



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Floristic evidence for alternative biome states in tropical Africa

Citation for published version:

Aleman, JC, Fayolle, A, Favier, C, Staver, AC, Dexter, KG, Ryan, CM, Azihou, AF, Bauman, D, Te Beest, M, Chidumayo, EN, Comiskey, JA, Cromsigt, JPGM, Dessard, H, Doucet, J, Finckh, M, Gillet, J, Gourlet-fleury, S, Hempson, GP, Holdo, RM, Kirunda, B, Kouame, FN, Mahy, G, Gonçalves, FMP, Mcnicol, I, Quintano, PN, Plumptre, AJ, Pritchard, RC, Revermann, R, Schmitt, CB, Swemmer, AM, Talila, H, Woollen, E & Swaine, MD 2020, 'Floristic evidence for alternative biome states in tropical Africa', *Proceedings of the National Academy of Sciences*, vol. 117, no. 45, pp. 28183-28190.
<https://doi.org/10.1073/pnas.2011515117>

Digital Object Identifier (DOI):

[10.1073/pnas.2011515117](https://doi.org/10.1073/pnas.2011515117)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Proceedings of the National Academy of Sciences

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



1 **Classification:** BIOLOGICAL SCIENCES, Environmental sciences

2
3 **Title:** Floristic evidence for alternative biome states in tropical Africa

4
5 **Author Line:**

6 J. C. Aleman^{1,2}, <https://orcid.org/0000-0003-0835-7015>

7 A. Fayolle¹, <https://orcid.org/0000-0002-6770-0031>

8 C. Favier³, <https://orcid.org/0000-0002-1126-6950>

9 A. C. Staver⁴,

10 K. G. Dexter^{5,6}, <https://orcid.org/0000-0001-9232-5221>

11 C. M. Ryan⁵, <https://orcid.org/0000-0002-1802-0128>

12 A. F. Azihou⁷, <https://orcid.org/0000-0002-7400-4528>

13 D. Bauman^{8,9}, <https://orcid.org/0000-0001-9115-6518>

14 M. te Beest^{10,11,12}, <https://orcid.org/0000-0003-3673-4105>

15 E. N. Chidumayo¹³, <https://orcid.org/0000-0002-1127-3106>

16 J. A. Comiskey^{14,15}, <https://orcid.org/0000-0001-6710-1269>

17 J. P. G. M. Cromsigt^{10,11,16}, <https://orcid.org/0000-0002-8632-9469>

18 H. Dessard^{17,18},

19 J-L. Doucet¹, <https://orcid.org/0000-0002-3277-898X>

20 M. Finckh¹⁹, <https://orcid.org/0000-0003-2186-0854>

21 J-F. Gillet¹,

22 S. Gourlet-Fleury^{17,18}, <https://orcid.org/0000-0002-1136-4307>

23 G. P. Hempson²⁰, <https://orcid.org/0000-0001-8055-4895>

24 R. M. Holdo²¹, <https://orcid.org/0000-0002-2484-0587>

25 B. Kirunda²²,

26 F. N. Kouame²³, <https://orcid.org/0000-0003-1112-858X>

27 G. Mahy¹,

28 F. Maiato P. Gonçalves²⁴, <https://orcid.org/0000-0002-8859-7491>

29 I. McNicol⁵, <https://orcid.org/0000-0003-3073-5668>

30 P. Nieto Quintano⁵, <https://orcid.org/0000-0001-8333-306X>

31 A. J. Plumptre^{22,25,26}, <https://orcid.org/0000-0002-9333-4047>

32 R. C. Pritchard^{5,27},

33 R. Revermann^{19,28}, <https://orcid.org/0000-0002-7044-768X>

34 C. B. Schmitt^{29,30}, <https://orcid.org/0000-0002-8527-9682>

35 A. M. Swemmer³¹, <https://orcid.org/0000-0003-1378-7394>

36 H. Talila³²,

37 E. Woollen⁵, <https://orcid.org/0000-0002-6504-4835>
38 M. D. Swaine³³, <https://orcid.org/0000-0002-8301-148X>

39

40 **Author Affiliation:**

41 ¹ Gembloux Agro-Bio Tech, Université de Liège, Passage des Déportés 2, 5030 Gembloux, Belgium.

42 ² Département de Géographie, Université de Montréal, Montréal (Québec), H2V 0B3, Canada.

43 ³ ISEM, Université de Montpellier, CNRS, EPHE, IRD, 34000 Montpellier, France.

44 ⁴ Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06520, USA.

45 ⁵ School of GeoSciences, University of Edinburgh, Edinburgh, EH8 9YL, UK.

46 ⁶ Royal Botanic Garden Edinburgh, Edinburgh, EH3 5LR, UK.

47 ⁷ Laboratory of Applied Ecology, Faculty of Agronomic Sciences, University of Abomey-Calavi, Cotonou,
48 Benin.

49 ⁸ Environmental Change Institute, School of Geography and the Environment, University of Oxford,
50 Oxford, OX1 3QY, UK.

51 ⁹ Plant Ecology and Biogeochemistry, Université Libre de Bruxelles, 1050 Bruxelles, Belgium.

52 ¹⁰ Copernicus Institute of Sustainable Development, Utrecht University, 3584 CS Utrecht, The
53 Netherlands.

54 ¹¹ Centre for African Conservation Ecology, Nelson Mandela University, Port Elizabeth, 6031, South
55 Africa.

56 ¹² South African Environmental Observation Network, Pietermaritzburg, 3201, South Africa.

57 ¹³ Makeni Savanna Research Project, PO Box 50323, Ridgeway, Lusaka, Zambia.

58 ¹⁴ Inventory and Monitoring Program, National Park Service, Fredericksburg, Virginia, 22405, USA.

59 ¹⁵ Smithsonian Institution, Washington, DC 20002, USA.

60 ¹⁶ Department of Wildlife, Fish, and Environmental Studies, Swedish University of Agricultural Sciences,
61 Skogsmarksgränd, SE-901 83, Umeå, Sweden.

62 ¹⁷ Forêts et Sociétés, Université de Montpellier, Cirad, 34000 Montpellier, France.

63 ¹⁸ Cirad, Forêts et Sociétés, 34398 Montpellier, France.

64 ¹⁹ Biodiversity, Evolution and Ecology of Plants, Institute of Plant Science and Microbiology, University of
65 Hamburg, 22609 Hamburg, Germany.

66 ²⁰ Centre for African Ecology, School of Animal, Plant and Environmental Sciences, University of the
67 Witwatersrand, Johannesburg, 2000, South Africa.

68 ²¹ Odum School of Ecology, University of Georgia, Athens, GA 30602, USA.

69 ²² Wildlife Conservation Society, Kampala, Uganda.

70 ²³ University Nangui Abrogoua, Nature Sciences Unit, 31 BP 165 Abidjan, Côte d'Ivoire.

71 ²⁴ Herbário do Lubango, ISCED – Huíla, Rua Sarmiento Rodrigues No. 2, C.P. 230, Lubango, Angola.

72 ²² Head KBA Secretariat, c/o BirdLife International, Cambridge, CB2 3QZ, UK.

73 ²⁵ Conservation Science Group, Zoology Department, Cambridge University, Cambridge, CB2 3EJ, UK.

74 ²⁷ Global Development Institute, University of Manchester, Manchester, M13 9PL, UK
75 ²⁸ Faculty of Natural Resources and Spatial Sciences, Namibia University of Science and Technology,
76 Windhoek, Namibia.
77 ²⁹ Center for Development Research (ZEF), University of Bonn, 53113 Bonn, Germany.
78 ³⁰ Nature Conservation and Landscape Ecology, University of Freiburg, 79106 Freiburg, Germany.
79 ³¹ South African Environmental Observation Network, Phalaborwa, 1389, South Africa.
80 ³² Department of Ecotourism and Biodiversity Conservation, College of Agriculture and Natural Resource,
81 Madda Walabu University, Bale Robe, Ethiopia.
82 ³³ Institute of Biological and Environmental Sciences, University of Aberdeen, Aberdeen, AB24 3FX, UK.
83

84 **Corresponding Author:** Adeline FAYOLLE, Gembloux Agro-Bio Tech, Université de Liège, Passage des
85 Déportés 2, 5030 Gembloux, Belgium, +32 81 62 22 98, adeline.fayolle@uliege.be
86

87 **Keywords:** Alternative stable states | Tropical biomes | Tree species composition | Precipitation and
88 seasonality | Fire | Paleo-ecological data
89

90 **Abstract:** The idea that tropical forest and savanna are alternative states is crucial to how we manage
91 these biomes and predict their future under global change. Large-scale empirical evidence for alternative
92 stable states is limited however and comes mostly from the multimodal distribution of structural aspects of
93 vegetation. These approaches have been criticized, as structure alone cannot separate out wetter
94 savannas from drier forests for example, and there are also technical challenges to mapping vegetation
95 structure in unbiased ways. Here, we develop an alternative approach to delimit the climatic envelope of
96 the two biomes in Africa, using tree species lists gathered for a large number of forest and savanna sites
97 distributed across the continent. Our analyses confirm extensive climatic overlap of forest and savanna,
98 supporting the alternative stable states hypothesis for Africa, and this result is corroborated by paleo-
99 ecological evidence. Further, we find the two biomes to have highly divergent tree species compositions
100 and to represent alternative compositional states. This allowed us to classify tree species as forest vs
101 savanna specialists, with some generalist species that span both biomes. In conjunction with geo-
102 referenced herbarium records, we mapped the forest and savanna distributions across Africa and
103 quantified their environmental limits, which are primarily related to precipitation and seasonality, with a
104 secondary contribution of fire. These results are important for the ongoing efforts to restore African
105 ecosystems, which depend on accurate biome maps to set appropriate targets for the restored states, but
106 also provide novel empirical evidence for broad scale bistability.
107

108 **Significance statement:** We develop a biogeographic approach to analyzing the presence of alternative
109 stable states in tropical biomes. Whilst forest-savanna bistability has been widely hypothesized and
110 modeled, empirical evidence has remained scarce and controversial, and here, applying our method to

111 Africa, we provide large scale evidence that there are alternative states in tree species composition of
112 tropical vegetation. Furthermore, our results have produced novel and more accurate maps of the forest
113 and savanna distributions in Africa, which takes into account differences in tree species composition, and
114 a complex suite of determinants. This result is not only important for understanding the biogeography of
115 the continent, but also to guide large-scaled tree planting and restoration efforts planned for the region.

116

117 **Introduction**

118 Tree cover and canopy openness are commonly used to differentiate tropical forests and savannas, but
119 the difference between the two biomes is not just a matter of structure (1). Whereas tropical moist forests
120 form a closed canopy with a complex vertical structure, savannas are more open, allowing fire- and
121 drought-adapted grasses to grow in the understory. Typically, forest dominates in wetter areas while
122 savannas occur in drier, seasonal areas (1, 2), although transitions between forest and savanna are not
123 rigidly determined by climate (2–4). Soils and topography can be locally and regionally important, but at
124 intermediate rainfall (between 1,000-2,500 mm globally), forest and savanna, both widespread, potentially
125 represent alternative stable states maintained by feedbacks between tree cover and disturbances —
126 specifically fire (3, 5–7) and chronic herbivory (8). As a result, forest and savanna tree species show
127 contrasting adaptations (9, 10), and transitions across the forest-savanna boundary are characterized by
128 high species turnover (10, 11).

