Unveiling African rainforest composition and vulnerability to global change

M. Réjou-Méchain, Frédéric Mortier, Jean-François Bastin, Guillaume Cornu, Nicolas Barbier, Nicolas Bayol, Fabrice Bénédet, Xavier Bry, Gilles Dauby, Vincent Deblauwe, et al.

► To cite this version:

M. Réjou-Méchain, Frédéric Mortier, Jean-François Bastin, Guillaume Cornu, Nicolas Barbier, et al.. Unveiling African rainforest composition and vulnerability to global change. Nature, 2021, 593, pp.90-94. 10.1038/s41586-021-03483-6 . hal-03205277

HAL Id: hal-03205277
https://hal.inrae.fr/hal-03205277v1

Title: Unveiling African rainforest composition and vulnerability to global change

Authors: Maxime Réjou-Méchain[1]*, Frédéric Mortier[2,3], Jean-François Bastin[1,2,3,4], Guillaume Cornu[2,3], Nicolas Barbier[1], Nicolas Bayol[5], Fabrice Bénédet[2,3], Xavier Bry[6], Gilles Dauby[1], Vincent Deblauwe[7,8], Jean-Louis Doucet[9], Charles Doumenge[2,3], Adeline Fayolle[9], Claude Garcia[2,3,10], Jean-Paul Kibambe[11,12], Jean-Joël Loumeto[13], Alfred Ngomanda[14], Pierre Ploton[1], Bonaventure Sonké[15], Catherine Trottier[6], Ruppert Vimal[16], Olga Yongo[17], Raphaël Pélissier[1], Sylvie Gourlet-Fleury[2,3]

Affiliations:

1. AMAP, Univ Montpellier, IRD, CNRS, CIRAD, INRA, Montpellier, France.
2. CIRAD, Forêts et Sociétés, F-34398 Montpellier, France.
3. Forêts et Sociétés, Univ Montpellier, CIRAD, Montpellier, France.
4. Crowther Lab, Department of Environmental Systems Science, Institute of Integrative Biology, ETH Zürich, Zürich, Switzerland.
5. Foret Ressources Management Ingénierie, 34130 Mauguio – Grand Montpellier, France.
6. Institut Montpelliérain Alexander Grothendieck (IMAG), Montpellier,France.
7. Center for Tropical Research, Institute of the Environment and Sustainability, University of California, Los Angeles, LosAngeles, CA 90095, USA.
8. International Institute of Tropical Agriculture, Yaounde, Cameroun.
9. Université de Liège, Gembloux Agro-Bio Tech, Forest is Life, TERRA Teaching and Research Centre, Passage des Déportés 2, 5030 Gembloux, Belgium.
10. Forest Management and Development Team, Department of Environmental System Science, Institute of Integrative Biology, ETH Zürich, Zürich, Switzerland.
11. Faculté des Sciences Agronomiques, Département de Gestion des Ressources Naturelles, Université de Kinshasa, Kinshasa I, Democratic Republic of Congo.
12. Wildlife Conservation Society, Kinshasa, Congo, Democratic Republic of Congo.

13. Faculté des Sciences et Techniques, Université Marien Ngouabi, BP 69 Brazzaville, Republic of Congo.
14. Institut de Recherche en Ecologie Tropicale (IRET/CENAREST), B.P. 13354 Libreville, Gabon.
15. Plant Systematic and Ecology Laboratory, Higher Teacher’s Training College, University of Yaoundé I, Yaoundé, Cameroon.
16. GEODE UMR 5602, CNRS, Université Jean-Jaurès, 5 Allée Antonio-Machado, Toulouse, France.
17. Laboratoire de Biodiversité Végétale et Fongique, Faculté des Sciences, Université de Bangui, Bangui, Central African Republic.

Africa is forecasted to experience large and rapid climate change[1] and population growth[2] during the XXIst century, threatening the world's second largest rainforest. Protecting and sustainably managing these forests requires an extended understanding of their current compositional heterogeneity, environmental drivers and vulnerability to ongoing changes. Here, using an unprecedented dataset of 6 million trees in more than 180,000 field plots, we jointly model the distribution in abundance of the most dominant central African tree taxa and produce the first continuous maps of the floristic and functional composition of central African forests. Our results show that the uncertainty in taxon-specific distributions averages out at the community level, revealing highly deterministic assemblages. We uncover contrasting floristic and functional compositions across climate, soil types and anthropogenic gradients, with functional convergence among floristically dissimilar forest types. Combining these spatial predictions with global change scenarios suggests a high vulnerability of the northern and southern forest margins, the Atlantic forests and of most forests from the Democratic Republic of Congo where both climate and anthropogenic threats are expected to increase sharply by 2085. These results constitute key quantitative benchmarks for scientists and policy makers to shape transnational conservation and management strategies aiming at providing a sustainable future for central African forests.

Concomitant increases in climate stress, human population needs and resource extraction in central Africa raise environmental concerns[3–5]. These threats may have considerable impacts on the carbon budget[6], climate[7] and biodiversity of central African forests[8], which shelter some of the world's most iconic wildlife species and that are already experiencing much drier and seasonal climate than other tropical forests[9]. However, the current composition of central African forests and its determinants at regional scale are still poorly known, being often studied on limited areas[10–12] and datasets[13] or at a very coarse grain with heterogeneous presence-only data[14]. Vast areas remain poorly scientifically explored[15] while most spaceborne systems of Earth observation provide very limited information on forest composition[16]. This limits our ability to understand how forest composition and functions vary regionally, to forecast how these forests will face upcoming global changes and, ultimately, to anticipate, on scientific bases, how to protect and manage them beyond national boundaries.

In this study, we used an unprecedented forest inventory dataset to (1) model the main floristic and functional gradients over central African forests, and (2) assess their expected vulnerability under forecasted global (climatic and anthropogenic) change conditions. We compiled the abundance distributions of 193 dominant tree taxa in 185,665 field plots (ca. $90,000$ ha) from commercial forest inventories spread over the five main forested countries in central Africa (Extended Data Fig. 1). We modeled the joint distributions of taxon abundances at a 10-km resolution using supervised component generalized linear regression (SCGLR)[17], a modelling method that extends partial least squares (PLS) regression to the multivariate generalized linear framework. SCGLR models a set of responses (here the abundances of taxa) from synthetic orthogonal explanatory components derived from 24 climatic variables (hereafter, climatic components, CCs) and additional soil type (here, sand vs. clay) and anthropogenic pressure covariates. We developed for this study an index based on population density and road network specifically designed and calibrated to predict recent human-induced forest disturbance intensity in central Africa - see Methods. Finally, thanks to the huge size of the dataset, the predicted floristic and functional gradients were cross-validated with spatially independent observations using Spearman correlation coefficients, ρ_{CV}.

Floristic composition in central Africa

Our model predicted individual taxon abundances with an overall median correlation ρ_{CV} of 0.48 (range of -0.11 to 0.83). This median was still as high as 0.45 when unoccupied sites were removed, showing that, beyond presence-absence, our model also captured variations in abundances within taxon’s distributional range. A correspondence analysis (CA) performed on the predicted taxon abundances revealed major regional floristic gradients (Fig. 1; Extended Data Fig. 2 and 3) highly correlated with the observed gradients (ρ_{CV}=0.89, 0.71 and 0.6 for CA axes 1 to 3, respectively; Fig. 1B-D). Contrary to Amazonian and Southeast Asian forests, where soil was shown to be the primary large scale driver of tree community composition[18,19], the most prominent floristic gradient predicted here (CA axis 1) was highly related to climate, in particular to the first predictive CC (Pearson's r=0.98), contrasting areas with a cool and light-deficient[20] dry season (coastal Gabon) and areas with high evapotranspiration rates (northern limit of the central African forests; Extended Data Fig. 4). The second predicted floristic gradient (CA axis 2) was highly correlated with the two other CCs (r=-0.86 and -0.72 for CC2 and CC3, respectively) related to seasonality and maximum temperature, thus contrasting equatorial areas with a low water deficit and areas with a high water deficit towards the limits of the tropics. Climate seasonality was also found to be an important driver of tree community composition in Amazonia[18] and maximum temperature has been recently identified as the most important pantropical driver of forest biomass, impacting woody productivity[21]. The third predicted floristic gradient (CA axis 3) revealed more local floristic variations highlighting human-impacted forests (r=0.67 with our index of human-induced forest disturbance intensity).

