Supplementary Information for

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

Supplementary Tables and Figures

Supplementary Table 1 | Overview of 18 field transects in the Democratic Republic of the Congo.

Transect name	Transect purpose	Regional group (spatial cross-validation)	Length (km)	Transect justification	Hypothesis / Peat prediction	Basic transect field description	% correctly predicted by Dargie et al.[9]
Bolengo	Hypothesis testing	Ruki River	8.0	Transect crosses through seemingly higher elevation on the left-bank of the Busira River. Traverses mostly predicted hardwood swamp, perpendicular to the river, before ending in likely upland *terra firme* forest that is predicted as peat by a preliminary ML model based on just HAND (not DEM or DEM+HAND). Potential gradient in peat thickness may be detected moving away from the river.	No peat is expected in areas predicted by the ML model based on HAND only. Effect of a nutrient gradient or inundation levels on peat thickness is expected when moving away from the river.	Largely hardwood-dominated peat swamp, traversing *terra firme* forest from 6.75 km onwards (as predicted by DEM+HAND). Peat thickness gradually increases towards a max. of 4.6 m at 4.75 km.	97.0
Ikelemba	Hypothesis testing	Ruki River	5.0	Transect crosses the floodplain next the Ikelemba River. Traverses both predicted hardwood and palm peat swamp forest before ending in upland *terra firme* forest. Potential effects of nutrient gradients or inundation levels (river-influenced or runoff) on peat thickness are expected.	Effect of a nutrient gradient or inundation levels (incl. upland runoff) on peat thickness to be detected when moving away from the river towards *terra firme*.	Seasonally inundated peat-forming hardwood swamp up to 2.5 km; after that permanently waterlogged hardwood peat swamp. Gradually thicker peat, with max. of 4.0 m at 4.75 km.	95.2
Lokolama	Hypothesis testing	Ruki River	5.0	Transect on edge of a suspected ovoid interfluvial basin next to the Congo River mainstem, which could be domed. However, radar data indicates potentially higher water-table depths than in ROC, which could be river-influenced or due to upland runoff. Traverses both predicted hardwood and palm-dominated peat swamp forests towards interior.	Thick peat deposits expected towards interior of suspected peat dome. No likely effect of inundation levels on peat thickness expected.	Mix of palm- and hardwood peat swamp forest from 500 m onwards. Thick peat deposits that gradually increase towards max. 6.0 m at 4.25 km and 5 km.	91.3
Boloko	Hypothesis testing	Ruki River	4.5	Transect crosses a small, dendritic river valley (Boloko River, tributary of Ruki River), of which many are found in DRC. Part of peatland water likely originates from upland runoff from surrounding *terra firme* forest.	Peat is expected in small dendritic river valleys because of high water levels.	Shallow hardwood peat swamp, traversed by river. Max. peat depth is 2.4 m at 3 km (1 km from river).	86.7
Bondamba	Hypothesis testing	Ruki River	7.0	Concentrating patterns are visible in optical and radar data along the right-bank of the Busira River, indicating potential peat domes. This transect runs from the margin to the centre of a potential dome, mostly traversing likely palm-dominated swamp.	Increasing peat thickness expected towards the interior of suspected peat dome. Effects of nutrient gradient or inundation levels to be detected.	Largely palm-dominated peat swamp, transitioning to hardwood after 7 km. Thick peat deposits present, reaching max. 6.4 m after 6.5 km.	86.2