129 Despite these functional and floristic differences, most recent efforts devoted to understanding
130 forest-savanna transitions have relied solely on information about the canopy structure, whether using
131 satellite-based maps of percent tree cover (3, 5, 6, 12), canopy height (13), or using field data on tree
132 basal area (14). However, these studies have two types of shortcomings. The first type is linked to the
133 use of remote sensing products which are ill-suited to capture the structural difference between savanna
134 and forest. Most tree cover products are parameterized to identify trees greater than 5 m in height (15,
135 16) and their spatial resolution is not sufficient to correctly characterize vegetation structure. Even though
136 Landsat-based tree cover is available at 30 m resolution and MODIS tree cover at 250 m, the accuracy of
137 these products is low in open systems (16), and although it increases with spatial aggregation, it then fails
138 to represent the fine-scale heterogeneity characteristic of savanna. Moreover, the tree cover threshold
139 used to differentiate forest from savanna seems to depend strongly on calibration, with higher thresholds
140 (75%) for Landsat tree cover (12) and lower (55-65%) for MODIS tree cover (3, 17). The second type of
141 shortcomings is even more important since defining forest and savanna using only canopy structure
142 misses key aspects of forest vs savanna function tied to differences in the structure of the grass layer and
143 in the tree species composition (1). The relatively open canopies of drier forests can be structurally similar
144 to those of wetter savannas (18) but are not separable with a structural approach, even though the two
145 differ dramatically in terms of species composition and ecosystem function (19). Also, systems with
146 similar physiognomies can have different histories, reflecting human land-use practices and recent
147 disturbances (12). For instance, forests subjected to human disturbances (such as clearing) can appear

148 more like savannas, while some practices (particularly fire suppression) can make savannas appear more
149 like forests (19, 20).

150 Biodiversity data available for sub-Saharan Africa (21–23) and for tropical Africa (24–27), offer
151 new opportunities to differentiating and mapping the tropical forest and savanna biomes at continental
152 scale. Here, we delineate the distribution of forest and savanna, and test for forest-savanna bistability,
153 using a biogeographic approach, which reflects the evolutionary history and ecology of each biome, and
154 does not rely on any structural metrics. We combined data on native tree species for sites identified
155 earlier as forest (ref 26, n=455) and savanna (ref 27, n=298) and covering the full extent of these biomes
156 in Africa (28). Each site consists of a species list assigned to either forest or savanna by the original
157 authors or by the data contributors who conducted the floristic surveys according to vegetation structure,
158 species composition and/or ecosystem functioning (see Materials and methods). We used tree species
159 composition from these 753 floristic surveys to determine the climatic envelopes of the forest and
160 savanna biomes (and their overlap) and to analyze the differences in tree species composition between
161 them. We then used georeferenced herbarium collections (24, 25) to expand the floristic information
162 spatially and describe the complete distribution of forest and savanna across Africa, and to identify the
163 determinants of their present day distribution.

164

165 **Extensive climatic overlap of forest and savanna in Africa**

166 Alongside the expected patterns of forest in wetter areas and savanna in drier and more seasonal ones
167 (1, 2), also retrieved here (Fig. 1A, S1), we found an extensive climatic area (Fig. 1B-C, S2) within which
168 both forest and savanna sites are widespread, which we term the '*bistable*' region. Mean Annual
169 Precipitation (MAP) partly differentiates the forest, the *bistable* region, and the savanna; although the
170 climatic gradient used to determine and map them is more complex and integrates precipitation and
171 seasonality (see Materials and methods for the details of the Principal Component Analysis on gridded
172 climate data). We find a large extent of forest, 1.8 million km², covering almost the whole Guineo-
173 Congolian Region (Fig. 1D), in line with the stability of forest in Lower Guinea suggested over the last two
174 millennia by pollen records (29), but challenging previous findings of bistability across the entire Congo
175 basin (3). Savanna was found to occupy the majority of the areas north, east and south of the Congo
176 Basin, under drier and more seasonal conditions, and totaling almost 8.4 million km². The savanna
177 notably includes the west coast of southern Africa (Fig. 1D), which contrasts with description of this area
178 as thicket, scrub or bush land (28). This is due to the presence of northern (Sahel) and eastern (Horn)
179 savanna sites with extremely arid climate in the floristic surveys, which extended the climatic envelope of
180 savanna this far, and also because northern and western savannas are not completely analogous
181 climatically to southern and eastern ones (27).

182 Meanwhile, the *bistable* region covers a vast area of more than 7.5 million km², often described
183 as woodland (28) and spans a broad MAP range (700-1900 mm, Fig. S2). The *bistable* region was
184 previously proposed to span a 1,000-2,000 mm MAP range in Africa (6) or, alternately, to extend to drier

185 sites (> 650 mm) in a study restricted to savannas (30) and to wetter sites (1,000-2,500 mm) in the global
186 tropics (3, 12). Our *bistable* region can be separated into two sub-regions: a wetter region more likely to
187 be forest (hereafter, *bistable* forest) and a drier and more seasonal one more likely to be savanna
188 (hereafter, *bistable* savanna). *Bistable* forest is located in the periphery of the Guineo-Congolian Region,
189 but also includes the Congolese Batéké Plateau and the Dahomey gap, which are currently dominated by
190 savanna. This stresses the importance of factors other than climate in the outcome of forest-savanna
191 bistability. The Batéké Plateau corresponds to the northern limit of the hyper-well-drained relict dunes of
192 the Kalahari sands, while the Dahomey gap is well known to have shifted to savanna between 4,500 and
193 3,400 years ago following an abrupt climatic change (31). Thus, the predominance of savanna in areas
194 more likely to be forest has arisen either because of soil characteristics (*i.e.*, sandy soils, 30) or the legacy
195 of past climatic changes (29) and/or past disturbances (31, 32).

196 Long-term paleo-ecological data (from lacustrine and soil cores) around the Gulf of Guinea
197 largely corroborates these areas of biome transition and stability (inset, Fig. 1D). For the *bistable* region,
198 forest-to-savanna transitions were predominant with 16 transitional vs nine stable sites, while for the
199 forest, stable biome identity was retrieved for ten out of 13 paleo-sites. For **the** savanna, the stability of
200 savanna vegetation has been demonstrated around Lake Tilla (#21 on the map) in Nigeria (33).

201 Current fire regime is also an important determinant of savanna distribution worldwide (2, 3, 6),
202 specifically in relatively wetter areas (20, 30), and here we found that *bistable* savanna is more likely to
203 burn than either **forest** or *bistable* forest, and also than savanna (Fig. S3). The latter is explained by the
204 lower productivity of drier savannas (30), but could also be linked to the presence of semi-desert
205 grasslands (28) in our study area (Sahel, Namib, Kalahari), where the discontinuous grass layer prevent
206 the spread of fires.

207

208 **African forests and savannas have distinct floristic composition**

209 We found a marked dissimilarity in tree species composition between forest and savanna sites, based on
210 floristic surveys. Pairwise comparisons showed a mean species turnover of 98% between forest and
211 savanna sites, much higher than the turnover within the same biome (Fig. S4). This result confirms the
212 divergence of the forest and savanna floras (22) and is consistent with **the** results of plot-based studies of
213 forest-savanna transition in South Africa (10), Ghana (11), and across the tropics (18). Because forest
214 and savanna represent alternative compositional states, we were confident in categorizing tree species
215 into significant indicators of forest, or forest specialists (n=825 species, 48%), and significant indicators of
216 savanna, or savanna specialists (n=523, 31%). Species that were not a significant indicator of either
217 forest or savanna were interpreted as generalists (n=359, 21%). Our group of generalists included true
218 generalists but also rare species or those otherwise rarely represented in the floristic surveys (Fig. 2).
219 Evidence for biome specialization was even stronger when only common species were analyzed (*i.e.*,
220 those present at ≥ 10 sites), with only 8% of common species classified as generalists. It is worth noting
221 that our results for 1,707 tree species represent only a fraction of the c. 45,000 flowering plant species

222 reported for sub-Saharan Africa (34). The specificity of the forest and the savanna flora is therefore likely
223 underestimated, since grasses and forbs were not included, despite being a highly distinctive component
224 of the savanna flora (22). Functionally, though, these results are demonstrative, since the tree flora
225 **already captures some functional differences between the forest and savanna biomes (10) but it would be**
226 **of great interest to extend the analyses to other growth forms to test the patterns observed for trees.**

227 The regions described above based on climate (Fig. 1) have distinct compositions (Fig. S5),
228 assessed using an independent occurrence dataset (Fig. 2; see Materials and methods) derived from
229 georeferenced herbarium records (24, 25). Forest and *bistable* forest are composed mainly of forest
230 specialists, and savanna and *bistable* savanna are composed mainly of savanna specialists. Indeed,
231 forest specialists were found to dominate the Guineo-Congolian Region and its periphery, and to be
232 abundant in East African Mountains (Fig. 2A). Within the Guineo-Congolian Region, Upper and Lower
233 Guinea were better sampled than Congolia, for which herbarium records are sparse, and Upper Guinea
234 includes the warm and wet forest sheltered by the Guinean Highlands in Sierra Leone and Guinea
235 (Conakry). The latter were not included in our mapped extent of forest (Fig. 1D) likely because the floristic
236 surveys we used do not cover this climate zone (positive scores on PC1 and PC2, Fig. 1A). Savanna
237 specialists dominate in the north and west of Africa, including the Dahomey gap, and in the east and
238 south (Fig. 2C). Interestingly, generalist species were found to be widespread and present within each
239 region, and while they are more frequent in the *bistable* region, the difference is small (Fig. S5B) and
240 generalists almost never dominate (Fig. 2B, S6).

241 242 **Savanna-forest co-existence is possible, but restricted spatially**

243 Based on a biome index, indicating the relative dominance of savanna (-1) and forest (1) specialists (see
244 **Materials and methods**), we find that forest-savanna coexistence is limited (Fig. 3) even though the
245 climate envelope where it is possible is large (42.4% of the study area and 25% of the African continent,
246 Fig. 1D). Vast and continuous areas are dominated by the extremes of the biome index, representing
247 either strict savanna or strict forest (Fig. S7), and these extremes of the biome index correspond to
248 recognized centers of endemism (28), *i.e.*, the Guineo-Congolian Region where forests are predominant,
249 and the Sudanian and Zambebian Regions where savannas and woodlands are predominant, and both
250 have specific floras and faunas (21–24). In contrast, intermediate values of the biome index, which
251 correspond to the third mode in the frequency distribution (inset, Fig. 3), appear spatially restricted (Fig.
252 S7). The zones of intermediate biome index, which arise either from frequent generalists (as in East
253 Africa) or from a mixture of forest and savanna specialists elsewhere (Fig. 2B, S6), correspond to long
254 recognized transition zones, forest-savanna mosaics (28), and to recently deforested areas (35). A
255 deeper investigation shows however that the possible intermediate state mostly corresponds to data-
256 deficient areas since the tri-modal signal shifted towards a bi-modal signal when including only pixels with
257 at least five or ten of our classified tree species (Fig. S8). Thus, the floristic surveys (26, 27) combined

258 with georeferenced herbarium records (24, 25) suggest that there are two floristic states in the intact
259 tropical African vegetation.

260 Finally, we provide estimates of the current distribution of the forest and savanna biomes across
261 the African continent combining both floristic and environmental information (Fig. 4). First, we re-
262 computed the biome index on interpolated species distributions derived from ordinary kriging (Fig. S9),
263 while accounting for spatial autocorrelation (inset, Fig. 4A). This shows that there is a strong spatial
264 structure for specialist species—whether forest or savanna— while generalist species do not show any
265 spatial structure, suggesting that they do not correspond to an ecologically meaningful third group. We
266 additionally include principle components of climate (PC1 and PC2 from Fig. 1) and soil information (sand
267 percentage and cation exchange capacity), as well as data on fire and herbivory (see Materials and
268 methods for the source of the datasets) as potential determinants of the biome index in a random forest
269 regression (Fig. 4B). Topographic and hydrologic factors (e.g., seasonal flooding, topographically
270 mediated frost occurrence or toxic heavy metal concentrations in soils) that are important determinant of
271 mosaic landscapes locally (28) were not included at this continental scale. The predictions of our random
272 forest model on a validation subset of 10% of data pixels were accurate ($R^2=0.81$ with a non-significant
273 intercept and a predicted vs observed slope of 0.82 ± 0.03). This analysis of the joint determinants of both
274 forest and savanna showed that precipitation and seasonality (PC1) are the main determinants of the
275 biome index. Fire and, to a lesser extent, temperature (PC2, which is related to altitude) were also
276 important drivers. Herbivory (including livestock and wild herbivore biomass) and sandy soils were found
277 to have an equivalent and moderate impact, at this continental scale. The environmental model notably
278 predicts a smaller forest area than the interpolated species distribution data (Fig. 4A) and than the raw
279 data (Fig. 2, 3). The latter possibly results from preferential sampling of forest trees in herbarium
280 collections. More importantly, our results support the primary role of precipitation (1, 2), precipitation
281 seasonality (36), and fire (2, 3, 6, 30) in the distribution of both forest and savanna in Africa, while most
282 analyses of the determinants of savanna distribution have so far ignored forest (2, 30). Therefore, the
283 predictions of the random forest (Fig. 4B) should be considered as our best estimate of the current
284 distribution of the forest and savanna biomes across the African continent.