As already shown in previous studies[22,23], the association between taxon distributions and climate patterns may appear by chance because both are spatially autocorrelated at the regional scale. We thus used a spatially explicit null model that randomized the predictive CCs while preserving their spatial (co)structures. When keeping the soil and human impact on forests unchanged, the null model did not predict significantly different abundances (P>0.1) from those predicted with original CCs for 67% of the taxa. This suggests that variation in taxon abundances directly depends on climate for a minimum of only one-third of the taxa, while for most of them, abundance may correlate with climate by chance only. In contrast, the association between climate and the main gradients of floristic assemblages was robust to autocorrelation artifacts (P=0.028, 0.006 and 0.06 for CA1 to 3, respectively). This result confirms that extrapolating assemblages from climate variables is more reliable than extrapolating individual taxon abundances[24]. Indeed, individual taxon abundances are likely less predictable on the basis of only current drivers since they are also affected by unknown past human disturbances[25], biotic interactions and biogeographical history[26], the idiosyncratic effects of which tend to average out at the community level.

Functional composition in central Africa

From the predicted floristic assemblages, we computed the community weighted mean[27] of three functional traits known to play an important role in ecosystem processes: wood density, deciduousness and maximum diameter (Fig. 2). The predicted functional composition was consistent with the observations (ρ_{CV}=0.47, 0.75 and 0.45 for the three traits, respectively; Extended Data Fig. 5). As in Amazonia[18], community wood density varies with soil type with the highest values found on sandy soils, at the boundaries of Cameroon, Republic of Congo and Central African Republic, where tree species with conservative resource use strategies predominate[11]. However, larger scale variations in wood density were primarily driven by human-induced forest disturbances, with a lower community wood density in human-impacted forests, indicating that they are mostly composed of fast-growing taxa[28]. However, these areas also feature a high proportion of trees that can potentially reach a large diameter. These two results indicate that human-impacted forests are dominated by long-lived pioneer taxa, which are characterized by a low wood density but a large potential stature, offering a fast and relatively long-lasting carbon sink potential in absence of disturbances[29]. Finally, a marked deciduousness gradient ran from the highly evergreen forests of coastal Gabon to the northern limit of the central African forests with, again, a well-known exception on the northern sandy soil plateau[11,30].

A reference map of forest types

To ease practical applications, we performed hierarchical clustering of the predicted floristic gradients (pixel scores on the first five CA axes), which are continuous by nature, and identified ten major forest types (Fig. 3; Extended Data Table 1). The strongest floristic dissimilarity appeared between Atlantic forests (types 1 to 3) and the other forest types (4 to 10), within which semideciduous seasonal forests were clearly distinguished (types 4 to 6). We also observed functional convergences among floristically dissimilar forest types and vice versa. For example, despite having a regional species pool similar to deciduous forests (types 4 and 6), sandstone forests (type 5) have a functional composition closer to remote forest groups (e.g. types 2, 3, 7 and 8), with high wood density and low deciduousness. Soil filtering indeed modifies the relative abundance of species (rather than their presence or absence[31]) favoring suitable functional attributes in poor sandy soils[11]. By contrast, while Atlantic forests (types 1 to 3) have little taxonomic affinity with the east-central and southern forests (types 7 and 8), they display similar functional composition due to more similar climate conditions, as represented on the first predictive CC (Extended Data Table 1). This confirms that while taxonomic composition has an important biogeographical component, the functional composition of tree communities can converge in similar environmental conditions.

Vulnerability to global change

For the ten forest types, most climate models predict current climate conditions either to virtually disappear from central Africa (e.g., types 2 and 5; Extended Data Fig. 6), or to move at spatial and temporal scales incommensurate with tree dispersal ability (e.g. types 4 and 6). This suggests that current forest communities will not be able to track their present climate envelopes and will face the emergence of novel climates, making the prediction of taxon distributions under future climate projection highly risky[32]. We thus assessed the vulnerability of central African forests to climate change through their sensitivity, exposure, and adaptive capacity, following the recommendation of the IPCC[33].

We quantified sensitivity at the community level using the inverse of the current climate niche breadth of taxa (Fig. 4C) and assuming that assemblages dominated by taxa with narrow environmental tolerances will be more vulnerable to upcoming changes[34]. Sensitivity appeared to be high in coastal Gabon (type 2), where a high level of species endemism exists[35] and in the driest northern margin of central African forests. Recent studies consistently showed that drier tropical forests exhibited larger functional changes than wetter forests in response to a long-term drought in west Africa[36] and are likely to be more sensitive to global warming[21]. By contrast, forests from the northwest Cameroon displayed a relatively low sensitivity to current climate conditions, probably because these forests are dominated by widespread tree taxa adapted to anthropogenic pressure (Fig. 2). Long-lived pioneers, typical of these human-impacted forests, are also expected to be favored by a possible acceleration in forest dynamics induced by global change[37,38].

Exposure to climate change was quantified as the extent to which the current climate determinants (CC1 to 3) are expected to change by 2085, using 18 unique bias-corrected climate model combinations (under the IPCC-AR5 RCP 4.5 scenario; see Extended Data Fig. 7 for other scenarios). We found that exposure to climate change was mostly driven by an increase in drought stress and maximum temperature (Fig. S2, see also[4,39]). The central and east part of central African forests are predicted to be the most exposed, particularly in the south of the Democratic Republic of Congo (DRC) (Fig. 4D). Note, however, that climate change predictions in central Africa are uncertain because meteorological data for model validation are lacking[4] (Fig. S3).

Finally, we assessed the adaptive capacity of tree communities through their evolutionary potential. We first found highly significant niche conservatism along the first two climate components ($P<0.002$). This indicates that closely related taxa tend to share similar climate niche spaces at the regional scale and suggests that they could be impacted similarly by future climate change. We thus assumed that higher local phylogenetic diversity provides a wider range of potential responses to novel climate conditions[40], similarly to the insurance hypothesis[41]. We thus used the phylogenetical diversity of predicted tree assemblages as a proxy of their adaptive capacity to climate change. Undisturbed semideciduous and transitional forests (types 6 and 10 in Fig. 3) appeared phylogenetically more diverse, thus having higher adaptive capacity, than evergreen forests (Fig. 4E). A recent study in Amazonia[42] also found a peak of phylogenetic diversity at intermediate precipitation level, where dry- and wet-adapted lineages are mixing. As expected[43], we also found that human-impacted areas tended to have a low phylogenetic diversity and hence a lower adaptive capacity to climate change.

The resulting vulnerability of tree communities to climate change did not correlate with the expected human impact on forests in 2085 (ρ= -0.08), here assessed using country-specific projections of human population (Fig. 4A; Extended Data Fig. 8). Vulnerability to climate change is expected to be higher for communities dominated by hard-wooded taxa (ρ=0.46 with wood density, Table S1). By contrast, forecasted human impact on forests is predicted to be higher in already disturbed communities, i.e., dominated by light-wooded taxa with a large potential size (ρ=-0.4 and 0.43, respectively). However, because we did not account for the appearance of new roads by 2085, we may underestimate the effect of future anthropogenic activities in remote, currently undisturbed forests. Vulnerability to both climate change and anthropogenic activities (pink color in Fig. 4A) is predicted to be high for forests from coastal Gabon, in large areas from DRC and in the northern margin of the forest domain. Forests from Cameroon and in the south of the Republic of Congo mostly appear vulnerable due to the high expected human impact on forests by 2085 (orange patches in Fig. 4A). By contrast, the tri-national Sangha transboundary forest complex and the northeastern part of Gabon appeared as the least vulnerable area in the region (the large green patch in Fig. 4A). Globally, DRC, which comprises most of the central African forests, mainly contains forests that are predicted to be vulnerable to climate change and/or to anthropogenic pressure (blues to pink patches in Fig. 4A).

Conclusions and perspectives

While some country-specific vegetation patterns were already suggested by phytogeographers, here we provide the first synoptic view of central African forest composition at a fine resolution, based on an unprecedented amount of quantitative data. Unveiling the functional composition of central African forests conveys important insights on their functioning, dynamics and carbon uptake potential and on the way they could respond to global change. Accounting for forest functional characteristics can also considerably reduce uncertainty in large-scale vegetation models[44] or improve remote sensing approaches, for example, by assimilating large-scale variation in wood density into forest carbon maps[45]. Our maps may also help scientists to design representative sampling to better understand the long-term impact of climate change on tree species and stand dynamics, e.g., monitoring under-represented forest types or areas highly vulnerable to climate change.

