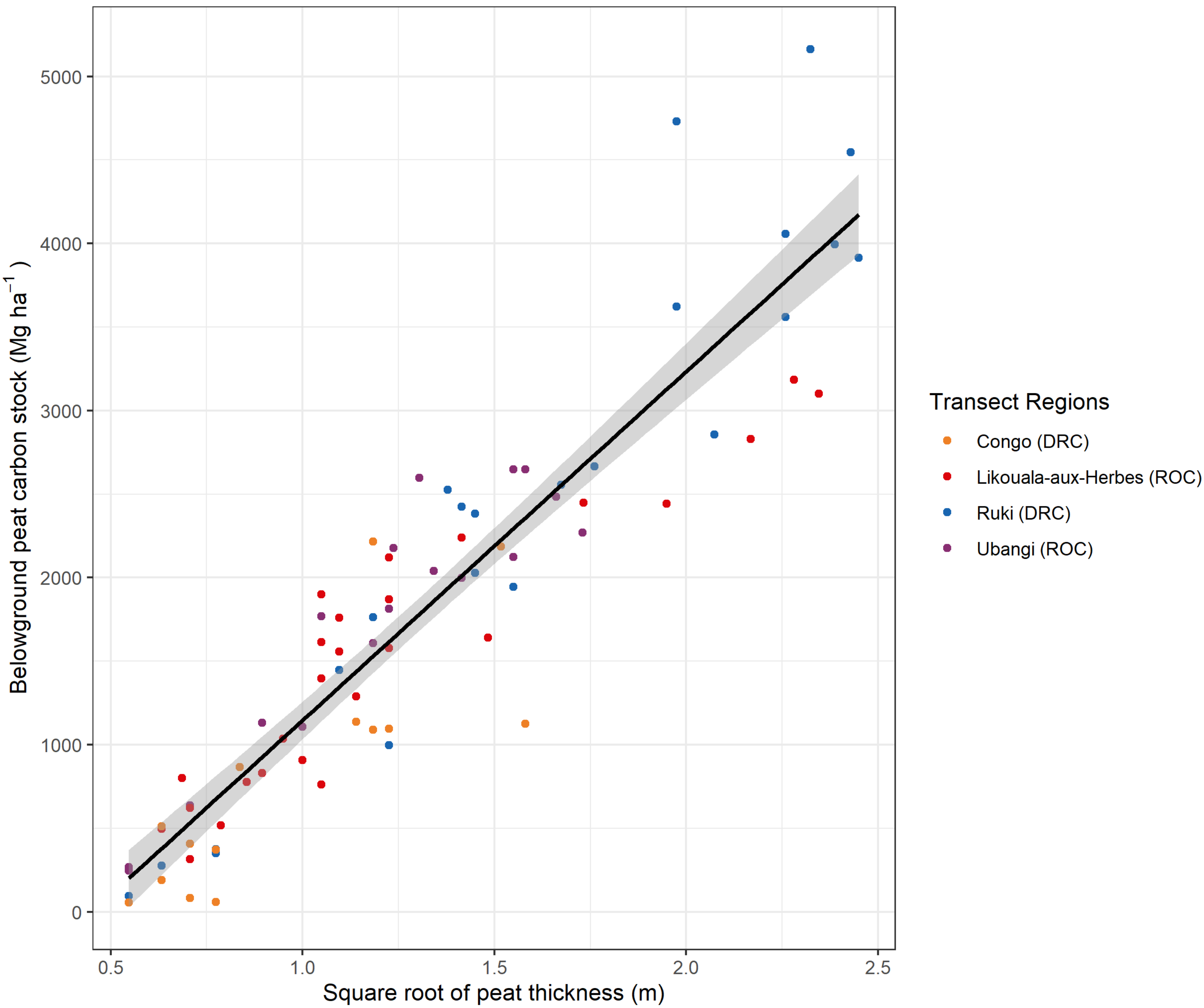


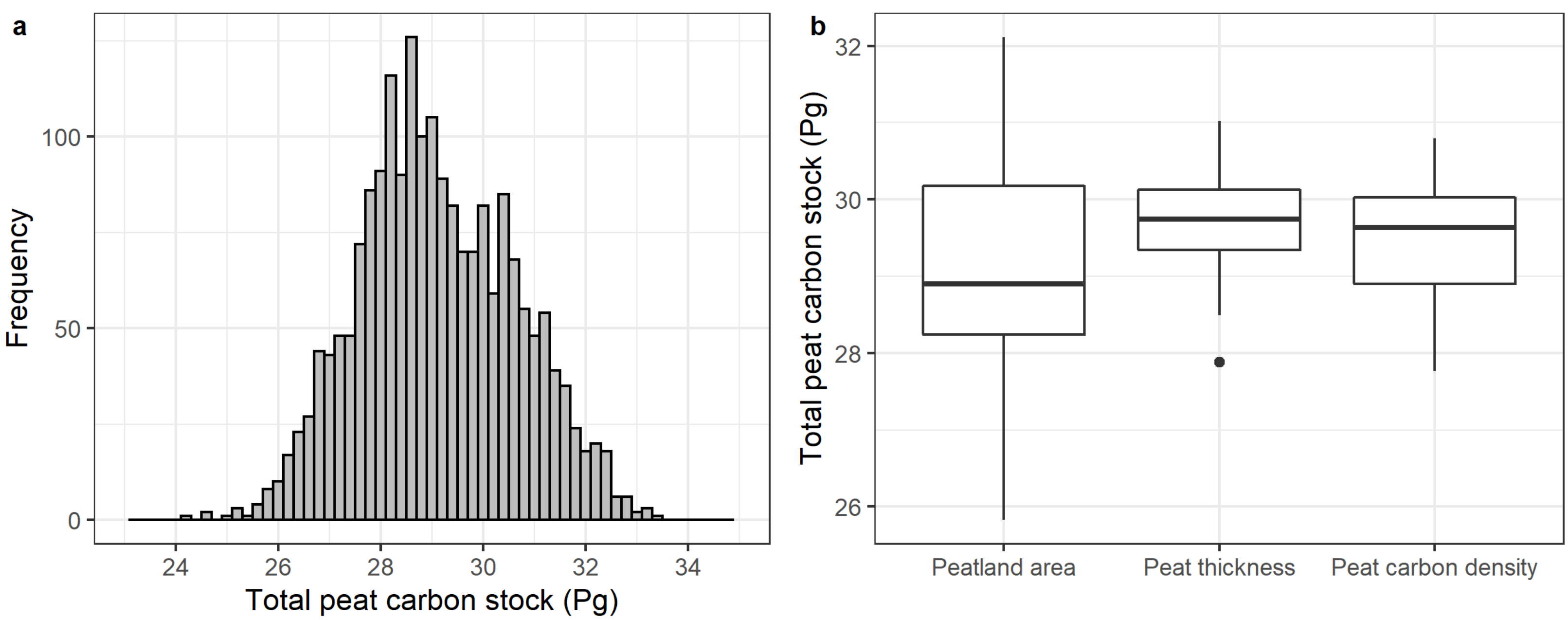
Tumba
Makodi
Bondoki
Ekolongouma
Centre
Boboka
-0.78
-0.17
0.06
0.25
0.26
0.26
Lobaka
Mbala
Ipombo
Itanga
Bolengo
Boloko
0.33
0.38
0.39
0.42
0.46
0.52
Lokolama
Mpeka
Ekondzo
Bondzale
Bondamba
Ikelemba
0.55
0.61
0.75
0.82
0.87
0.88
Peat thickness (m)
Distance from peatland margin (km)
Transect Region
Congo (DRC)
Likouala-aux-Herbes (ROC)
Ruki (DRC)
Ubangi (ROC)