285

286 **Conclusions and practical implications**

287 In this study, we provided evidence for two compositional states **across the African continent**, alternatively
288 dominated by forest and savanna species. These two states can coexist in a wide range of climates, but
289 our biome index showed that current day spatial co-existence is spatially restricted, with one state or the
290 other dominating over vast areas. This supports the notion that the two biomes are stable alternatives,
291 and, consistent with the predictions of alternative stable state theory, suggests the potential for abrupt
292 shifts in composition under external perturbations such as climate change, or altered fire regimes, as
293 demonstrated by paleo-ecological evidence. **With our correlative approach, we were not able to**
294 **determine the ecological mechanisms allowing savanna persistence, but the facts that fire is more**

295 frequent in the bistable savanna area, and that fire is the second determinant of the biome index after the
296 precipitation and seasonality gradient, support the hypothesis of fire acting as a major feedback
297 mechanism, allowing savanna under climates favorable to forest. These results are important for the
298 ongoing efforts to restore African ecosystems, which depend on accurate biome maps to set appropriate
299 targets for the restored states. This will help avoid, for example, inappropriate planting of forest trees, or
300 attempts to “restore” forest in savanna areas. Up to now, because of the wide climatic zone where forest
301 and savanna are both widespread, savanna has often been mischaracterized as degraded forest (28, 37)
302 and has been seen as a target for higher carbon storage via afforestation (37, 38). This viewpoint
303 neglects several features of savannas, including their substantial belowground carbon storage (39), their
304 high biodiversity and endemism (40), and their socio-economic value (41). Our biome index map
305 therefore is a useful tool for restoration, as it could help set appropriate species targets, and to identify
306 degraded forest and encroached savanna areas which should be restored using forest and savanna
307 specialists, if needed. Indeed, the current trend within *bistable* savannas is toward woody encroachment
308 (42) and in this case restoration strategies mainly involve tree-cutting and prescribed fire to restore
309 biodiversity and ecosystem services (43). Conversely, restoring arid savannas to prevent desertification
310 might involve tree and shrub planting (44), and in this case using savanna specialists is and has been
311 warranted. In the spatial extent where both biomes currently coexist, potentially both may be restored but
312 a deeper investigation is needed, for example using long-term data, to determine the baseline biome (45),
313 and human preference should also be considered.

314

315 **Materials and methods**

316 To delineate the distribution of forest and savanna across the African continent and to identify their
317 underlying determinants, we developed a biogeographic approach consisting of four steps, each testing a
318 specific hypothesis. First, we used the location of specific sites originally classified as forest (26) or
319 savanna (27) to test the forest-savanna bistability, *i.e.*, whether forest and savanna can both occur in
320 areas with similar climates, earlier identified *via* remote-sensing products at the global (3, 5) and
321 landscape (7) scales. In addition, for a selected area around the Gulf of Guinea with good paleo-
322 ecological data, we tested for historical biome transitions to confirm forest-savanna bistability. Second, we
323 used tree species lists for the same forest and savanna sites to test the biome specialization of the tree
324 species, and to evaluate the overall distinctiveness of forest and savanna tree floras (22). Third, we
325 gridded the geo-referenced herbarium records available for tropical Africa (24, 25), and tested the degree
326 of biome specialization characterizing each pixel, thereby testing the forest-savanna bistability from a
327 floristic perspective. Finally, we used spatial and environmental predictions to delineate the distribution of
328 the forest and savanna biomes at the continental scale and to test the relative influence of climate, fire,
329 herbivory, and soils on the forest and savanna distributions (2, 30).

330

331 **Datasets**

332 (i) *Floristic data*. We combined native tree species lists for 455 forest (26) and 298 savanna sites (27)
333 across sub-Saharan Africa (Fig. 1D, Table S1). Each site consists of a species list assigned to either
334 forest (excluding 'montane forest', and 'thicket') or savanna (including 'woodland') by the original authors
335 or by the data contributors who conducted the floristic surveys. Forest sites corresponded to moist
336 broadleaved forest, though restricted to lowland and terra firme, while savanna sites both included
337 wooded grasslands and woodlands. Indeed, several southern African woodland types, such as *Miombo*,
338 *mopane* and *Baikiaea* woodlands (28), are considered as savannas since they burn regularly and have a
339 relatively continuous grass layer (1, 20). Inclusion of a species list corresponding to either forest or
340 savanna was first based on the definition given in the source according to vegetation structure (savannas
341 do not show closed canopy and have a grass layer) and/or on clear indications of species composition
342 and ecosystem functioning. The combined floristic dataset analyzed consisted of a total of 1,707 species
343 occurring in more than five of our combined forest and savanna sites, belonging to 590 genera and 110
344 families (Table S1). The majority of the species were trees, but some shrubs were included in the
345 savanna sites. The taxonomy was standardized according to the African Plant Database (<http://www.ville-ge.ch/musinfo/bd/cjb/africa/recherche.php?langue=an>) in January 2018.

347 To spatially extrapolate the floristic information from our forest and savanna sites, we used
348 independent data of plant species occurrence available online (<https://gdauby.shinyapps.io/rainbio/>) and
349 assembled in the RAINBIO project (<http://rainbio.cesab.org/>). These represent the largest ever collation of
350 georeferenced herbarium records for tropical Africa (24, 25).

351 We additionally gathered information on past vegetation for 42 paleo-ecological sites located
352 around the Gulf of Guinea (Table S2) from published records of lacustrine fossil pollen (n=24), lacustrine
353 phytoliths (n=1), and phytoliths (n=4) and $\delta^{13}\text{C}$ of soil organic matter (n=14) from soil profiles. We
354 selected records with at least 2,000 years of vegetation history. From the original studies, we identified 20
355 paleo-sites with stable vegetation during the Holocene, either forest or savanna, and 22 which
356 experienced a biome transition, always from forest to savanna. Two sites, Lakes Bosumtwi and Barombi
357 Mbo, date back to the Last Glacial Maximum (~21,000 years BP), during which the vegetation was
358 probably a mosaic of savanna (identified by an increase in Poaceae) and forest elements.

359
360 (ii) *Environmental drivers*. Our study area corresponded to the geographical extent covered by the floristic
361 surveys, but excluding two extreme sites in the very south (Fort Hare, South Africa) and east (Ogo,
362 Somalia), ending up with latitudes ranging from 28°S to 17.5°N, and longitudes from 17°W to 42.5°E.

363 To characterize the climatic conditions across the study area, we used the bioclimatic variables
364 from Worldclim version 2 at 0.5° resolution (46), which consist of 19 variables describing precipitation and
365 temperature, and including mean annual temperature (MAT), mean diurnal range (mdrT), isothermality
366 (isoT), temperature seasonality (Tsea), maximum temperature of the warmest month (Twm), minimum
367 temperature of the coldest month (Tcm), temperature annual range (Tar), mean temperature of the
368 wettest quarter (Tweq), mean temperature of the driest quarter (Tdq), mean temperature of the warmest

369 quarter (Twaq), mean temperature of the coldest quarter (Tcq), mean annual precipitation (MAP),
370 precipitation of the wettest month (Pwm), precipitation of the driest month (Pdm), precipitation seasonality
371 (Psea), precipitation of the wettest quarter (Pweq), precipitation of the driest quarter (Pdq), precipitation of
372 the warmest quarter (Pwaq), and precipitation of the coldest quarter (Pcq).

373 Beyond climate, we also considered disturbance, specifically fire and chronic herbivory (herbivore
374 biomass), and soils (sand percentage and cation exchange capacity) because they have been identified
375 as determinants of savanna distribution (2, 30). Estimates of fire frequency were derived from the burned
376 area product from Moderate Resolution Imaging Spectroradiometer Active Fire (MODIS) data at 1-km
377 resolution (47) over the 2003-2012 period. Current-day estimates of herbivore biomass at 0.5° resolution
378 across the African continent included livestock and also historical wild herbivore biomass filtered by
379 landscape change indices (48). In this map, originally developed at 1° resolution excluding grid cells with
380 > 50 % in inland water bodies, and then later refined at 0.5° resolution, there is no information for a large
381 edge along the coast and around inland waterbodies. Among available soil variables at 250-m resolution
382 for sub-Saharan Africa (<https://soilgrids.org/>), we retained the percentage of sand and the cation
383 exchange capacity in the top layer (0-5 cm) of soil (49). Soil variables were first projected and resampled
384 to fit the 0.5° resolution.

385 The Global Lakes and Wetlands Database (GLWD, [https://www.worldwildlife.org/pages/global-](https://www.worldwildlife.org/pages/global-lakes-and-wetlands-database)
386 [lakes-and-wetlands-database](https://www.worldwildlife.org/pages/global-lakes-and-wetlands-database)) was used to map major waterbodies, including lakes (GLWD-1) and rivers
387 (GLWD-2).

388

389 **Data analyses**

390 *(i) Climatic envelope of forest and savanna.* To detect the major climatic gradients over the entire study
391 area, we performed a Principal Components Analysis (PCA) on the matrix of climate variables at 0.5°
392 resolution (Fig. 1, S1). On the factorial plane defined by the first two Principal Components (PC1 and
393 PC2), the pixels containing floristic surveys were colored and the climatic space they cover was defined
394 by a convex hull (Fig. 1A). The first climatic gradient (PC1, describing 42.3% of the total variance) is
395 characterized primarily by variation in precipitation, and by seasonality of precipitation and of temperature
396 (Fig. 1A, S1), and differentiates forest and savanna sites. The second climatic gradient (PC2, 31.1% of
397 variance) is driven by temperature (Fig. 1A), and differentiates northern and western savannas from
398 southern and eastern savannas and woodlands (Fig. S1C), which are generally found at higher altitudes
399 (28) and thereby exposed to colder climates (27), except for coastal plains (*e.g.*, in Mozambique).

400 Tropical forest and savanna have been shown to co-occur within the same MAP range (3, 5, 6),
401 which was also apparent along PC1, the more complex climatic gradient integrating precipitation and
402 seasonality (Fig. 1A, C). We defined this climatic space as the region of potential bistability, or '*bistable*'
403 region, because forest and savanna are both widespread, and because the stability over time has been
404 previously demonstrated through mechanistic modelling (6). Because of the non-uniform distribution of
405 sites along PC1, we sub-sampled the sites with a stratified random sampling by taking, with replacement,

406 20 sites for each bin of 0.1 units PC1 (*i.e.*, 20 samples with PC1 between -4 and -3.9, and so on).
407 Numerically, the *bistable* region was defined as the region over which the slope in the frequency of the
408 two biomes at each point over PC1 exceeded the mean of the slope averaged over the entire range (Fig.
409 S2A). In practice, we plotted the frequency of forest and savanna sites along PC1, and computed the
410 slope of these curves for each PC1 bin. The slopes were averaged for the entire PC1 range. The bistable
411 region was then defined as the area of the plot for which individual slopes are above the average. We
412 found that forest and savanna sites coexist between PC1 values of 0.141 and 4.235 (Fig. 1C), and that
413 their frequency distribution is equal to 0.5 for a PC1 value of 1.943, such that below this value, the
414 probability of savanna occurring is higher than the probability of forest, and *vice versa*.