The forest types and vulnerability maps should guide the development of new land use plans that preserve the full range of evolutionary and functional potential of today's forests or, at least, that maintain their connectivity to attenuate the threats related to expected changes. In central Africa, protected areas and logging concessions, which cover almost half of the forest domain (14.9% and 32.2%, respectively; Extended Data Fig. 9), are important to consider in such plans. Protected areas do not equally cover the ten identified forest types (4 to 54%; Extended Data Table 1) and should therefore be extended to reach a better representativity. How estimated vulnerability should be accounted for when designing protected areas, e.g., by extending the network in vulnerable areas to minimize biodiversity loss, or in areas with low anthropogenic pressure to improve their protection, is subject to debate[46]. Logging concessions can also contribute to the maintenance of forest cover and functions, providing that they are well managed[47,48], and currently likely act like biodiversity corridors between protected areas[49]. However, this will only prove effective in the long term if they strictly comply with legislation and, ideally, with standard certification requirements. These good practices are especially important in forests dominated by evergreen taxa with high wood density, where disturbances may have a higher impact on community composition. In areas expected to be under high anthropogenic pressure, forest connectivity could be preserved by promoting agroforestry and restoration programs, strictly controlling access to logging roads and stabilizing shifting agriculture[50]. Over central Africa, the highest uncertainties for the future of forests remain in DRC, where substantial areas, belonging to the state, are not yet attributed to any land use category and should deserve particular attention due to their high vulnerability (Fig. 4).

Main references

1. Diffenbaugh, N. S. & Giorgi, F. Climate change hotspots in the CMIP5 global climate model ensemble. *Climatic change* **114**, 813–822 (2012).

2. United Nations, Department of Economic and & Social Affairs, Population Division. World Population Prospects: The 2017 Revision, Key Findings and Advance Tables. *Working Paper* **ESA**/**P**/**WP**/**248**, (2017).

3. Malhi, Y., Adu-Bredu, S., Asare, R. A., Lewis, S. L. & Mayaux, P. African rainforests: Past, present and future. *Philosophical Transactions of the Royal Society B: Biological Sciences* **368**, 20120312 (2013).

4. James, R., Washington, R. & Rowell, D. P. Implications of global warming for the climate of African rainforests. *Philosophical Transactions of the Royal Society B: Biological Sciences* **368**, 20120298 (2013).

5. Abernethy, K., Maisels, F. & White, L. J. Environmental issues in central Africa. *Annual Review of Environment and Resources* **41**, 1–33 (2016).

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

7. De Wasseige, C., Tadoum, M., Atyi, E. & Doumenge, C. *The forests of the Congo Basin-Forests and climate change*. (Weyrich, 2015).

8. Stévart, T. *et al.* A third of the tropical African flora is potentially threatened with extinction. *Science advances* **5**, eaax9444 (2019).

9. Parmentier, I. *et al.* The odd man out? Might climate explain the lower tree alpha-diversity of African rain forests relative to Amazonian rain forests? *Journal of Ecology* **95**, 1058–1071 (2007).

10. Réjou-Méchain, M. *et al.* Regional variation in tropical forest tree species composition in the Central African Republic: An assessment based on inventories by forest companies. *Journal of Tropical Ecology* **24**, 663–674 (2008).

11. Réjou-Méchain, M. *et al.* Tropical tree assembly depends on the interactions between successional and soil filtering processes. *Global Ecology and Biogeography* **23**, 1440–1449 (2014).

12. Fayolle, A. *et al.* Geological substrates shape tree species and trait distributions in african moist forests. *PLoS ONE* **7**, e42381 (2012).

13. Fayolle, A. *et al.* Patterns of tree species composition across tropical African forests. *Journal of Biogeography* **41**, 2320–2331 (2014).

14. Droissart, V. *et al.* Beyond trees: Biogeographical regionalization of tropical Africa. *Journal of biogeography* **45**, 1153–1167 (2018).

15. Sosef, M. S. *et al.* Exploring the floristic diversity of tropical Africa. *BMC biology* **15**, 15 (2017).

16. Parmentier, I. *et al.* Predicting alpha diversity of African rain forests: Models based on climate and satellite-derived data do not perform better than a purely spatial model. *Journal of Biogeography* **38**, 1164–1176 (2011).

17. Bry, X., Trottier, C., Verron, T. & Mortier, F. Supervised component generalized linear regression using a pls-extension of the fisher scoring algorithm. *Journal of Multivariate Analysis* **119**, 47–60 (2013).

18. Steege, H. ter *et al.* Continental-scale patterns of canopy tree composition and function across Amazonia. *Nature* **443**, 444–447 (2006).

19. Slik, J. F. *et al.* Soils on exposed Sunda Shelf shaped biogeographic patterns in the equatorial forests of Southeast Asia. *Proceedings of the National Academy of Sciences* **108**, 12343–12347 (2011).

20. Philippon, N. *et al.* The light-deficient climates of western Central African evergreen forests. *Environmental Research Letters* **14**, 034007 (2019).

21. Sullivan, M. J. P. *et al.* Long-term thermal sensitivity of Earth's tropical forests. *Science* **368**, 869–874 (2020).

22. Beale, C. M., Lennon, J. J. & Gimona, A. Opening the climate envelope reveals no macroscale associations with climate in European birds. *Proceedings of the National Academy of Sciences* **105**, 14908–14912 (2008).

23. Deblauwe, V. *et al.* Remotely sensed temperature and precipitation data improve species distribution modelling in the tropics. *Global Ecology and Biogeography* **25**, 443–454 (2016).

24. Maguire, K. C. *et al.* Controlled comparison of species-and community-level models across novel climates and communities. *Proc. R. Soc. B* **283**, 20152817 (2016).

25. Morin-Rivat, J. *et al.* Present-day central African forest is a legacy of the 19th century human history. *Elife* **6**, (2017).

26. Ricklefs, R. E. Intrinsic dynamics of the regional community. *Ecology Letters* **18**, 497–503 (2015).

27. Violle, C. *et al.* Let the concept of trait be functional! *Oikos* **116**, 882–892 (2007).

28. Díaz, S. *et al.* The global spectrum of plant form and function. *Nature* **529**, 167–171 (2016).

29. Rüger, N. *et al.* Demographic trade-offs predict tropical forest dynamics. *Science* **368**, 165–168 (2020).

30. Ouédraogo, D.-Y. *et al.* The determinants of tropical forest deciduousness: Disentangling the effects of rainfall and geology in central Africa. *Journal of Ecology* **104**, 924–935 (2016).

31. Shipley, B. *From plant traits to vegetation structure: Chance and selection in the assembly of ecological communities*. (Cambridge University Press, 2010).

32. Feeley, K. J. & Silman, M. R. Biotic attrition from tropical forests correcting for truncated temperature niches. *Global Change Biology* **16**, 1830–1836 (2010).

33. Parry, M. *et al. Climate change 2007-impacts, adaptation and vulnerability: Working group II contribution to the fourth assessment report of the IPCC*. **4**, (Cambridge University Press, 2007).

34. Foden, W. B. *et al.* Identifying the world's most climate change vulnerable species: A systematic trait-based assessment of all birds, amphibians and corals. *PloS one* **8**, (2013).

35. Lachenaud, O., Stévart, T., Ikabanga, D., Ndjabounda, E. C. N. & Walters, G. The littoral forests of the Libreville area (Gabon) and their importance for conservation: Description of a new endemic species (Rubiaceae). (2013). doi:info:doi/10.5091/plecevo.2013.744

36. Aguirre-Gutiérrez, J. *et al.* Drier tropical forests are susceptible to functional changes in response to a long-term drought. *Ecology Letters* **22**, 855–865 (2019).

37. Claeys, F. *et al.* Climate change would lead to a sharp acceleration of Central African forests dynamics by the end of the century. *Environmental Research Letters* **14**, 044002 (2019).

38. McDowell, N. G. *et al.* Pervasive shifts in forest dynamics in a changing world. *Science* **368**, (2020).

39. Zhou, L. *et al.* Widespread decline of Congo rainforest greenness in the past decade. *Nature* **509**, 86 (2014).

40. Purvis, A. Phylogenetic Approaches to the Study of Extinction. *Annual Review of Ecology, Evolution, and Systematics* **39**, 301–319 (2008).

41. Yachi, S. & Loreau, M. Biodiversity and ecosystem productivity in a fluctuating environment: The insurance hypothesis. *Proceedings of the National Academy of Sciences* **96**, 1463–1468 (1999).

42. Neves, D. M. *et al.* Evolutionary diversity in tropical tree communities peaks at intermediate precipitation. *Scientific reports* **10**, 1–7 (2020).

43. Letcher, S. G. Phylogenetic structure of angiosperm communities during tropical forest succession. *Proceedings of the Royal Society B: Biological Sciences* **277**, 97–104 (2010).

44. Fyllas, N. M., Quesada, C. A. & Lloyd, J. Deriving Plant Functional Types for Amazonian forests for use in vegetation dynamics models. *Perspectives in Plant Ecology, Evolution and Systematics* **14**, 97–110 (2012).