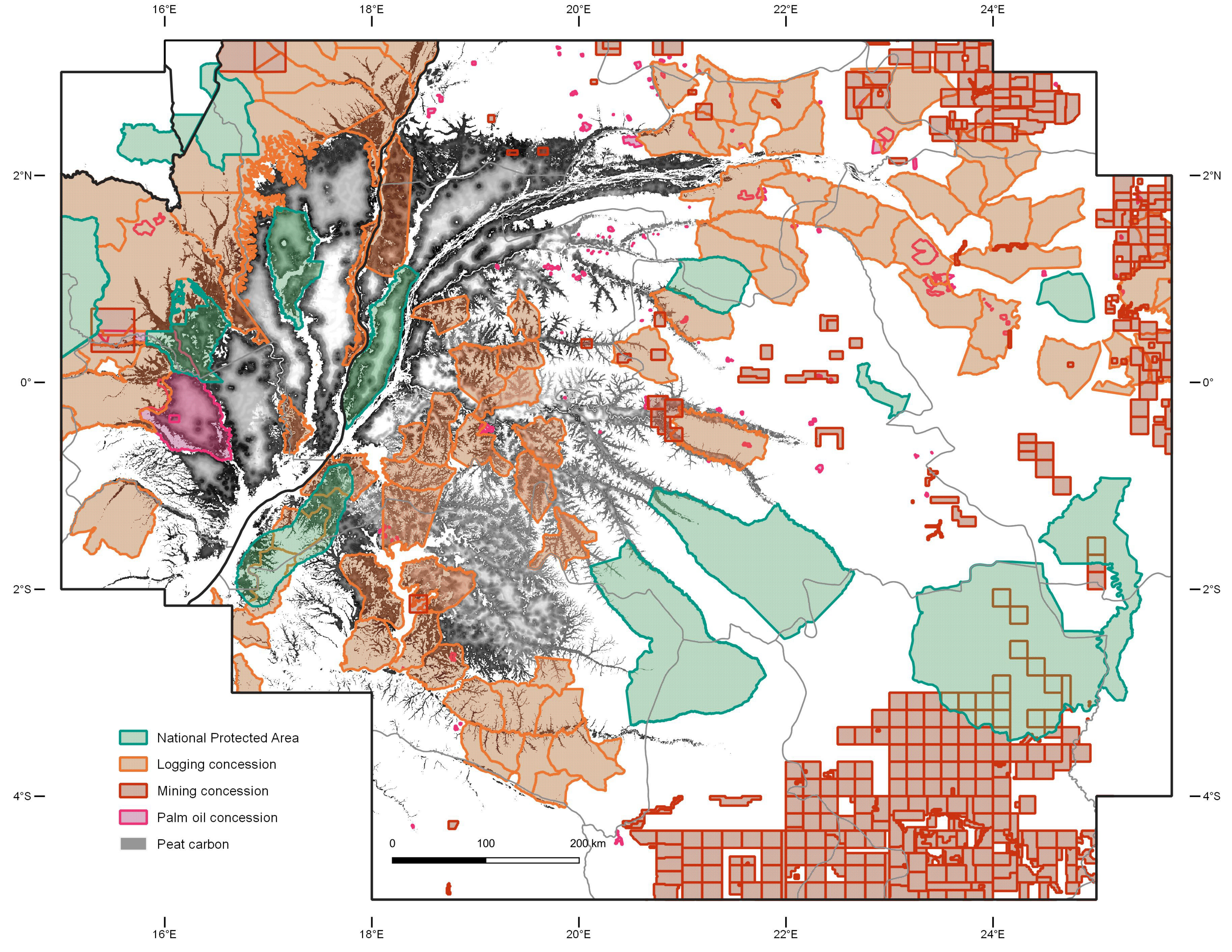

a
b
Predicted peat thickness (m)
Observed peat thickness (m)
0
1
2
3
4
5
6
7



a
Distance from peatland margin (km)
< 4
4-8
8-12
12-16
16-20
> 20
b
Precipitation seasonality (Coefficient of Variation)
< 30
30-35
35-40
40-45
45-50
> 50
c
Climatic water balance (mm)
< 100
100-250
250-400
400-550
550-700
> 700
d
Distance from nearest drainage point (km)
< 1
1-2
2-3
3-4
4-5
> 5
0
100
200 km
16°E
18°E
20°E
22°E
24°E
2°N
0°
2°S
4°S


Belowground peat carbon stock (Mg ha^{-1})
5000
4000
3000
2000
1000
0
0.5
1.0
1.5
2.0
2.5
Square root of peat thickness (m)
Transect Regions
Congo (DRC)
Likouala-aux-Herbes (ROC)
Ruki (DRC)
Ubangi (ROC)


a
b
Frequency
100
50
0
24
26
28
30
32
34
Total peat carbon stock (Pg)
32
30
28
26
Total peat carbon stock (Pg)
Peatland area
Peat thickness
Peat carbon density



16°E
18°E
20°E
22°E
24°E
2°N
0°
2°S
4°S
National Protected Area
Logging concession
Mining concession
Palm oil concession
Peat carbon
0
100
200 km

Country	Concessions / Protected areas	Peatland area (km^2)	Peat thickness (m)	Peat carbon density ($Mg\ C\ ha^{-1}$)	Peat carbon stock (Pg C)
Republic of the Congo (ROC)	Industrial logging / mining / palm oil concessions	13,539 (25%)	1.2 ± 0.6	1,299 ± 451	2.0 (22%)
	National-level protected areas	6,402 (12%)	1.4 ± 0.6	1,463 ± 478	1.0 (11%)
Democratic Republic of the Congo (DRC)	Industrial logging / mining / palm oil concessions	29,712 (26%)	1.6 ± 0.7	1,671 ± 567	5.4 (28%)
	National-level protected areas	8,105 (7%)	1.5 ± 0.8	1,552 ± 592	1.4 (7%)
ROC and DRC combined	Industrial logging / mining / palm oil concessions	43,250 (26%)	1.5 ± 0.7	1,551 ± 560	7.4 (26%)
	National-level protected areas	14,511 (9%)	1.5 ± 0.7	1,513 ± 547	2.4 (8%)