415 After defining these thresholds, we mapped the potential biomes based on gridded climatic
416 variables, but we restricted our spatial predictions to the geographic extent and to the climate space
417 covered by the floristic surveys. The latter was defined with a convex hull on the site scores along PC1
418 and PC2 (Fig. 1A). Predictions were thus not possible for some mountain areas in eastern Africa and for
419 a large area in southern Africa corresponding to (semi-desert) grasslands (28). First, savanna, is defined
420 as having values of PC1 < 0.141, and forest is defined as having PC1 ≥ 4.235 (Fig. 1C, S2A). Then,
421 within the *bistable* region, *bistable* savanna, which corresponds to a greater probability of savanna in the
422 climatic space of co-occurrence (Fig. 1C), is defined by $0.141 \leq PC1 < 1.943$ and *bistable* forest, which
423 corresponds to a greater probability of forest in the climatic space of co-occurrence, is defined by $1.943 \leq$
424 PC1 < 4.235. Finally, for clarity and comparison with earlier results, we conducted the same analysis with
425 MAP, sampling 20 sites for each MAP bin of 50 mm (Fig. S2B), and found a range of 700-1900 mm for
426 the *bistable* region (Fig. S2C).

427 To confirm the forest-savanna bistability we mapped the paleo-ecological evidence of biome
428 transitions (Table S2). To specifically test whether the *bistable* region is more likely to burn than the other
429 potential biomes, we computed mean fire frequency for each potential biome (Fig. S3). The latter analysis
430 was performed both at the pixel- and at the site-scale.

431
432 (ii) *Floristics of forest and savanna*. We first computed the overall dissimilarity in tree species composition
433 between all pairs of forest and savanna sites (Fig. S4), using the Simpson index of beta-diversity (β_{sim}),
434 which is known to be richness-independent. We then performed the *IndVal* procedure (50) on the
435 presence matrix of the 1,707 native tree species encountered in the 455 forest sites and the 298 savanna
436 sites. This allowed us to identify significant indicators for categorizing species as forest or savanna
437 specialists, or generalists (= no significant indicators, Table S3). We further applied this classification of
438 our 1,707 tree species on an independent dataset of species occurrence derived from georeferenced
439 herbarium records (24, 25). We computed the number (and percentage) of forest specialists, generalists,
440 and savanna specialists per pixel of the 0.5° resolution grid representing our study area (Fig. 2).

441 We verified the concordance between the potential biomes derived from climate, and the species
442 composition in terms of percentage of forest specialists, generalists, and savanna specialists (Fig. S5).

443 We used Kruskal-Wallis tests to compare each pair of potential biomes (*forest* and *bistable forest*,
444 *bistable savanna* and savanna), and specifically test whether the *bistable* region is more composed of
445 generalist species, or is a mix of forest and savanna specialist species.

446

447 *(iii) Spatial distribution of the biome index.* The spatial distributions of our forest specialists, generalists,
448 and savanna specialists were then integrated into a biome index at the scale of 0.5° pixel (Fig. 3), with the
449 following formula:

$$450 \text{ biome index} = \frac{\#sp \text{ for} - \#sp \text{ sav}}{\#sp \text{ for} + \#sp \text{ sav} + \#sp \text{ gen}}$$

451 where *#sp for* is the number of forest specialists, *#sp sav* the number of savanna specialists, and *#sp gen*
452 the number of generalist species in each pixel. The biome index tracks the specialization of each pixel,
453 with values toward -1 representing the dominance of savanna specialists, and values toward +1
454 representing the dominance of forest specialists.

455 To further examine the underlying composition of the biome index, and to specifically test whether
456 the intermediate values of the biome index rather correspond to more generalist species or to a mix of
457 forest and savanna specialists, we used spline regressions between the percentage of each of the three
458 species groups and the biome index (Fig. S6). Since the frequency distribution of the biome index was
459 found to be slightly tri-modal (inset, Fig. 3), we mapped the three states: the savanna state, the
460 intermediate state, and the forest state, using different thresholds on the biome index (Fig. S7). Finally, to
461 examine the effect of herbarium record availability, we constructed the map of the biome index for pixels
462 containing at least five or ten of our species and showed the associated frequency distribution (Fig. S8).

463

464 *(iv) Spatial and environmental predictions of the biome index.* Because species occurrences were missing
465 in remote and/or undersampled areas (Fig. 3), we re-computed our biome index (Fig. 4A) from spatially
466 interpolated values of the percentage of forest specialists, generalists and savanna specialists using
467 ordinary kriging (Fig. S9A-C). The spherical autocorrelation function provided a good fit to the
468 experimental variograms (inset, Fig. 4A). We evaluated the spatial accuracy of the kriged map outside the
469 pixels containing herbarium records of our study species by computing confidence intervals from the
470 variance of the predictions (Fig. S9D-F).

471 We also used a random forest approach to identify the environmental determinants of the biome
472 index, and to map the biome index according to climate (described by PC1 and PC2, Fig. 1A, S1), fire
473 (Fig. S3), herbivory, and soil factors (described by the percentage of sand and CEC in the top 0-5 cm).
474 For the random forest, our dataset corresponded to the pixels containing at least 5 of our 1,707 species
475 and with available data for all environmental determinants. We kept 10% of the dataset for validation
476 (n=169 pixels) and fit the random forest model on a calibration dataset (n=1,523 pixels). We also
477 examined the importance of each environmental determinant, which measures how the accuracy of the
478 results is affected if the input variable is randomly permuted.

479

480 All the analyses were performed within the open source R environment (51) using the *raster* package (52)
481 for raster constructions and most spatial analyses. The *ade4* package (53) was used for the PCA and the
482 *dismo* package (54) for the convex hull. The *vegan* package (55) was used for the dissimilarity analysis
483 and the *labdsv* package (56) for computing species indicator values (and significance) for each biome.
484 The *gstat* package (57) was used for the spatial interpolation by ordinary kriging, while the *RandomForest*
485 package (58) for the Random Forest regression used to identify the environmental determinants and to
486 provide environmental predictions.

487

488 **Acknowledgments:** We warmly thank all the members of the Tropical Forestry group in Gembloux Agro-
489 Bio Tech, University of Liège, not listed as co-authors of the study. This work was funded by the Belgian
490 Science Policy (BR/143/A3/HERBAXYLAREDD, A.F) and the Natural Environment Research Council
491 (NE/P008755/1, C.M.R. and K.G.D.). D.B. received support from Fondation Wiener-Anspach.

492

493 **References:**

- 494 1. J. Ratnam, *et al.*, When is a 'forest' a savanna, and why does it matter? *Glob. Ecol. Biogeogr.* **20**,
495 653–660 (2011).
- 496 2. C. E. Lehmann, S. A. Archibald, W. A. Hoffmann, W. J. Bond, Deciphering the distribution of the
497 savanna biome. *New Phytol.* **191**, 197–209 (2011).
- 498 3. A. C. Staver, S. Archibald, S. A. Levin, The global extent and determinants of savanna and forest as
499 alternative biome states. *Science* **334**, 230–232 (2011).
- 500 4. P. L. Silva de Miranda, *et al.*, Using tree species inventories to map biomes and assess their
501 climatic overlaps in lowland tropical South America. *Glob. Ecol. Biogeogr.* **27**, 899–912 (2018).
- 502 5. M. Hirota, M. Holmgren, E. H. V. Nes, M. Scheffer, Global Resilience of Tropical Forest and
503 Savanna to Critical Transitions. *Science* **334**, 232–235 (2011).
- 504 6. A. C. Staver, S. Archibald, S. Levin, Tree cover in sub-Saharan Africa: Rainfall and fire constrain
505 forest and savanna as alternative stable states. *Ecology* **92**, 1063–1072 (2011).
- 506 7. C. Favier, *et al.*, Abrupt shifts in African savanna tree cover along a climatic gradient. *Glob. Ecol.*
507 *Biogeogr.* **21**, 787–797 (2012).
- 508 8. G. P. Hempson, S. Archibald, J. E. Donaldson, C. E. Lehmann, Alternate Grassy Ecosystem States
509 Are Determined by Palatability–Flammability Trade-Offs. *Trends Ecol. Evol.* **34**, 286–290 (2019).
- 510 9. W. A. Hoffmann, *et al.*, Ecological thresholds at the savanna-forest boundary: how plant traits,
511 resources and fire govern the distribution of tropical biomes. *Ecol. Lett.* **15**, 759–768 (2012).
- 512 10. T. Charles-Dominique, A. C. Staver, G. F. Midgley, W. J. Bond, Functional differentiation of biomes
513 in an African savanna/forest mosaic. *South Afr. J. Bot.* **101**, 82–90 (2015).
- 514 11. M. D. Swaine, J. B. Hall, J. M. Lock, The forest-savanna boundary in west-central Ghana. *Ghana J.*
515 *Sci.* **16**, 35–52 (1976).

- 516 12. J. C. Aleman, A. C. Staver, Spatial patterns in the global distributions of savanna and forest. *Glob. Ecol. Biogeogr.* **27**, 792–803 (2018).
517
- 518 13. C. Xu, *et al.*, Remotely sensed canopy height reveals three pantropical ecosystem states. *Ecology* **97**, 2518–2521 (2016).
519
- 520 14. V. de L. Dantas, M. Hirota, R. S. Oliveira, J. G. Pausas, Disturbance maintains alternative biome states. *Ecol. Lett.* **19**, 12–19 (2016).
521
- 522 15. M. C. Hansen, *et al.*, Towards an operational MODIS continuous field of percent tree cover algorithm: examples using AVHRR and MODIS data. *Remote Sens. Environ.* **83**, 303–319 (2002).
523
- 524 16. M. C. Hansen, *et al.*, High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* **342**, 850–853 (2013).
525
- 526 17. J. C. Aleman, O. Blarquez, C. A. Staver, Land-use change outweighs projected effects of changing rainfall on tree cover in sub-Saharan Africa. *Glob. Change Biol.* **22**, 3013–3025 (2016).
527
- 528 18. M. Torello-Raventos, *et al.*, On the delineation of tropical vegetation types with an emphasis on forest/savanna transitions. *Plant Ecol. Divers.* **6**, 101–137 (2013).
529
- 530 19. J. W. Veldman, Clarifying the confusion: old-growth savannahs and tropical ecosystem degradation. *Philos. Trans. R. Soc. B Biol. Sci.* **371**, 20150306 (2016).
531
- 532 20. C. P. Osborne, *et al.*, Human impacts in African savannas are mediated by plant functional traits. *New Phytol.* (2018) <https://doi.org/10.1111/nph.15236> (June 1, 2018).
533
- 534 21. H. P. Linder, *et al.*, The partitioning of Africa: statistically defined biogeographical regions in sub-Saharan Africa. *J. Biogeogr.* **39**, 1189–1205 (2012).
535
- 536 22. H. P. Linder, The evolution of African plant diversity. *Front. Ecol. Evol.* **2**, 38 (2014).
537
- 538 23. G. P. Hempson, S. Archibald, W. J. Bond, A continent-wide assessment of the form and intensity of large mammal herbivory in Africa. *Science* **350**, 1056–1061 (2015).
539
- 540 24. V. Droissart, *et al.*, Beyond trees: Biogeographical regionalization of tropical Africa. *J. Biogeogr.* **45**, 1153–1167 (2018).
541
- 542 25. M. S. Sosef, *et al.*, Exploring the floristic diversity of tropical Africa. *BMC Biol.* **15**, 15 (2017).
543
- 544 26. A. Fayolle, *et al.*, Patterns of tree species composition across tropical African forests. *J. Biogeogr.* **41**, 2320–2331 (2014).
545
- 546 27. A. Fayolle, *et al.*, A sharp floristic discontinuity revealed by the biogeographic regionalization of African savannas. *J. Biogeogr.* **46**, 454–465 (2019).
547
- 548 28. F. White, *The vegetation of Africa: A descriptive memoir to accompany the UNESCO/AETFAT/UNSO vegetation map of Africa* (ORSTOM - UNESCO, 1983).
549
- 550 29. A. Vincens, *et al.*, Forest response to climate changes in Atlantic Equatorial Africa during the last 4000 years BP and inheritance on the modern landscapes. *J. Biogeogr.* **26**, 879–885 (1999).
551
- 550 30. M. Sankaran, *et al.*, Determinants of woody cover in African savannas. *Nature* **438**, 846–849 (2005).
551