45. Mitchard, E. T. A. *et al.* Markedly divergent estimates of Amazon forest carbon density from ground plots and satellites. *Global Ecology and Biogeography* **23**, 935–946 (2014).

46. Visconti, P., Pressey, R. L., Bode, M. & Segan, D. B. Habitat vulnerability in conservation planning—when it matters and how much. *Conservation Letters* **3**, 404–414 (2010).

47. Putz, F. E. *et al.* Sustaining conservation values in selectively logged tropical forests: The attained and the attainable. *Conservation Letters* **5**, 296–303 (2012).

48. Gourlet-Fleury, S. *et al.* Tropical forest recovery from logging: A 24 year silvicultural experiment from Central Africa. *Phil. Trans. R. Soc. B* **368**, 20120302 (2013).

49. Clark, C. J., Poulsen, J. R., Malonga, R. & Jr, P. W. E. Logging Concessions Can Extend the Conservation Estate for Central African Tropical Forests. *Conservation Biology* **23**, 1281–1293 (2009).

50. Curtis, P. G., Slay, C. M., Harris, N. L., Tyukavina, A. & Hansen, M. C. Classifying drivers of global forest loss. *Science* 361, 1108–1111 (2018).

Figures

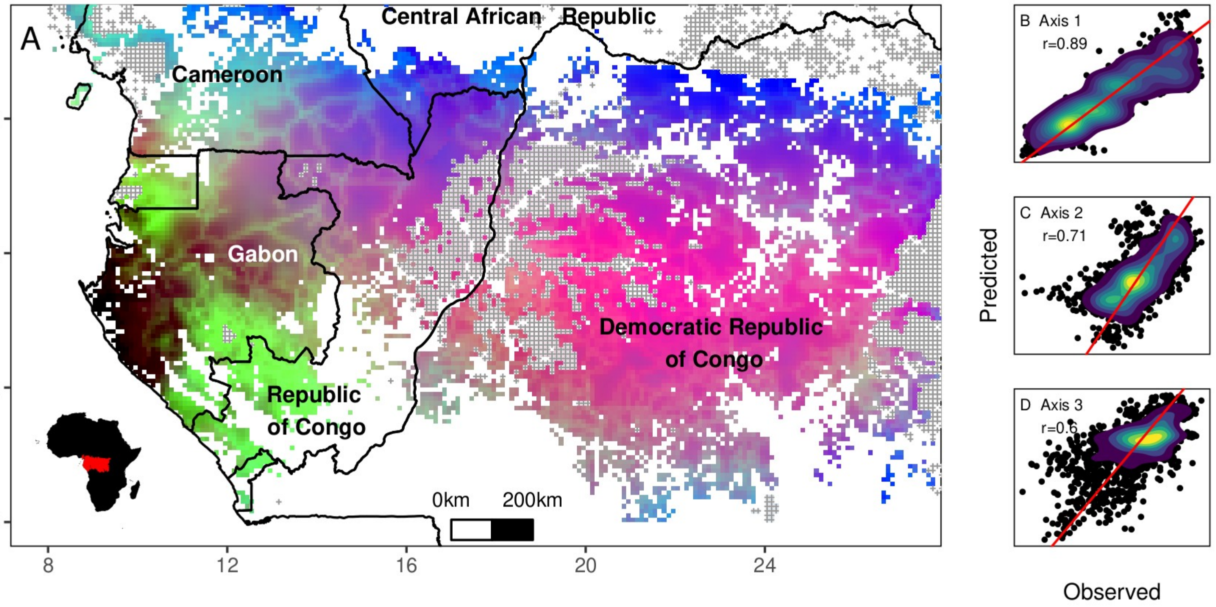


Figure 1: Floristic composition of central African forests. A, Spatialized RGB composition of the three first axes of a correspondence analysis (CA) performed on jointly predicted taxon abundances at 10-km resolution (n=193 taxa; axis 1: blue, axis 2: red, axis 3: green). Grey crosses represent forested areas outside the calibration domain, including permanently flooded forests and country boundaries are represented in black. Right panels **B-D** provide cross-validation results comparing the observed and predicted CA gradients (1:1 line in red). Taxon CA planes 1-2 and 1-3 are given in Extended Data Fig. 2.

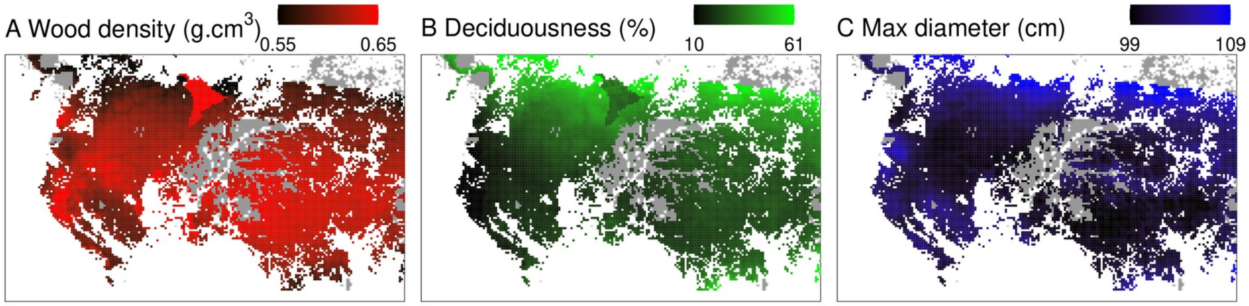


Figure 2: Predicted functional composition of central African forests. A-C, Predicted community weighted functional trait values at 10-km resolution.

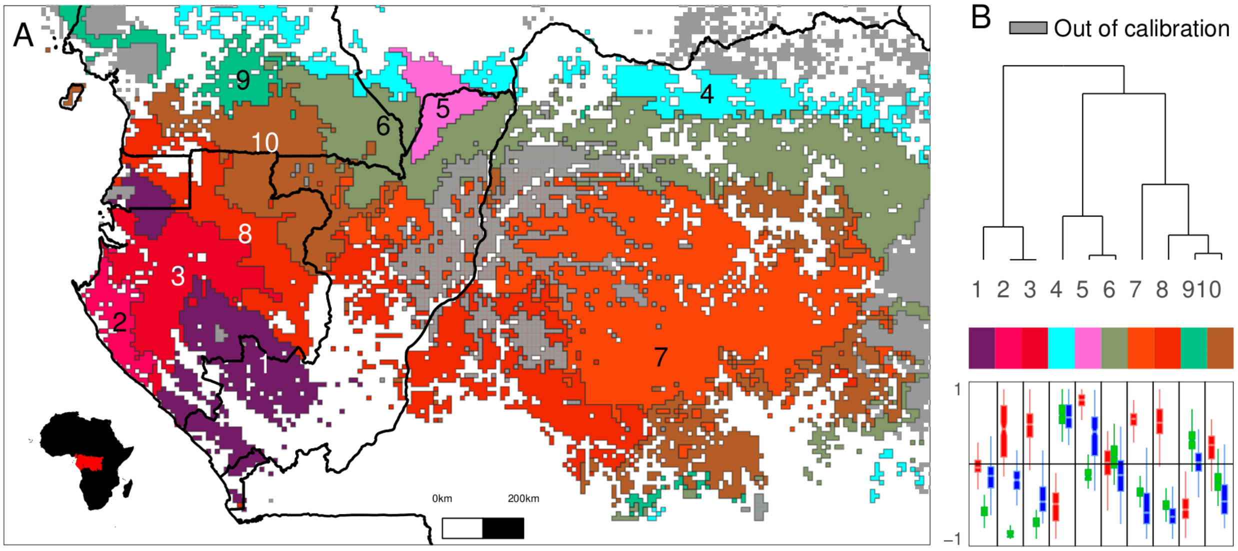


Figure 3: Main forest types across central Africa and their functional composition. **A,** Forest type classification obtained by hierarchical clustering of the predicted floristic gradients. Colors represent a RGB composite of the mean values of the three functional traits per forest type (see Fig. 2), viz. wood density (red), deciduousness (green) and maximum diameter (blue). Thus similar colors illustrate similar functional composition. **B**, Taxonomic relationships among the forest types illustrated by a clustering dendrogram (top) and a boxplot of the standardized predicted functional composition over the 12,295 grid cells (bottom), with wood density in red, deciduousness in green and maximum diameter in blue (median is reported at the center, the lower and upper hinges correspond to the first and third quartiles and the two whiskers extend from these two quartiles to the largest and smallest values, at most 1.5 times the inter-quartile range from the hinge). Forest type names and additional information are provided in Extended Data Table 1. Clustering uncertainty is reported in Fig. S1.