Tumba	Hypothesis testing	Ruki River	4.0	Transect crosses a likely low-lying floodplain forest between the Bonsole and Boloko Rivers (both tributaries of Ruki), with little elevation change. Traverses predicted palm-dominated peat swamp forest.	No or little peat is expected because of erosion from likely migrating rivers and streams.	Shallow seasonally inundated peat-forming swamp (mixed hardwood and palm vegetation). Max. peat thickness is 2.3 m at 0.75 km.	83.3
Mpeka	Hypothesis testing	Ruki River	10.0	Transect runs perpendicular to the Ruki River towards the interior of a likely low-lying floodplain, with mostly palm swamp predicted. Traverses a river levee, before entering a likely depression. Potential effects of nutrient gradient or inundation levels (river-influenced) on peat thickness.	Effects of a nutrient gradient or inundation levels on peat thickness to be expected when moving away from the river.	Largely palm-dominated peat swamp forest, transitioning to hardwood from 8 km onwards. Peat thickness gradually increases to 6.0 m at 9.75 km. *Terra firme* forest at 1 km because of a levee.	80.5
Boboka	Hypothesis testing	Congo River	11.0	Transect extends into the interior of a likely interfluvial basin towards a lake, across largely hardwood swamp forest. Likely different hydrology than other Congo River transects, due to the presence of a lake and limited inundations at higher elevations.	Thick peat deposits are expected towards the interior. Little effect of river-influenced inundations expected.	Non-peat forming seasonally inundated forest close to the river, transitioning to shallow hardwood-dominated peat swamp after 3 km. Max. peat thickness is 2.8 m after 7.25 km.	88.9
Lobaka	Hypothesis testing	Congo River	6.0	Transect runs perpendicular to the Congo mainstem into the interior of a likely interfluvial basin. Mostly expected palm swamp forest. Potential effect of a nutrient gradient or inundation levels (river-influenced) may be detected moving away from the river.	Thick peat deposits expected towards interior. Effect of nutrient gradient, inundation levels or erosion to be detected closer to the river.	Non-peat forming seasonally inundated forest close to the river, transitioning to shallow palm-dominated peat swamp after 2 km. Max. peat thickness is 2.8 m at 5.75 km.	68.0
Ipombo	Hypothesis testing	Congo River	6.0	Transect runs perpendicular to Congo River. It first crosses suspected river-influenced floodplain channels, then *terra firme* forest at a levee, before entering a depression with likely palm-dominated peat swamp forest.	Peat expected towards interior, after the levee. No or little peat expected in river-influenced floodplain due to nutrient-rich river water.	Non-peat forming seasonally inundated forest close to the river. Shallow peat-forming, seasonally inundated hardwood forest, after 2.25 km. Max. peat thickness is 3.7 m at 3.5 km.	68.0
Pombi	Assessing mapping capabilities	Ruki River	3.0	Transect located opposite to Mpeka on the left-bank of Ruki River, on higher elevation in a likely small channel of the dendritic drainage system. Crosses sites either predicted as peat by the ML model based on DEM or HAND alone as elevation input in the model, but not both.	Likely false positive by Dargie et al.[9]. No peat expected because of higher relative elevation above the river.	Non-peat forming seasonally inundated forest, transitioning to *terra firme* forest further inland. No peat found.	46.2
Bondamba 2	Assessing mapping capabilities	Ruki River	0.5	Short transect close to the main Bondamba transect, traverses an area predicted as peat by the ML model based on DEM only (but not by ML models based on HAND or DEM+HAND).	Likely false positive by Dargie et al.[9]. No peat expected because of higher relative elevation above river.	Seasonally inundated forest. Shallow peat (0.5 m) at 0.25 km only, but otherwise non-peat forming.	33.3