- 552 31. U. Salzmann, P. Hoelzmann, The Dahomey Gap: an abrupt climatically induced rain forest
553 fragmentation in West Africa during the late Holocene. *The Holocene* **15**, 190–199 (2005).
- 554 32. J. C. Aleman, *et al.*, Palaeo-trajectories of forest savannization in the southern Congo. *Biol. Lett.* **15**,
555 20190284 (2019).
- 556 33. U. Salzmann, Are modern savannas degraded forests?-A Holocene pollen record from the
557 Sudanian vegetation zone of NE Nigeria. *Veg. Hist. Archaeobotany* **9**, 1–15 (2000).
- 558 34. R. R. Klopper, L. Gautier, C. Chatelain, G. F. Smith, R. Spichiger, Floristics of the angiosperm flora
559 of sub-Saharan Africa: an analysis of the African Plant Checklist and Database. *Taxon* **56**, 201–208
560 (2007).
- 561 35. J. C. Aleman, M. A. Jarzyna, A. C. Staver, Forest extent and deforestation in tropical Africa since
562 1900. *Nat. Ecol. Evol.* **2**, 26–33 (2018).
- 563 36. A. Vincens, Y. Garcin, G. Buchet, Influence of rainfall seasonality on African lowland vegetation
564 during the Late Quaternary: pollen evidence from Lake Masoko, Tanzania. *J. Biogeogr.* **34**, 1274–
565 1288 (2007).
- 566 37. L. Laestadius, *et al.*, Opportunities for forest landscape restoration. *Unasylva* **62**, 238 (2011).
- 567 38. M. Greve, B. Reyers, A. M. Lykke, J.-C. Svenning, Spatial optimization of carbon-stocking projects
568 across Africa integrating stocking potential with co-benefits and feasibility. *Nat. Commun.* **4**, 2975
569 (2013).
- 570 39. R. C. Abreu, *et al.*, The biodiversity cost of carbon sequestration in tropical savanna. *Sci. Adv.* **3**,
571 e1701284 (2017).
- 572 40. B. P. Murphy, A. N. Andersen, C. L. Parr, The underestimated biodiversity of tropical grassy
573 biomes. *Philos. Trans. R. Soc. B Biol. Sci.* **371**, 20150319 (2016).
- 574 41. C. M. Ryan, *et al.*, Ecosystem services from southern African woodlands and their future under
575 global change. *Philos. Trans. R. Soc. B Biol. Sci.* **371**, 20150312 (2016).
- 576 42. N. Stevens, C. E. Lehmann, B. P. Murphy, G. Durigan, Savanna woody encroachment is
577 widespread across three continents. *Glob. Change Biol.* **23**, 235–244 (2017).
- 578 43. E. Buisson, *et al.*, Resilience and restoration of tropical and subtropical grasslands, savannas, and
579 grassy woodlands. *Biol. Rev.* **94**, 590–609 (2019).
- 580 44. A. T. Kaptué, L. Prihodko, N. P. Hanan, On greening and degradation in Sahelian watersheds.
581 *Proc. Natl. Acad. Sci.* **112**, 12133–12138 (2015).
- 582 45. L. Gillson, R. Marchant, From myopia to clarity: sharpening the focus of ecosystem management
583 through the lens of palaeoecology. *Trends Ecol. Evol.* **29**, 317–325 (2014).
- 584 46. S. E. Fick, R. J. Hijmans, WorldClim 2: new 1-km spatial resolution climate surfaces for global land
585 areas. *Int. J. Climatol.* **37**, 4302–4315 (2017).
- 586 47. L. Giglio, J. Descloitres, C. O. Justice, Y. J. Kaufman, An enhanced contextual fire detection
587 algorithm for MODIS. *Remote Sens. Environ.* **87**, 273–282 (2003).
- 588 48. G. P. Hempson, S. Archibald, W. J. Bond, The consequences of replacing wildlife with livestock in
589 Africa. *Sci. Rep.* **7**, 1–10 (2017).

- 590 49. T. Hengl, *et al.*, Mapping soil properties of Africa at 250 m resolution: Random forests significantly
591 improve current predictions. *PloS One* **10** (2015).
- 592 50. M. Dufrière, P. Legendre, Species assemblages and indicator species: the need for a flexible
593 asymmetrical approach. *Ecol. Monogr.* **67**, 345–366 (1997).
- 594 51. R Development Core Team, R: A language and environment for statistical computing. (2018).
- 595 52. R. J. Hijmans, *raster: Geographic Data Analysis and Modeling.* (2017).
- 596 53. S. Dray, A. B. Dufour, The ade4 package: implementing the duality diagram for ecologists. *J. Stat.*
597 *Softw.* **22**, 1–20 (2007).
- 598 54. R. J. Hijmans, S. Phillips, J. Leathwick, J. Elith, *dismo: Species Distribution Modeling* (2017).
- 599 55. J. Oksanen, *et al.*, *vegan: Community Ecology Package* (2011).
- 600 56. D. W. Roberts, *labdsv: Ordination and multivariate analysis for ecology* (2010).
- 601 57. E. J. Pebesma, Multivariable geostatistics in {S}: the gstat package. *Comput. Geosci.* **30**, 683–691
602 (2004).
- 603 58. A. Liaw, M. Wiener, Classification and regression by randomForest. *R News* **2**, 18–22 (2002).
- 604

605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641

Figure Legends:

Figure 1. Climatic envelope of the forest and savanna biomes in Africa.

To determine the climatic envelopes of the two tropical biomes, the dominant climatic gradients were first identified with a Principal Component Analysis of gridded climatic variables. (A) Each point corresponds to the center of a 0.5° pixel, with pixels containing floristic surveys in forest (green) and savanna (orange) sites indicated. Red and blue arrows indicate the influence of temperature and precipitation variables, respectively. (B) Distribution of forest (green) and savanna (orange) sites along an axis of mean annual precipitation (MAP, in mm). (C) Frequency distribution of forest (green) and savanna (orange) sites along a complex precipitation and seasonality gradient (PC1, A), with the climatic area where savanna and forest both occur plotted in light orange (where savanna is more common) and light green (where forest is more common). The dashed line corresponds to an equal probability of savanna and forest. (D) Map of forest and *bistable* forest, *bistable* savanna and savanna, with the locations of floristic surveys in forest (green) and savanna (orange) sites. White pixels are outside the geographic extent and/or the environmental range covered by the floristic surveys. The latter was defined by a convex hull on the site scores on PC1 and PC2 (dashed line in A). Major waterbodies and rivers are shown in blue. The paleo-ecological data available around the Gulf of Guinea retrieved from published records of lacustrine fossil pollen (n=24), lacustrine phytoliths (n=1), phytoliths (n=4) and $\delta^{13}\text{C}$ of soil organic matter (n=14) from soil profiles are shown in inset. The paleo-sites provided information on stable (black) and transitional (grey) sites during the Holocene (see Table S2 for the code correspondence of paleo-sites).

Figure 2. Distribution of forest specialist, generalist, and savanna specialist tree species.

To test for specialization toward the forest and savanna biomes by individual tree species, we applied the *IndVal* procedure (50) to the presence matrix of the 1,707 species in the 753 sites (455 forest and 298 savanna sites), and obtained a classification of species into forest specialists (n=825 species) and savanna specialists (n=523), with non-significant indicators resulting in species being interpreted as generalists (n=359). For each 0.5° pixel containing herbarium records for at least 5 of our 1,707 species, we computed and mapped the percentage of (A) forest specialists, (B) generalists and (C) savanna specialists. White pixels thus correspond to a paucity of georeferenced herbarium records for our classified tree species. Major waterbodies are shown in blue. The frequency distribution of the number of sites in which (A) forest specialists, (B) generalists, and (C) savanna specialists occurred in the original floristic surveys is given on the bottom of each panel, illustrating the higher frequency of extremely infrequent species in the group of generalist species, in comparison to the specialists.

Figure 3. Spatial distribution of the biome index across Africa.

The distributions of forest specialists, generalists, and savanna specialists derived from georeferenced herbarium records were used to devise a biome index based on tree species composition and computed

642 at 0.5° resolution. The biome index tracks the biome specialization of each pixel, with values toward -1
643 representing the dominance of savanna specialists and values toward +1 representing the dominance of
644 forest specialists. White pixels correspond to a lack of georeferenced herbarium records for our species.
645 Major waterbodies are shown in blue. The frequency distribution of the biome index is given in inset.
646

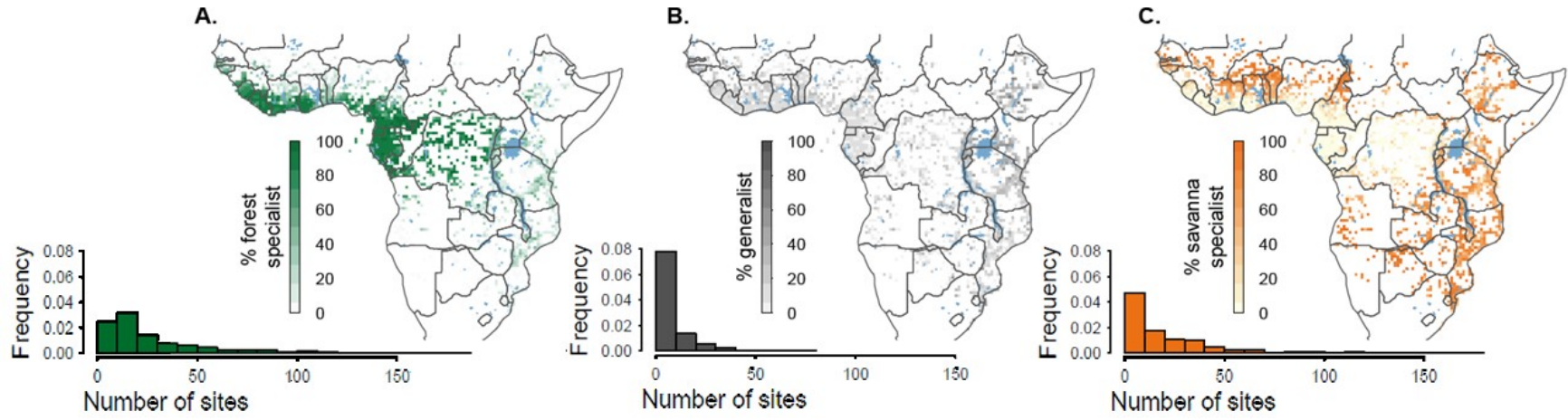
647 **Figure 4. Spatial and environmental predictions of the biome index.**

648 We predicted the distribution of the biome index across the climatic space covered by the floristic surveys
649 using (A) spatial information only vs (B) a random forest approach based on environmental determinants.
650 For the spatial predictions, we interpolated the distribution of the percentage of forest specialists,
651 generalists and savanna specialists (Fig. S8) using ordinary kriging, and recomputed the biome index.
652 The empirical semi-variogram (points) and the spherical semi-variogram model (lines) used for kriging
653 and shown as inset in (A) indicate that the spatial structure of generalist species is weaker and more
654 homogeneous than that of forest and savanna specialists. For the environmental predictions of the biome
655 index, we used climate (described by PC1 and PC2, Fig. 1A, S1), fire (Fig. S3), herbivory (total biomass
656 of both livestock and wild herbivores), and soil factors (described by the percentage of sand and CEC in
657 the first 0-5 cm). The relative importance of each environmental determinant is shown as inset in (B). The
658 importance (% increase in MSE) tests how the accuracy of the results is affected if the input variable is
659 randomly permuted. White pixels in (A) correspond to areas outside the geographical extent and the
660 environmental range covered by the floristic surveys. White pixels in (B) additionally contain pixels for
661 which herbivory data were not available (coastal edge and edge of inland waterbodies). Major
662 waterbodies are shown in blue.