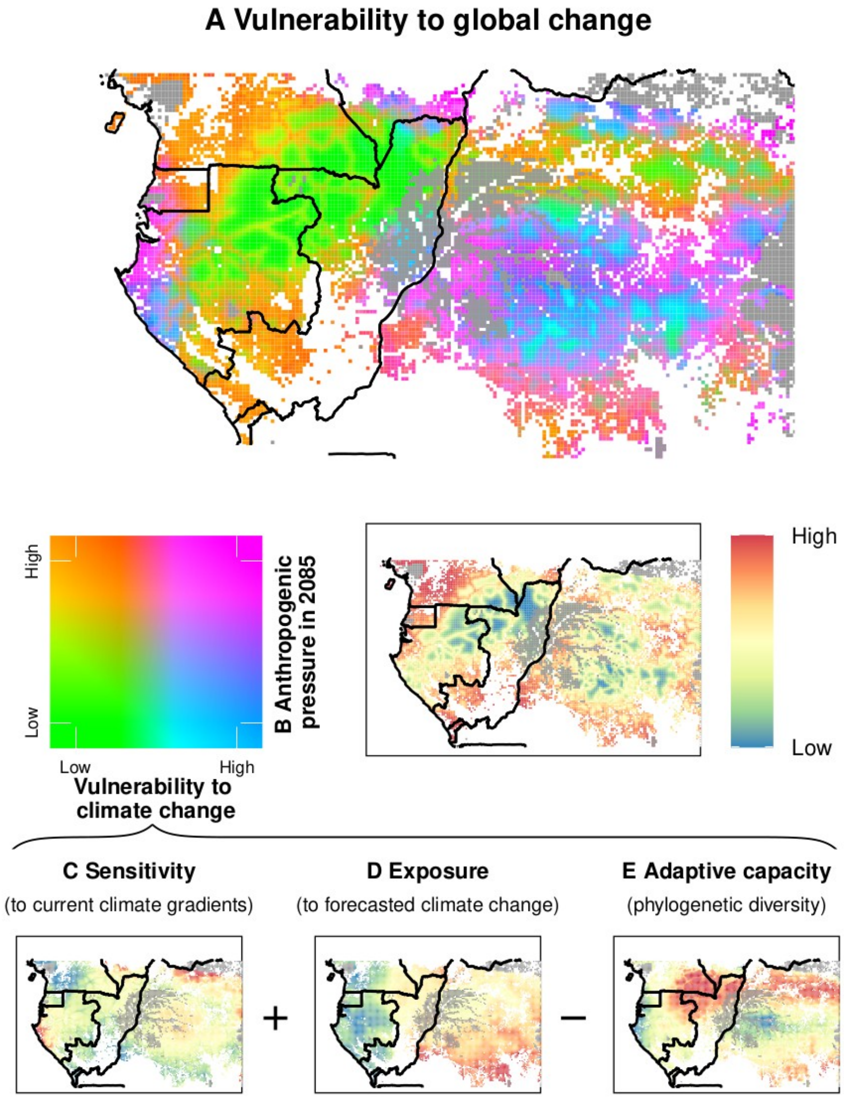


Figure 4: Predicted vulnerability of central African tree communities to global changes. A, Composite map of the vulnerability to climate change and of the forecasted human-induced forest disturbance intensity by 2085. **B,** Projected human-induced forest disturbance intensity in 2085. Vulnerability to climate change was estimated as the sensitivity to current climate (**C**) plus the exposure to forecasted climate changes by 2085 (under RCP scenario 4.5) (**D**) minus the adaptive capacity of tree communities using phylogenetic diversity as a proxy (**E**).

METHODS

Floristic and functional trait data

Forestry data were extracted from management forest inventories conducted in 105 logging concessions covering ca. 1.6×10^5 km^2 (Extended Data Fig. 1). Most companies followed a standardized inventory protocol similar to that described in Réjou-Méchain et al.[51]. In most cases, it consisted of continuous and parallel transects 20 m or 25 m wide, often 2-3 km apart, and subdivided into rectangular 0.4 or 0.5-ha plots. Overall, the full dataset had a total of 192,972 plots. Within each plot, trees with a diameter at breast height (DBH) ≥ 30 cm were allocated into 10-cm wide diameter classes and identified at the species or genus level whenever possible through either commercial or local names[51]. Independent analyses performed on 298 scientific plots (≥ 1 ha in size) showed that the floristic gradients obtained with large trees are representative of the ones obtained with trees ≥ 10 cm in diameter (Pearson r>0.94; Fig. S4). Overall, ca. 7×10^6 trees were recorded. Taxonomy was revised and homogenized using the African Flowering Plants Database (http://www.ville-ge.ch/musinfo/bd/cjb/africa/index.php?langue=an, last access on 01/09/2019) and the Angiosperm Phylogeny Group III for orders and families[52]. A total of 1,092 taxa were recorded in the original dataset. Extensive botanical controls demonstrated that the patterns of both intra (alpha)- and inter (beta)- plot diversity extracted from these data were highly reliable[51].

For the purpose of the present paper, we conducted an additional assessment according to botanical experts and by comparing the distributional range of our taxa with that in other datasets[53,54] to select a set of species and genera deemed to be reliably identified over the whole study area (n=195). The abundances of these taxa were aggregated in 10x10-km^2 grid cells across the study area, but we only kept the taxa occurring in at least 5% of the cells to discard taxa that cannot be studied at the regional scale (n=2). For the statistical analyses, we kept the 10x10-km^2 grid cells having a field plot sampling area ≥ to 10 ha and where the selected taxa represented at least 75% of the total number of individuals originally inventoried to ensure that our dataset was representative of the within-cell tree community composition. The final dataset contains

6.1×10^6 tree individuals belonging to 193 taxa, of which 96 were analyzed at the species and 97 at the genus levels (Table S2), recorded in 185,665 plots aggregated in 1,571 10x10-km^2 grid cells. Overall, the selected taxa represented 90% of the total number of individuals originally inventoried in the selected grid cells.

For each taxon, we compiled information on three important functional traits. First, we extracted an average wood density using the global wood density database[55,56] as well as other wood density data[57]. Wood density is an integrative trait that reflects a trade-off between tree growth potential and mortality risk[28] and is thus highly informative on community dynamics[58]. It ultimately directly impacts the amount of carbon that can be stored in the vegetation[59]. Second, we extracted the leaf phenology (deciduous or evergreen) of all taxa from the large unpublished CoForTraits database[60]. This database compiles information on more than 1000 species from central Africa with values extracted from the literature (mostly from local floras, academic papers and unpublished theses). When several values were available for a given species from different sources, we attributed the one with the maximum of occurrences (ambiguities were left as unknown). At the genus level, we first computed this step for all species belonging to the genus and then attributed the phenology with the maximum of occurrences at the species level to the genus so that all congeneric species have the same weight in the phenology attribution. This approach relies on the assumption that leaf phenological traits are highly phylogenetically conserved[61]. For a few taxa (n=5), the phenology information was obtained from Ouédraogo et al.[30] and following these authors we considered *Lophira alata* Banks ex C. F. Gaertn. and *Musanga cecropioides* R. Br. as leaf exchangers, i.e., with a trait value of 0.5, intermediate between evergreen (0) and deciduous (1). Leaf phenology is one of the few traits considered in dynamic global vegetation models as it impacts the dynamics of forest productivity[62]. In particular, deciduousness indicates that tree photosynthetic activity, and thus growth, is seasonally depressed, which has a direct impact on carbon, water and nutrient cycling[63]. Deciduousness has often been interpreted as a strategy to avoid water stress and is thus expected to depend on climate and soil conditions[30,64]. Lastly, we used the original inventory data to calculate the maximum diameter as the 95th percentile value of the diameter distribution for each taxon. Maximum potential diameter, which is often used as a proxy of maximum height[65], informs both on tree competitive ability for light and on the carbon sequestration potential. At the community level, it is expected to vary along gradients of productivity and disturbance[66]. Leaf phenology was successfully assigned to 89% of the taxa (98% of the individuals), wood density to 91% of the taxa (96% of the individuals) and maximum diameter to all taxa.

Climate and soil data

We considered 24 climatic predictors derived from the open Climatic Research Unit (CRU) dataset[67] (Extended Data Table 2). We decided to rely on the CRU dataset as other datasets, such as WorldClim[68], contain erroneous observations for some climatic stations (e.g., Ngoundi in Cameroon) that severely impacted our model. Furthermore, our cross-validation approach showed that the CRU database led to higher correlations between observed and predicted taxa abundances, correspondence analyses scores and community weighted trait values than the WorldClim[68] and CHIRPS[69] databases (results not shown).