Bondamba 3	Assessing mapping capabilities	Ruki River	0.5	Short transect close to the main Bondamba transect, traverses an area predicted as peat by the ML model based on DEM only (but not by ML models based on HAND or DEM+HAND).	Likely false positive by Dargie et al.[9]. No peat expected because of higher relative elevation above the river.	Seasonally inundated forest. Shallow peat (1.0 m) at 0.5 km only, but otherwise non-peat forming.	33.3
Boleke	Assessing mapping capabilities	Ruki River	2.0	Transect in likely floodplain in bend of the Busira River. Crosses expected savanna, *terra firme* and peat swamp forests. Peat predicted by ML model based on DEM or DEM+HAND, but not HAND alone.	Shallow peat expected because of frequent river inundations in low-lying floodplain, but with possible effects of river erosion.	Non-peat forming seasonally inundated forest, transitioning to *terra firme* forest. No peat found.	22.2
Boboka 2	Assessing mapping capabilities	Congo River	0.5	Short transect close to the main Boboka transect, traverses an area of likely *terra firme* forest next to the peatland margin.	No peat expected.	Non-peat forming seasonally inundated forest throughout.	100.0
Boboka 3	Assessing mapping capabilities	Congo River	0.75	Short transect close to the main Boboka transect, traverses an area of likely *terra firme* forest next to the peatland margin.	No peat expected.	Seasonally inundated forest. Shallow peat (0.5 m) at 0.75 km only, but otherwise non-peat forming.	100.0
Bonzembo	Assessing mapping capabilities	Congo River	1.5	Short transect on the right-bank of the Congo River, traverses a potential palm-dominated channel, into likely *terra firme* forest on higher elevation.	No or little peat expected in floodplain close to the Congo River, due to nutrient-rich river water.	Largely non-peat forming seasonally inundated forest. Shallow peat at 0.75 and 1.5 km (0.7 and 0.3 m).	28.6
Bolombo	Assessing mapping capabilities	Congo River	3.0	Short transect in floodplain on the left-bank of the Congo River, with suspected channels. Likely palm swamp forest present, but potentially limited peat due to nutrient-rich water from the Congo River, or effects of river erosion.	Increasing peat thickness expected towards interior. Effects of nutrient gradient, inundation levels or erosion to be detected on peat thickness.	Non-peat forming seasonally inundated forest throughout.	0.0

Expected landcover descriptions are from Dargie et al.[9]. Preliminary Maximum Likelihood (ML) classification predictions, based on either SRTM-derived DEM, HAND or DEM+HAND as surface elevation data (with a subset of ground-truth data), were used to select sites that could help differentiate the modelling effect of these elevation products. Expected inundations are based on Lee et al.[15] or ALOS PALSAR radar data. Peat thickness measurements were taken every 250 m using the corrected pole-method along all transects (Extended Data Figure 3). Full peat cores were sampled every other kilometre along the ten transects used for hypothesis testing. The eight transects used to assess mapping capabilities were sometimes specifically chosen in locations expected not to have peat or be likely false predictions, hence some transects have low or even 0% correct predictions in Dargie et al.'s map[9].

Supplementary Table 2 | Overview of landcover classes, number of ground-truth datapoints and their sources.

Data source	Water	Savanna	Non-peat forming forest *†*	Palm-dominated peat swamp forest	Hardwood-dominated peat swamp forest	Total peat swamp forest *#*	Total
DRC fieldwork (This study)	19	32	105 (55)	90	136	226	382
ROC fieldwork (ref. [9]**)**	0	13	66 (34)	90	123	213	292
Archaeological database (ref. [30-31]**)**	0	128	171	0	0	0	299
AfriTRON / ForestPlots (ref. [32-34]**)**	0	0	186 (1)	0	5	5	191
Forest and savanna sites around Lomami NP (*pers.comm.,* R.B., G.I., A.C-S.)	0	134	95	0	0	0	229
Savanna around Lomami NP (ref. [35]**)**	0	24	0	0	0	0	24
Palaeo-archaeological research (ref. [11]**)**	0	2	9 (7)	8	4	12	23
Google Earth *	153	143	0	0	0	0	296
Total	172	476	632 (97)	188	268	456	1,736

* *Manually selected from Google Earth for the visually unambiguous water and savanna classes only, spread out across the region.*

† *Non-peat forming forest is* terra firme *forest, including non-peat forming seasonally inundated forest (number of datapoints in parentheses, if present).*