663
664

669
670

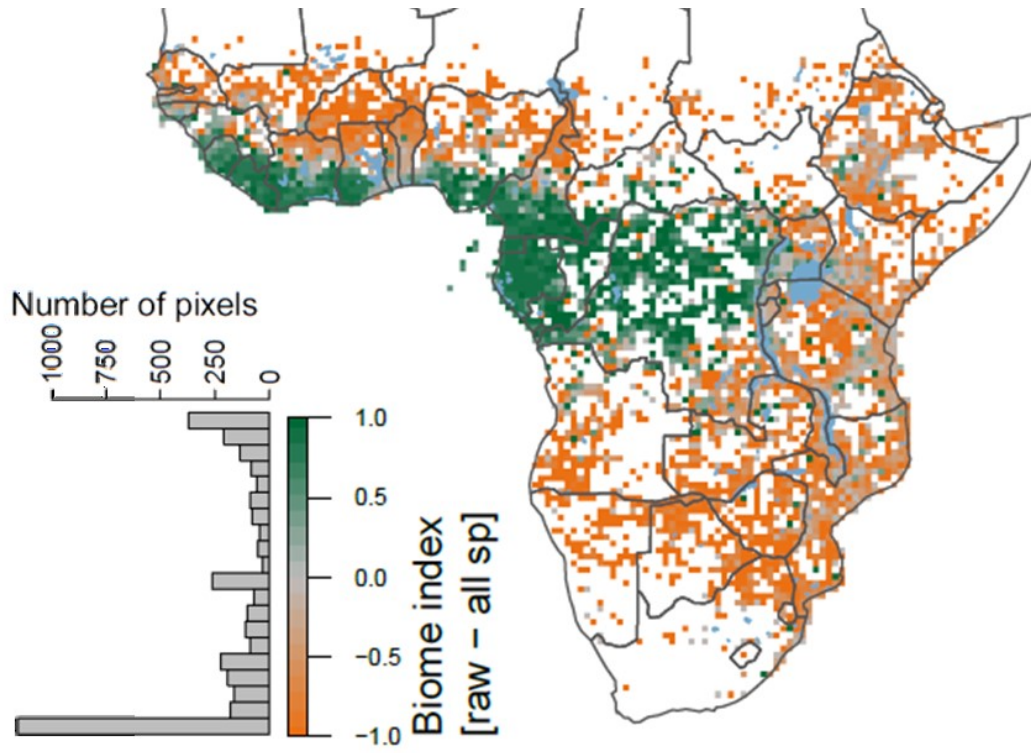
Figure 2



671
672

673
674

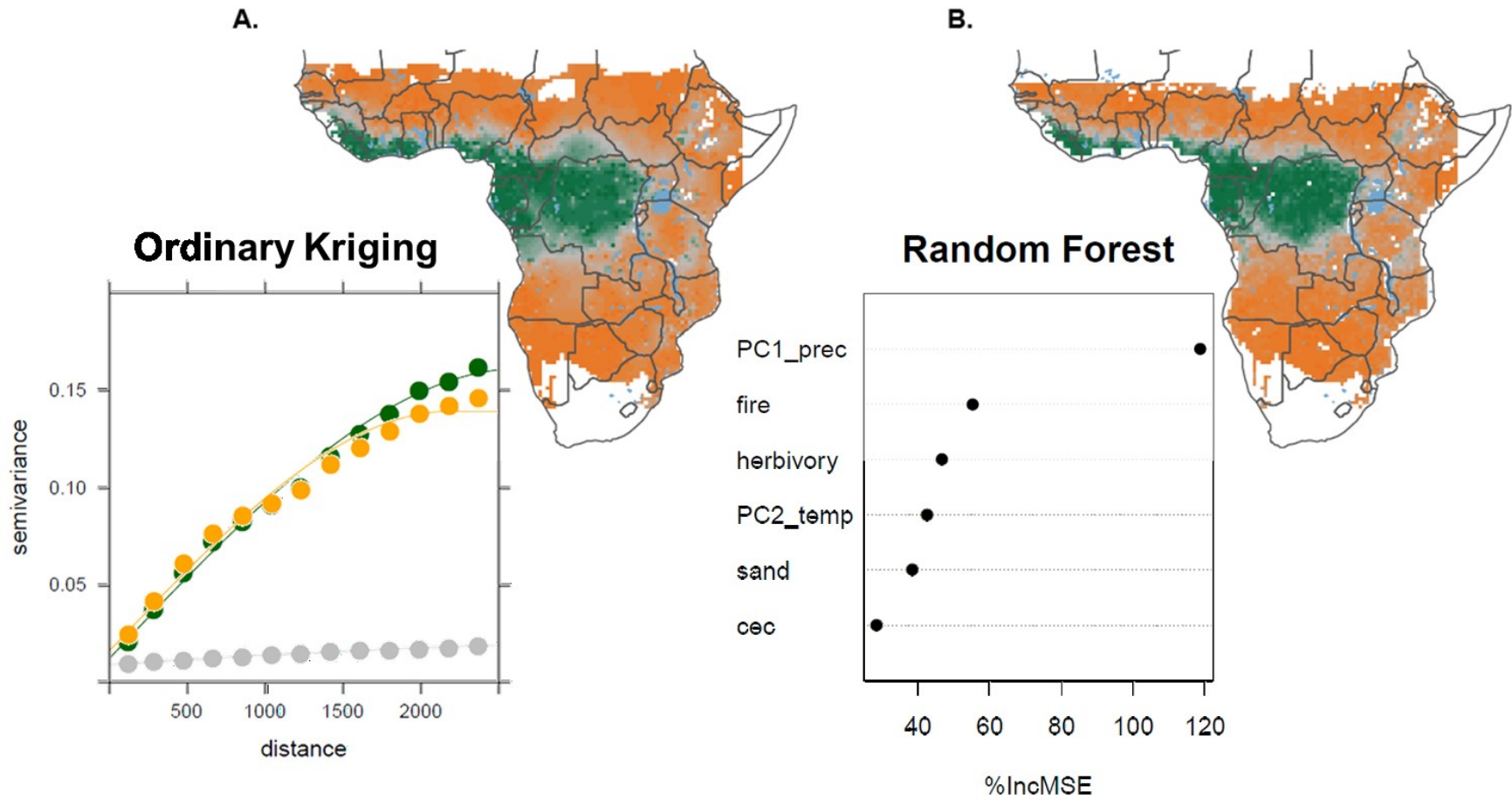
Figure 3



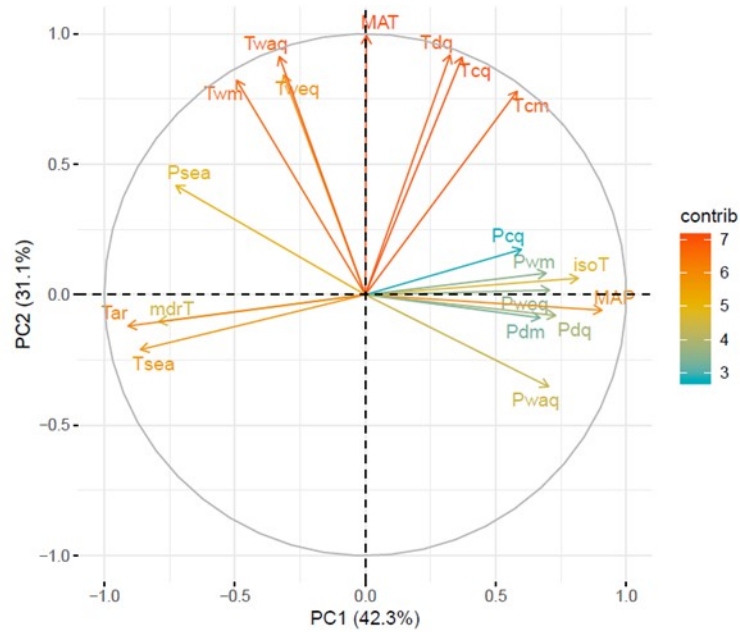
675
676

677
678

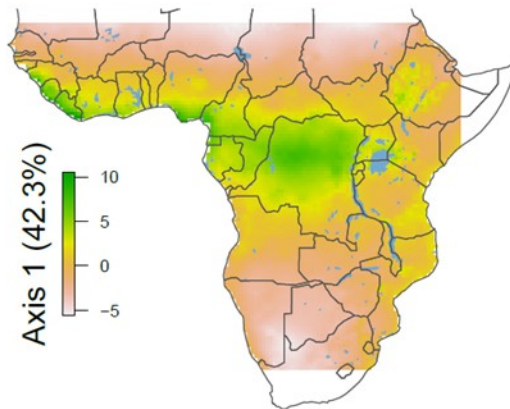
Figure 4



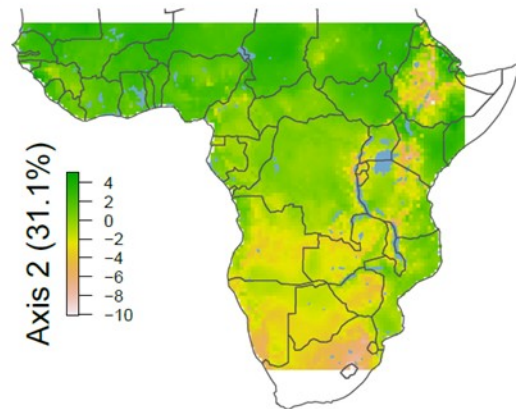
679
680
681



B.



C.

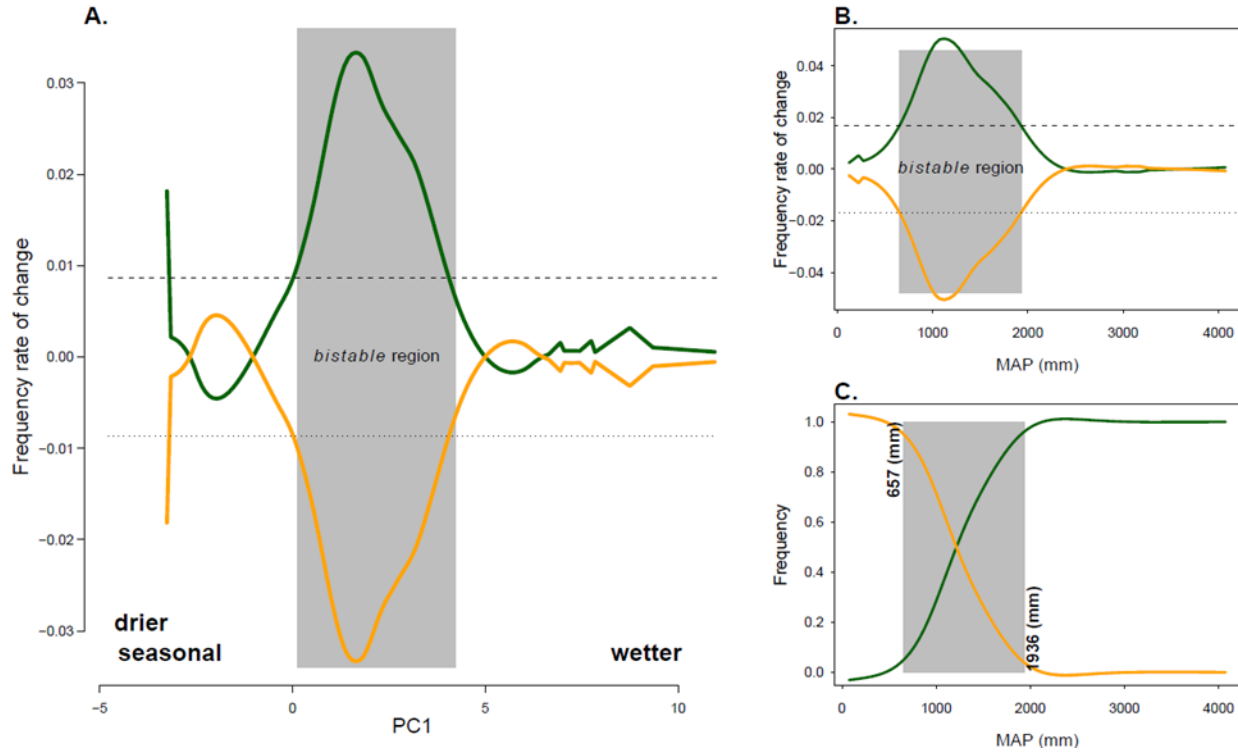


682

683 **Figure S1. Climatic gradients.**

684 (A) The correlation circle of bioclimatic variables (see Materials and methods, Datasets, (ii) Environmental
 685 drivers, for the code correspondence) from the PCA performed on the matrix covering the entire study
 686 area (Fig. 1A) highlights the contribution of precipitation and temperature variables to PC1 and PC2 (color
 687 gradient). The other panels represent the map of (B) PC1 and (C) PC2 scores over the entire study area.
 688 White pixels in the maps are thus outside the geographic extent covered by the floristic surveys. Major
 689 waterbodies are shown in blue.