Soil maps have been published at the country scale in central Africa and their homogenization is very challenging. We thus relied on a global dataset, the Harmonized World Soil Database (HWSD)[70], to attribute a soil type to each grid cell. A cross-validation analysis of our joint distribution model revealed that soil significantly improved predictions, mostly due to the contrast between Arenic Acrisols and the other soil types (Fig. S5). To keep the model parsimonious and maximize its robustness, we thus merged all soil categories but the Arenic Acrisols soils into a single category and discarded the permanently flooded areas as mapped in the open ESA-CCI landcover product (V.1.6), where no tree inventory data were available.

Human-induced forests disturbance intensity

Many studies have attempted to spatialize human impacts on environment at a large scale. In most cases, these human footprint maps have consisted of cumulative threat maps, assuming for instance population density and infrastructure effects[71–73]. Moreover, most of these maps relied on population statistics obtained at the level of administrative entities, resulting in human footprint indices with sharp changes at administrative boundaries[74]. We thus developed a statistical model to link the probability for a forest pixel i to be impacted by anthropogenic activities depending on human population density and road proximity through nonlinear relationships. This resulted in a spatially continuous index representing human-induced forest disturbance intensity that can be projected in space and/or time following predefined human population dynamics scenarios (Extended Data Fig. 8).

We calibrated this index with the "Settlement Points" dataset produced under the "Global Rural Urban Mapping Project" (Grumpv1). This dataset provides estimates of human population (counts, in persons) for the year 2000 using a proportional allocation gridding algorithm (1-km² grid) based on more than 1,000,000 national and subnational geographic units. Focusing on central Africa, we compared this product with the Natural Earth Populated Places product (version 3.0.0; http://www.naturalearthdata.com/downloads/10m-cultural-vectors/10m-populated-places/; last access the 07/10/2018) derived from the LandScan (https://earthworks.stanford.edu/catalog/stanford-yj715rc4110#iso-metadata-reference-info) dataset (pixels with fewer than 200 persons per km² were discarded). The total number of populated points in central Africa (longitude 5.6 to 39.8, latitude -9.8 to 7.5 in decimal degrees) was 807 and 376 for the Grumpv1 and Natural Earth products, respectively. We thus performed a random manual check of the populated places present in Grumpv1 and absent from Natural Earth (the reverse rarely occurred) using Google Earth images and found that in all cases Grumpv1 was correct. We thus finally used the Grumpv1 dataset, which mostly provides information on populated places with more than ca. 1000 people. Because smaller populations may have a significant impact on forests, we added to this dataset the populated locations of the category "towns" from OpenStreetMap (https://data.maptiler.com/downloads/planet/#1.59/-17.3/19.7; last access 02/10/2018) assuming by default that they all contained 500 people (OpenStreetMap does not provide systematic information on population size).

The road network was extracted from the Global Roads Open Access Data Set, version 1 (https://data.maptiler.com/downloads/planet/#1.59/-17.3/19.7; last access the 14/09/2018) a dataset combining road data by country. Note that logging roads are not fully represented in this dataset, so we may underestimate their effect in this study. A few roads from the northern Republic of Congo were corrected using data from OpenStreetMap. Preliminary analyses revealed that further accounting for the railway and river networks did not improve predictions of tree taxon and community distributions.

Our index was thus calculated as followed. Let $z_i, i=1,\ldots,n$ be n random variables indicating the disturbance status of pixel i: 0 if the pixel is undisturbed and 1 if disturbed.

We assumed that z_i is distributed as a Bernoulli variable: $z_i = Bern(p_i)$ with $p_i = \frac{IP_i(\theta)}{IP_i(\theta)+IR_i^r}$

where $IP_i(\theta)$ is a synthetic index describing the influence of the population density of all populated places on pixel i (see below), θ is an unknown parameter to be inferred, and IR_i^r expresses the road influence on pixel i, defined as the normalized square root distance of pixel i to the nearest road r:

$$IR_i^r = \frac{\min_{r \in R} \sqrt{DR_i^r}}{max_{i=1,\ldots,n}\left(\min_{r \in R} \sqrt{DR_i^r}\right)}$$

where DR denotes the distance to the nearest road in the study area and R denotes all roads in the study area.

Population influence, IP_i^θ, is defined as the normalized square root of the weighted sum of the population size of place j. Note that the weight depends on both the distance between pixel i and populated place j, δ_{ij}, and on the population size N_j:

$$IP_i^{\theta}=\frac{\sqrt{\sum_{j}^{n} N_j e^{-\frac{\delta_{ij}}{\log(N_j)^{\theta}}}+1}}{max\sqrt{\sum_{j}^{n} N_j e^{-\frac{\delta_{ij}}{\log(N_j)^{\theta}}}+1}}$$

We finally calibrated the θ parameter using two reference areas of ca. 190,000 km^2 (Fig. S6). These two areas were chosen because they cover contrasting conditions, are well known by our team and were found to be little influenced by atmospheric pollution in the MODIS data. Degraded versus intact forests were identified from a recently published MODIS-based regional vegetation map[20]. Using a likelihood optimization approach in these two areas, we obtained θ=1.27 and 1.71 in calibration areas 1 and 2, respectively, indicating that under a similar anthropogenic context, forests tend to be disturbed at a greater distance from anthropogenic disturbance sources in the second calibration area. The final human-induced forest disturbance intensity index was thus calculated with θ =1.49, the average estimate for the two calibration areas, over the whole central African forest domain, thus avoiding the risk of artefacts related to atmospheric pollution from which suffer satellite products, especially over Gabon.

This index, built independently from our floristic dataset, outperformed previously published indices to predict floristic composition in our study area. Using a simple linear model, with individual anthropogenic indices as single predictors, the mean wood density of tree communities was better predicted with our new index (r=0.33) than with the WorldPop[75] (r=0.30), LandScan (r=0.15) and Venter[72] (r=0.23) indices. Similarly, using a simple generalized linear model with a Poisson distribution to predict the abundance of *Musanga cecropioides*, the most widespread and abundant short-lived pioneer taxon over central African forests, revealed a better performance of our index (r=0.35) compared to previous indices (r=0.31, 0.11, 0.26 for WorldPop, LandScan and Venter, respectively).

Statistical model

To predict the joint taxa distributions we relied on a recently developed methodology called supervised component generalized linear regression (SCGLR)[17], which identifies the most predictive dimensions among a large set of potentially multicollinear predictors. SCGLR is an extension of partial least-squares regression (PLSR) to the uni- and multivariate generalized linear framework. PLSR is particularly well suited for analyzing a large array of correlated predictor variables, and many studies have demonstrated its ability for prediction in various biological fields, such as genetics[76] or ecology[77]. While PLSR is well adapted for continuous variables, SCGLR is suited for non-Gaussian outcomes and noncontinuous covariates. It is a model-based approach that extends PLSR[78], PCA on instrumental variables[79], canonical correspondence analysis[80], and other related empirical methods by maximizing a trade-off between goodness of fit of the model and the quantity of information the components capture from the climatic variables. This information is measured through an indicator of “structural relevance” (SR)[81], which uses bundles of highly correlated variables to attract components to rich and robust informational dimensions.

The components are sought as K linear combinations of environmental variables common to all species with coefficient vectors denoted $u=(u_1,\ldots,u_K)$ (under norm and orthogonality constraints). SCGLR also estimates the corresponding $q\times K$ (number of species by number of components) matrix of unknown parameters γ to maximize the following convex sum:

$$s\log\psi(u,\gamma)+(1-s)\log\phi_l(u)$$

where ψ is the likelihood and ϕ_l is the SR. s and l are tuning parameters. s is related to the trade-off between goodness of fit and structural relevance. l is a nonnegative scalar related to the narrowness of the bundles of climatic variables the components are wanted to align with. The K climatic components (CCs) are then equal to $CC_k=Xu_k, k=1,\ldots,K$ and can be understood as the main environmental directions predicting all species simultaneously, while $\gamma_j, j=1\ldots,q$ are the magnitude of the effects of the K components on the abundances of each species. Then, the species abundances of each taxon $j=1,$, 193 on the grid cell $i=1,$, 1571 are modeled using a generalized linear Poisson regression such that:

$$y_{ij} \sim P(S_i \lambda_{ij})$$

$$log(\lambda_{ij}) = X_i \beta_j + T_i \alpha_j = X_i u \gamma_j + T_i \alpha_j = CC_i \gamma_j + T_i \alpha_j$$

where X denotes climatic variables (Extended data Table 2), S_i is an offset corresponding to the number of plots within each grid cell, and T is a set of covariates known to impact species abundances: here, the soil type and the human-induced forest disturbance intensity index, as well as its logarithm to account for nonlinear responses. The number of components (K) as well as the tuning parameters (l and s) must appropriately be chosen. This was done with a 10% cross-validation procedure in which the criterion used was the harmonic mean of the mean square prediction error (MSPE) across the 194 taxa. A dedicated R package, SCGLR[82], is available (see also https://github.com/SCnext/SCGLR).