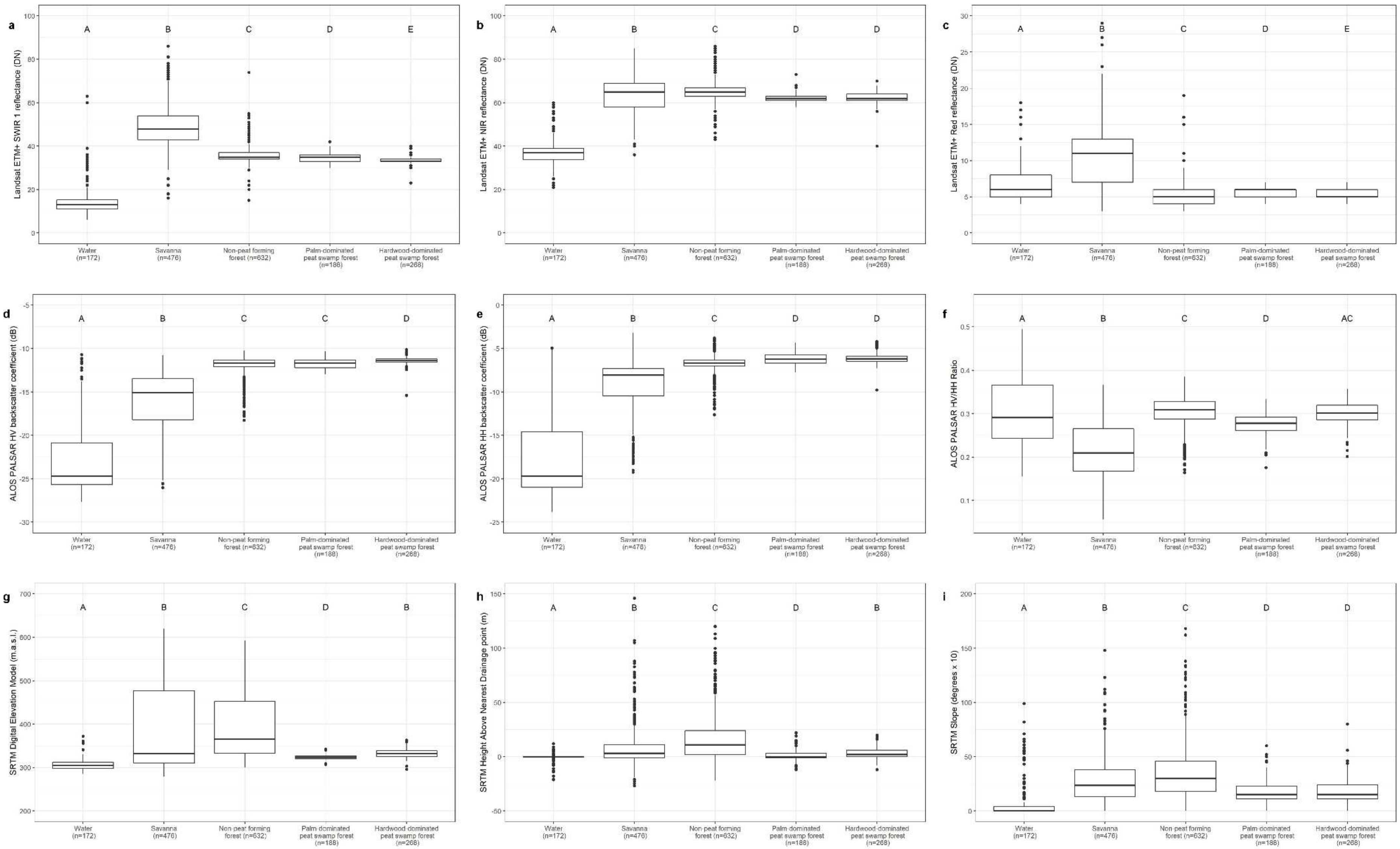
\# *Total peat swamp forest includes both palm-dominated and hardwood-dominated peat swamp forest.*

Supplementary Table 3 | Area estimates and random/spatial cross-validated accuracy results for 3 classification algorithms.