690

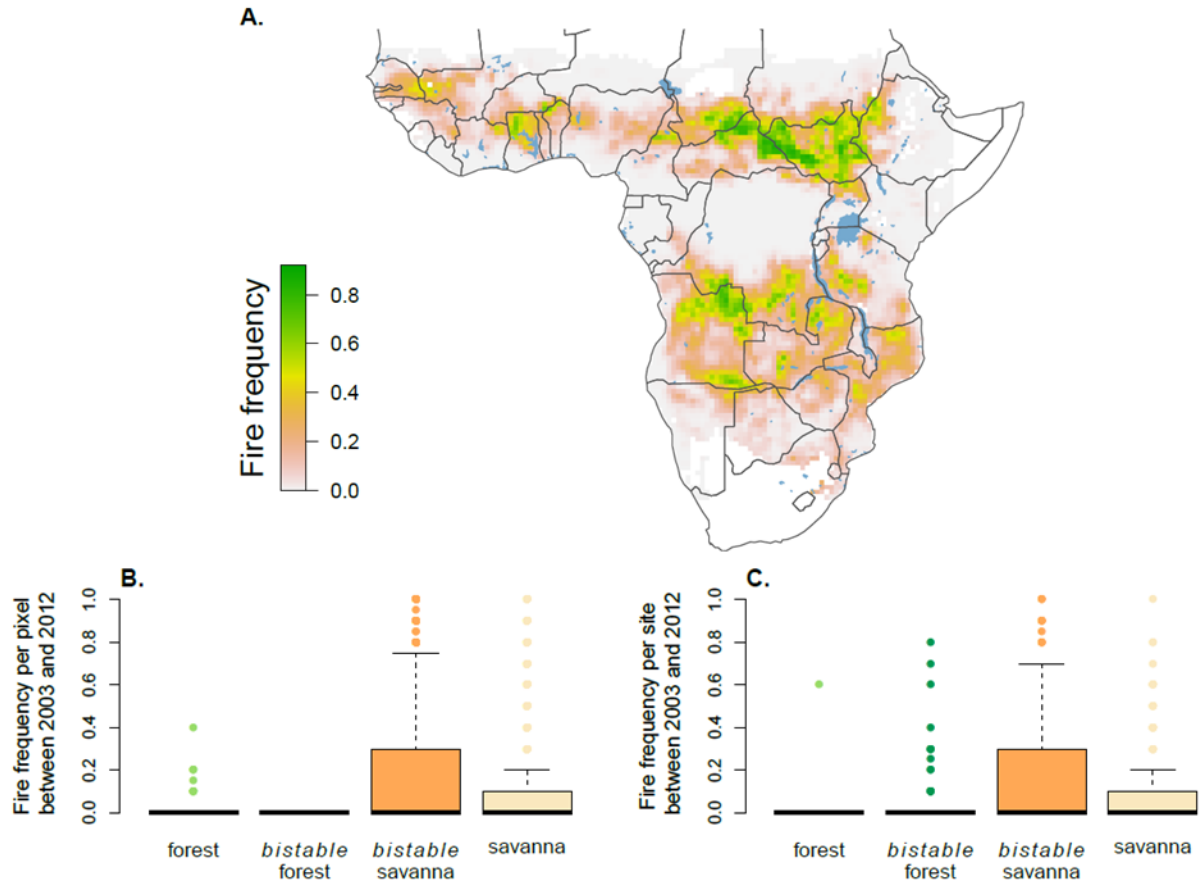


691

692 **Figure S2. Climatic overlap.**

693 The putative *bistable* region where forest and savanna may represent alternative stable states was
 694 defined as the region of climatic overlap between forest and savanna sites. Slopes of both forest (green)
 695 and savanna (orange) frequency distributions are shown with respect to PC1, the complex climatic
 696 gradient integrating precipitation and seasonality (Fig. 1A, S1). The climate space for which forest and
 697 savanna are both widespread (0.141 – 4.235 on PC1) is represented in grey, this is the *bistable* region
 698 (Fig. 1C). The black lines correspond to the mean of the rate of change for each distribution (dashed for
 699 forest, and dotted for savanna) over the entire PC1 axis. The same analysis was conducted along an axis
 700 of (B) Mean Annual Precipitation (MAP, in mm) for comparison with other studies and leads to (C) a MAP
 701 range of circa 700-1900 mm for the *bistable* region.

702

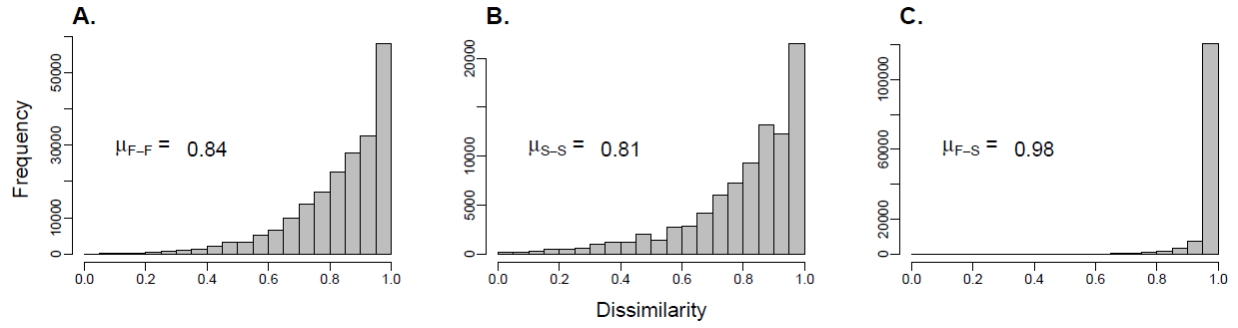


703

704 **Figure S3. Fire frequency.**

705 (A) A map of fire frequency over 2003-2012 was derived from MODIS Active Fire burned area. White
 706 pixels in the map are outside the geographic extent and/or the environmental range covered by the
 707 floristic surveys. Major waterbodies are shown in blue. Fire frequency was examined according to the
 708 potential biomes, *forest*, *bistable forest*, *bistable savanna*, and savanna as defined in Fig. 1C-D, (B) for
 709 the entire study area, and (C) for the pixels containing floristic surveys.

710



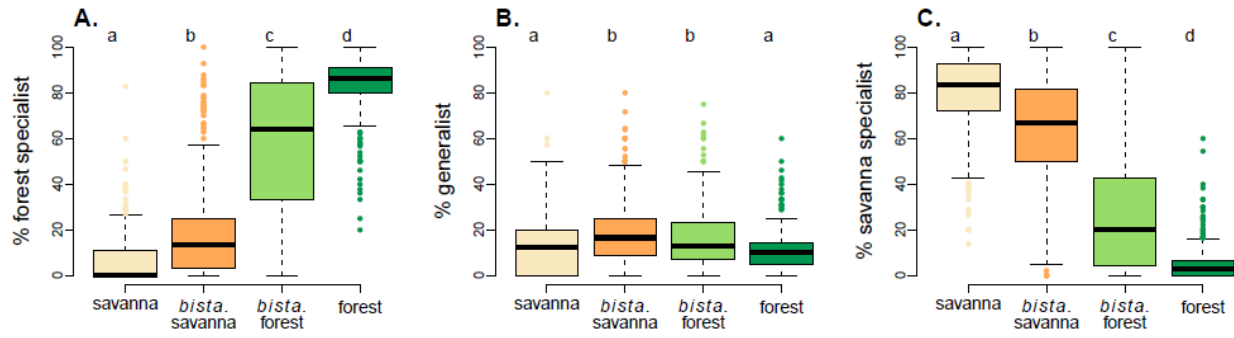
711

712 **Figure S4. Floristic dissimilarity.**

713 The Simpson index of beta-diversity was computed between all pairs of sites. The frequency distribution
 714 of floristic dissimilarity (A) within forest sites, (B) within savanna sites, and (C) between forest and
 715 savanna sites shows a marked floristic dissimilarity between forest and savanna sites in comparison to
 716 that observed within forest sites, and within savanna sites. The mean of all pairwise dissimilarities is
 717 shown for each comparison, (A) within forest sites, (B) within savanna sites, and (C) between forest and
 718 savanna sites.

719

720

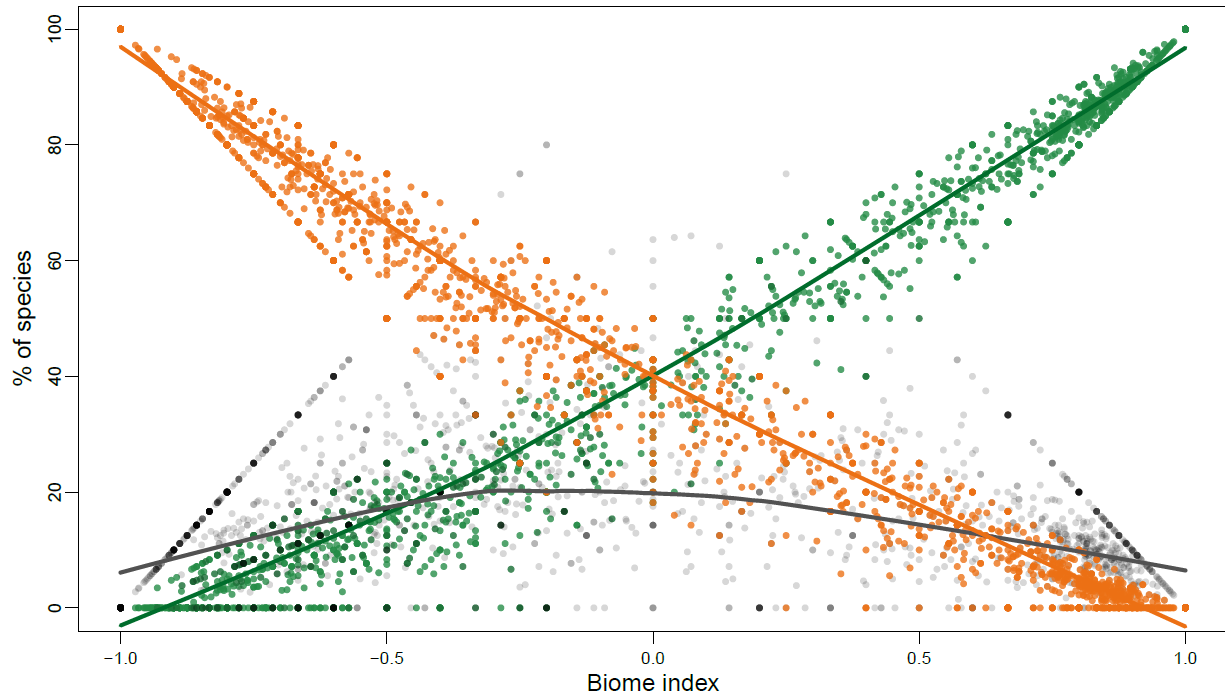


721
722

723 **Figure S5. Composition of the *bistable* region.**

724 Relationships between potential biomes derived from climate and floristic composition derived from
 725 georeferenced herbarium records were examined. The boxplots depict for each potential biome (Fig. 1)
 726 the percentage of (A) forest specialist species, (B) generalist species, and (C) savanna specialist species,
 727 for pixels containing at least 5 of our 1,707 species. The letters (a, b, c and d) in these panels show the
 728 results of Kruskal-Wallis tests applied for each pair of potential biomes, with different letters indicating a
 729 significant statistical difference at $p < 0.05$.