To assess the predictive power of our model, we performed a leave-one block out cross-validation in which our dataset was divided into 40 spatial clusters identified with a Ward's hierarchical clustering[83] of plot coordinates (Fig. S7). All clusters but one were used for training the model (i.e., calibration dataset) and the remaining cluster was used for validating the model. We repeated this procedure 40 times such that all clusters were used once in the validation dataset and participated in the model assessment. Model validation was performed through the use of the nonparametric Spearman's rank correlation coefficients between observations and predictions. For individual taxon abundances, correlations were estimated using observed and predicted abundance per taxon. For taxon assemblages, a correspondence analysis (CA) was performed on the grid cell × observed species abundance matrix, providing the observed CA axes. The predicted site scores on each CA axis were then obtained by projecting the grid cell × predicted species abundance matrix in the observed CA planes. Correlations were computed on the observed and predicted site scores (i.e., loadings) enabling us to assess the ability of our model to predict the main floristic gradients across our area.

Finally, for the three functional traits, correlations were estimated on the grid cell-based community weighted mean (CWM) traits[27] calculated on observed and predicted taxon assemblages.

Taxon abundances and community composition were predicted across the entire study area in a regular 10-km grid. To predict the floristic composition of the existing forests, we first used the ESA-CCI landcover product (V.1.6) to only keep grid cells that are likely to be forested (i.e., category “broadleaved evergreen”). Then, we only selected grid cells that had a combination of predictor values similar to those in the calibration dataset. To do this, we used a 3-dimensional convex hull algorithm on the three climatic components to exclude all the grid cells that had a combination of predictors different from that represented in the calibration dataset. This resulted in 12,295 grid cells, representing 85% of the central African forests, i.e., an area of ca. $1{,}230{,}000$ km^2. We finally used the Ward’s hierarchical clustering method to classify the predicted floristic composition into broad floristic types. Group classification was done on the first five axes of a CA performed on predicted taxon abundances, accounting for 77% of the total inertia. The number of retained types was chosen based on our expert knowledge. The uncertainty associated with this classification was then assessed through Gaussian finite mixture models[84] (repeated 500 times) using a spherical, equal volume model (EII).

Spatially explicit null models

Whenever predictors and observations are spatially structured, model errors of type I (false positive associations) are inflated[85], hindering our capacity to extrapolate predictions in space or time[22]. We thus built a spatialized null model to test the degree to which the successfulness of our predictions resulted from an actual relationship with climatic variables or was simply due to spatial costructures between taxon distributions and climatic gradients that arose by chance. We used the RGEOSTAT R package[86] to simulate sets of SCGLR climatic components (CCs) having similar spatial properties to those of the observed CCs as well as similar spatial costructures between them. This step consisted of fitting theoretical variograms and covariograms to empirical ones to simulate random realizations that can be then used as “null” spatialized predictors (Fig. S8 and S9). We simulated 500 sets of “null” spatialized predictors and used them as predictors in our GLMs using the leave-one block out cross-validation described above. The resulting correlations between observed and predicted taxon abundances, and axes scores (for taxon assemblages) were finally compared with the correlations obtained when observed climatic predictors were considered. The resulting p-values were calculated as the number of times a simulated correlation was higher than the observed one, divided by the total number of realizations (n=501).

Forest vulnerability to global change

Vulnerability to climate change, as assessed through the IPCC framework, is composed of three components: (1) sensitivity, (2) exposure, and (3) adaptive capacity to climate change.

Sensitivity to climate change, $Sensitivity_{clim}$, was firstly estimated at the taxon level in a similar way to Foden et al.[34]. For each taxon, we calculated the mean of the weighted standard deviation (SDw) of the three climatic components (CCs) at the present time, using locally observed taxon abundances as weights. SDw thus represents the width of the climatic niche currently occupied by the taxa. Taxon-specific climate sensitivity was then measured as $1/\overline{SDw}$ (it increases as niche width decreases). To upscale tree sensitivity to climate change at the community level and over our study area, sensitivity was measured as the community weighted mean (CWM) of taxon-specific climate sensitivity scores, using predicted taxon assemblages.

Exposure to climate change, $Exposure_{clim}$, was assessed using projected changes in climate from 18 unique climate model combinations provided by the AFRICLIM V3.0 dataset[87] (last access on 03/02/2020). These models corresponded to pairwise combinations of five regional climate models (RCMs) driven by ten general circulation models (GCMs), with an unequal number of GCMs models per RCM (ten models for the Swedish Meteorological and Hydrological (SMHI) RCM, four for the Climate Limited-area Modelling Community (CLMCOM) RCM, two for the Royal Netherlands

Meteorological Institute (KNMI) RCM, one for the Canadian Centre for Climate Modelling (CCCMA) RCM and one for the Danish Meteorological Institute (DMI) RCM). These models were generated using change-factor downscaling approaches to model spatial variation at local scales while correcting for differences between observed and simulated baseline climates (see Platts et al.[87] for more details). We here concentrated on one representative concentration pathway of the IPCC-AR5 (RCP 4.5) for the late 21st century (2071-2100, hereafter named 2085) and reconstructed the three SCGLR selected CCs from the climatic predictions as follows: let $X_{rcp4.5}$ be the predicted future climatic conditions. Let $m=\overline{X}$ and $S=sd(X)$ be the mean and standard deviation matrices of the current climatic conditions. The predictive climatic components under future scenarios are then equal to $f_{rcp4.5}=(X_{rcp4.5}-m)S\hat{u}$, where $\hat{u}$ represents SCGLR CCs. We then calculated the euclidean distance between the three current and the three predicted CCs for each of the 18 models and then estimated the exposure to climate change as the mean distance over the 18 models.

Adaptive capacity to climate change, $Adaptive_{clim}$, was assessed through the phylogenetic diversity of predicted assemblages at the genus level. We used a recently published dated phylogeny[88], covering 167 out of our 180 genera (representing 94% of predicted individuals). We first tested if the studied taxa exhibited a significant conservatism in their climate niches using Abouheif's permutation tests (Abouheif, 1999) on the taxa-specific score (γ) values on the three SCGLR climate components (γ represents the influence of a CC on a given taxa distribution, see above). We then measured the phylogenetic diversity (PD) of predicted assemblages at the genera level using the Chao's PD index with an order q of 1 (equivalent to the Shannon index)[89] that we used as a proxy of adaptive capacity.

Vulnerability to climate change, $Vulnerability_{clim}$, was finally estimated as the sum of the three standardized (0 to 1) components:

$$Vulnerability_{clim}=\left(Sensitivity_{clim}^{st}+Exposure_{clim}^{st}-Adaptive_{clim}^{st}\right).$$

$Vulnerability_{clim}$ theoretically ranges from -1 (low vulnerability) to 2 (high vulnerability) and, due to the standardization of its three components, it expresses a relative vulnerability over the study area and is thus little impacted by the IPCC scenario chosen (RCP 4.5 or 8.5) because different scenarios predict different amplitudes of changes but similar spatial patterns (Extended Data Fig. 7).

Forecasted human impact on forests in 2085 was assessed using our human-induced forest disturbance intensity index combined with country-specific projections of human populations in 2085. We assigned to each current town a country-specific relative population increase drawn from the World population prospects (United Nations)[90] and rebuild our index based on this modified dataset. This approach did not account for new roads that may established by 2085, and thus tended to underestimate the increase in anthropogenic pressure.

Software and packages

All analyses were performed and figures were created with the R statistical software[91], mostly using the ade4[92], alphashape3d[93], ggplot2[94], raster[95], RgeoStat[96], entropart[97] and SCGLR (https://github.com/SCnext/SCGLR/) packages. Data are archived in a public repository[98].

Data availability

All maps and data used for this article are accessible online in a public repository at http://dx.doi.org/10.18167/DVN1/UCNCA7. Raw floristic data are, however, archived in a private data repository, due to the highly sensitive nature of commercial inventory data, and access may be granted for research purpose using the form provided in the public repository.

Code availability

R scripts are available at https://github.com/MaximeRM/ScriptNature.

Methods references

51. Réjou-Méchain, M. *et al.* Detecting large-scale diversity patterns in tropical trees: Can we trust commercial forest inventories? *Forest Ecology and Management* **261**, 187–194 (2011).