Model	Total peatland area (hardwood- and palm-dominated peat swamp forest, km^2)	Random cross-validation (BA, %)	Random cross-validation (MCC, %)	Spatial cross-validation (MCC, %)					
				Likouala-aux-Herbes (n = 261)	Ubangi (n = 140)	Congo (n = 371)	Ruki (n = 143)	Interfluvial basin peatlands (n = 401)	River-influenced peatlands (n = 540)
ML	167,648 (159,378 - 175,079)	91.9 (90.2-93.6)	78.0 (74.2-81.6)	78.1 (76.5-79.7)	66.6 (63.6-71.4)	41.9 (37.3-45.1)	73.2 (70.1-76.0)	73.9 (71.4-76.0)	65.0 (61.1-67.5)
SVM	135,359 (124,847 - 145,991)	87.0 (84.1-89.7)	77.5 (72.4-81.9)	76.6 (73.0-78.2)	75.1 (71.9-79.0)	34.9 (26.4-44.7)	74.0 (68.1-78.3)	73.2 (67.1-75.8)	68.2 (63.2-71.8)
RF	101,988 (92,596 - 111,358)	89.6 (86.7-93.3)	79.0 (73.5-84.1)	70.4 (60.1-76.6)	71.3 (64.7-78.4)	37.1 (25.9-46.8)	49.2 (39.5-58.5)	66.8 (58.0-74.1)	46.2 (39.2-52.5)

Binary classification performances (random and spatial Matthews correlation coefficient [MCC], %) are reported for supervised Maximum Likelihood (ML), Support Vector Machine (SVM) and Random Forest (RF) classification algorithms. Balanced accuracy results (BA, %) using random cross-validation (CV) are additionally reported to facilitate correct comparison with the first-generation peatland map (median BA 89.8%; 95% CI, 86.0-93.4)[9]. Random CV and area results are obtained from 1,000 randomly stratified 2/3 data splits of 1,736 datapoints. Spatial CV results are obtained for each region by selecting 1,000 randomly stratified 2/3 data splits as training data from all remaining datapoints, validated against all datapoints of the selected (omitted) region. The 'interfluvial basin peatlands' group comprises the Likouala-aux-Herbes and Ubangi River regional gropus in ROC; the 'river-influenced peatlands' group comprises the Congo and Ruki River regional groups in DRC. Models were implemented in IDL-ENVI (ML) or GEE (SVM/RF) for random CV and area calculations. All spatial CV analyses were implemented in R. All figures are median values with 95% confidence intervals in parentheses.

Supplementary Figure 1 | Remote sensing spectral signatures of five landcover classes.



Boxplots showing the distribution of ground-truth datapoints per landcover class for: ***a.*** Landsat ETM+ SWIR I band (Digital Number [DN]); ***b.*** Landsat ETM+ NIR band (DN); ***c.*** Landsat ETM+ Red band (DN); ***d.*** ALOS PALSAR HV backscatter coefficient (dB); ***e.*** ALOS PALSAR HH backscatter coefficient (dB); ***f.*** ALOS PALSAR HV/HH ratio; ***g.*** SRTM Digital Elevation Model (DEM; m.a.s.l.); ***h.*** SRTM Height Above Nearest Drainage point (HAND; m); and ***i.*** SRTM Slope (degrees x 10). Black lines show the median of each class, boxes show the upper and lower quartiles, and the vertical lines show maximum and minimum values. Circles represent potential outlying values. Land cover classes which do not share a common letter have significantly different means for the respective remote sensing product ($P < 0.05$, Kruskall-Wallis multiple comparison [Dunn's] test with BH adjustment). Non-peat forming forest includes both *terra firme* forest (n = 535) and non-peat forming seasonally inundated forest (n = 97).