730
731

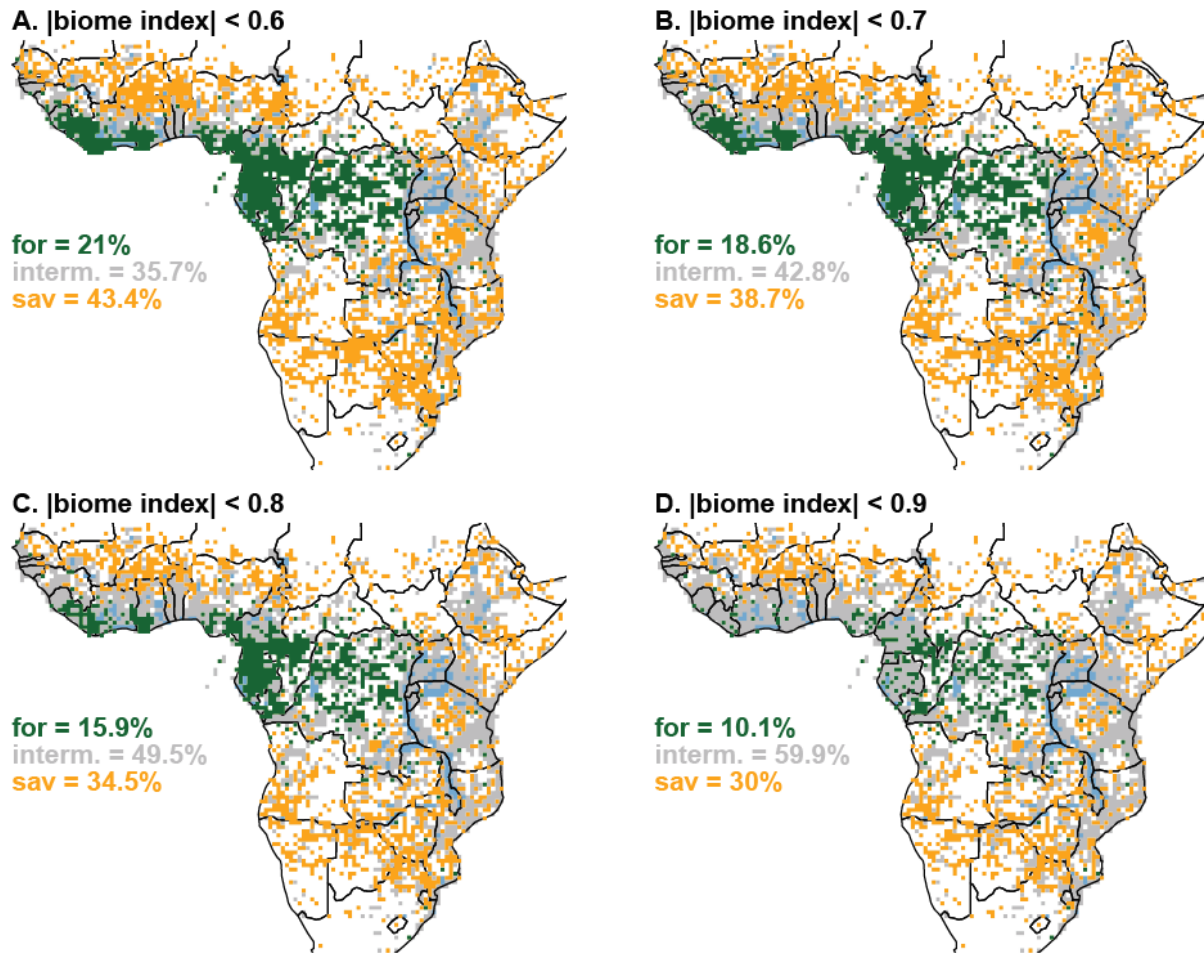


732

733 **Figure S6. Composition of the biome index.**

734 The underlying composition in terms of percentage of forest specialist (green), generalist (grey) and
 735 savanna specialist (orange) species was examined for the biome index computed with distribution data
 736 derived from georeferenced herbarium records. The thick lines represent a spline regression of the data
 737 points, illustrating a continuous increase/decrease of the percentage of forest /savanna specialists with
 738 the biome index, and a relatively stable and low (< 20) percentage of generalist species across the whole
 739 range of biome index, that, however, tended to show a slight peak for intermediate biome index.

740

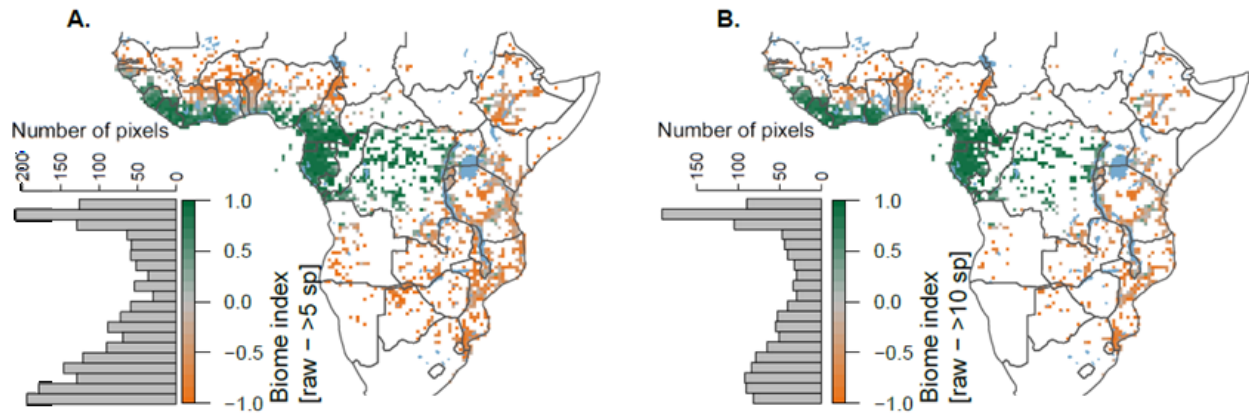


741
742

743 **Figure S7. Intermediate biome index.**

744 Since the frequency distribution of the biome index was found to be slightly tri-modal (inset, Fig. 3), we
 745 used different threshold to map the intermediate state: (A) |biome index| < 0.6, (B) |biome index| < 0.7,
 746 (C) |biome index| < 0.8 and (D) |biome index| < 0.9. The percentage of pixels belonging to the forest state
 747 (in green), to the intermediate state (in grey), and to the savanna state (in orange), is given for each
 748 threshold. Blue pixels are dominated by waterbodies. In West and Central Africa, intermediate values of
 749 the biome index correspond to long recognized transition zones between major floristic regions (28), to
 750 savanna areas such as the Dahomey gap that were forested during the Holocene (31), and where forest-
 751 savanna mosaics are common, and to recently deforested areas (35). In East Africa, intermediate values
 752 of the biome index also correspond to the high elevation forest of Ethiopia, much of which are not
 753 included in our species lists, and to densely-populated areas, such as the coastal forest and the Lake
 754 Victoria mosaic (28).

755
756



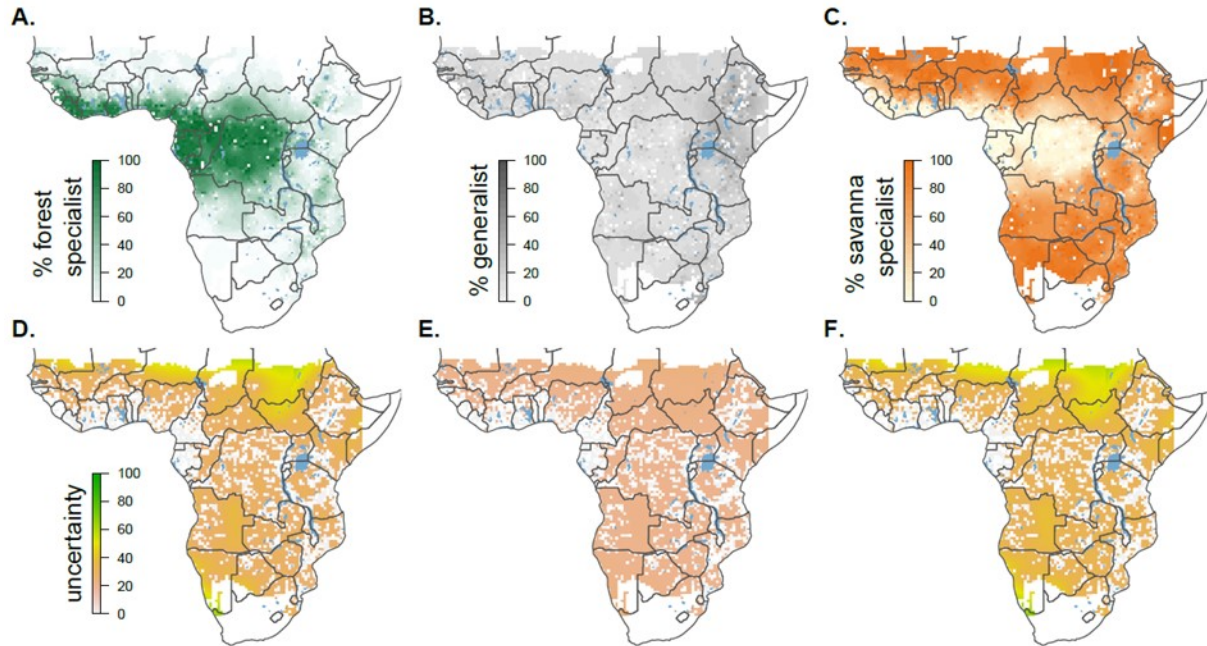
757

758 **Figure S8. Biome index and availability of herbarium records.**

759 The distribution of forest specialists, generalists, and savanna specialists derived from georeferenced
 760 herbarium records was used to devise a biome index based on tree species composition and computed
 761 at 0.5° resolution, and here, pixels containing at least (A) five and (B) ten of our 1,707 tree species are
 762 solely shown. White pixels correspond to a lack of georeferenced herbarium records for our species.
 763 Major waterbodies are shown in blue. The frequency distributions of the biome index given in inset
 764 demonstrated that the tri-modal signal observed with all pixels (Fig. 3) gradually disappears under
 765 increased sampling (A and B), giving only a savanna (-1) and a forest (+1) mode.

766

767



768

769 **Figure S9. Interpolated distributions and associated uncertainty.**

770 The percentage of (A) forest specialist, (B) generalist and (C) savanna specialist species was spatially
 771 interpolated at 0.5° resolution using ordinary kriging. Associated uncertainty was also computed for (D)
 772 forest specialist, (E) generalist and of (F) savanna specialist species, using the confidence interval of the
 773 spatial predictions. White pixels in the maps (A-C) are outside the geographical extent and/or the
 774 environmental range covered by the floristic surveys. White pixels in the maps (D-F) also include the
 775 training data (pixels containing georeferenced herbarium records for at least 5 of our 1,707 species) for
 776 which there is no uncertainty. Major waterbodies are shown in blue.

777

778 **Table S1. Floristic surveys.**

779 List of 455 forest sites and of 298 savanna sites with floristic information on tree species.

780

781 **Table S2. Paleo-ecological evidence.**

782 List of 42 paleo-sites for which at least 2,000 years of vegetation history is available.

783

784 **Table S3. Indicator species.**

785 List of forest specialist, generalist and savanna specialist species among the 1,707 tree species present

786 in the 455 forest and 298 savanna sites and retained for analyses.