52. Chase, M. W. *et al.* An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society* **161**, 105–121 (2009).

53. Dauby, G. *et al.* RAINBIO: A mega-database of tropical African vascular plants distributions. *PhytoKeys* 1 (2016).

54. African Plant Database (version 3.4.0). Conservatoire et Jardin botaniques de la Ville de Genève and South African National Biodiversity Institute, Pretoria. Retrieved the 10/02/2017.

55. Chave, J. *et al.* Towards a worldwide wood economics spectrum. *Ecology Letters* **12**, 351–366 (2009).

56. Zanne, A. E. *et al.* Global wood density database. *Dryad. Identifier: http://hdl. handle. net/10255/dryad* **235**, (2009).

57. Gourlet-Fleury, S. *et al.* Environmental filtering of dense-wooded species controls above-ground biomass stored in African moist forests. *Journal of Ecology* **99**, 981–990 (2011).

58. Westoby, M. & Wright, I. J. Land-plant ecology on the basis of functional traits. *Trends in Ecology & Evolution* **21**, 261–268 (2006).

59. Chave, J. *et al.* Improved allometric models to estimate the aboveground biomass of tropical trees. *Global Change Biology* **20**, 3177–3190 (2014).

60. Bénédet, F. *et al.* CoForTraits, African plant traits information database . Version 1.0. (2013).

61. Davies, T. J. *et al.* Phylogenetic conservatism in plant phenology. *Journal of ecology* **101**, 1520–1530 (2013).

62. Cramer, W. *et al.* Global response of terrestrial ecosystem structure and function to CO2 and climate change: Results from six dynamic global vegetation models. *Global change biology* **7**, 357–373 (2001).

63. Menzel, A. Phenology: Its importance to the global change community. *Climatic change* **54**, 379 (2002).

64. Borchert, R., Rivera, G. & Hagnauer, W. Modification of Vegetative Phenology in a Tropical Semi-deciduous Forest by Abnormal Drought and Rain 1. *Biotropica* **34**, 27–39 (2002).

65. Kraft, N. J. B., Valencia, R. & Ackerly, D. D. Functional traits and niche-based tree community assembly in an amazonian forest. *Science* **322**, 580–582 (2008).

66. Schamp, B. S. & Aarssen, L. W. The assembly of forest communities according to maximum species height along resource and disturbance gradients. *Oikos* **118**, 564–572 (2009).

67. New, M., Lister, D., Hulme, M. & Makin, I. A high-resolution data set of surface climate over global land areas. *Climate research* **21**, 1–25 (2002).

68. Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. & Jarvis, A. Very high resolution interpolated climate surface for global land areas. *International journal of climatology* **25**, 1965–1978 (2005).

69. Funk, C. *et al.* The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific data* **2**, 150066 (2015).

70. Nachtergaele, F. *et al.* The harmonized world soil database. in *Proceedings of the 19th World Congress of Soil Science, Soil Solutions for a Changing World, Brisbane, Australia, 1-6 August 2010* 34–37 (2010).

71. Woolmer, G. *et al.* Rescaling the human footprint: A tool for conservation planning at an ecoregional scale. *Landscape and Urban Planning* **87**, 42–53 (2008).

72. Venter, O. *et al.* Sixteen years of change in the global terrestrial human footprint and implications for biodiversity conservation. *Nature Communications* **7**, 12558 (2016).

73. Geldmann, J., Joppa, L. N. & Burgess, N. D. Mapping change in human pressure globally on land and within protected areas. *Conservation biology* **28**, 1604–1616 (2014).

74. Linard, C., Gilbert, M., Snow, R. W., Noor, A. M. & Tatem, A. J. Population distribution, settlement patterns and accessibility across Africa in 2010. *PloS one* **7**, e31743 (2012).

75. Lloyd, C. T. *et al.* Global spatio-temporally harmonised datasets for producing high-resolution gridded population distribution datasets. *Big Earth Data* **3**, 108–139 (2019).

76. Boulesteix, A.-L. & Strimmer, K. Partial least squares: A versatile tool for the analysis of high-dimensional genomic data. *Briefings in bioinformatics* **8**, 32–44 (2007).

77. Carrascal, L. M., Galván, I. & Gordo, O. Partial least squares regression as an alternative to current regression methods used in ecology. *Oikos* **118**, 681–690 (2009).

78. Tenenhaus, M. *La régression PLS: Théorie et pratique*. (Editions Technip, 1998).

79. Sabatier, R., Lebreton, J. D. & Chessel, D. Principal component analysis with instrumental variables as a tool for modelling composition data. *Multiway data analysis* 341–352 (1989).

80. Ter Braak, C. J. The analysis of vegetation-environment relationships by canonical correspondence analysis. in *Theory and models in vegetation science* 69–77 (Springer, 1987).

81. Bry, X. & Verron, T. THEME: THEmatic model exploration through multiple co-structure maximization. *Journal of Chemometrics* **29**, 637–647 (2015).

82. Cornu, G., Mortier, F., Trottier, C. & Bry, X. SCGLR: Supervised Component Generalized Linear Regression. (2016).

83. Ward Jr, J. H. Hierarchical grouping to optimize an objective function. *Journal of the American statistical association* **58**, 236–244 (1963).

84. Scrucca, L., Fop, M., Murphy, T. B. & Raftery, A. E. Mclust 5: Clustering, classification and density estimation using gaussian finite mixture models. *The R journal* **8**, 289 (2016).

85. Dormann, C. F. *et al.* Methods to account for spatial autocorrelation in the analysis of species distributional data: A review. *Ecography* **30**, 609–628 (2007).

86. Renard, D. *et al.* RGeostats: The Geostatistical package 11.0. 1. *MINES ParisTech* (2014).

87. Platts, P. J., Omeny, P. A. & Marchant, R. AFRICLIM: High-resolution climate projections for ecological applications in Africa. *African Journal of Ecology* **53**, 103–108 (2015).

88. Janssens, S. B. *et al.* A large-scale species level dated angiosperm phylogeny for evolutionary and ecological analyses. *Biodiversity data journal* **8**, (2020).

89. Chao, A., Chiu, C.-H. & Jost, L. Phylogenetic diversity measures based on Hill numbers. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**, 3599–3609 (2010).

90. Nations, U. World population prospects: The 2017 revision, key findings and advance tables. *United Nations, New york* (2017).

91. R Core Team. *R: A language and environment for statistical computing*. (R Foundation for Statistical Computing, 2017).

92. Chessel, D., Dufour, A. B. & Thioulouse, J. The ade4 package-I- One-table methods. *R News* **7**, 47–52 (2004).

93. Lafarge, T. & Pateiro-Lopez, B. *Alphashape3d: Implementation of the 3D alpha-shape for the reconstruction of 3D sets from a point cloud*. (2017).

94. Wickham, H. *Ggplot2: Elegant graphics for data analysis*. (Springer-Verlag New York, 2016).

95. Hijmans, R. J. *Raster: Geographic data analysis and modeling*. (2017).

96. Renard, D. *et al. RGeostats: Geostatistical package*. (2017).

97. Marcon, E. & Hérault, B. entropart: An R package to measure and partition diversity. *Journal of Statistical Software* **67**, 1–26 (2015).

98. Réjou-Méchain, M. et al. *Maps of central African rainforest composition and vulnerability*, doi:10.18167/DVN1/UCNCA7, CIRAD Dataverse, V2, (2021).

Acknowledgements

We thank the 105 forest companies that provided access, albeit restricted, to their inventory data for research purposes and members of the central african plot network (https://central-african-plot-network.netlify.app/), Y. Yalibanda, F. Allah-Barem, F. Baya, F. Boyemba, M. Mbasi Mbula, P. Berenger, M. Mazengue, V. Istace, I. Zombo, E. Forni, Nature+ and the CEB-Precious Woods company for giving access to the scientific inventories described in Fig. S4, some of which were funded by the AFD and the FFEM (e.g. DynAfFor and P3FAC projects). We thank J. Chave, P. Couteron, S. Lewis and M. Tadesse for their helpful comments and discussions on previous versions. We also thank the five reviewers for their highly constructive comments on previous versions, B. Sultan for useful discussions on climate projections, O.J. Hardy for advices on phylogenetical analyses, B. Locatelli for advices on vulnerability analyses, G. Vieilledent for our discussions on the human-induced forest disturbance intensity index and A. Stokes for English editing. This work was supported by the CoForTips project (ANR-12-

EBID-0002) funded by the ERA-NET BiodivERsA, with the national funders ANR, BELSPO and FWF, as part of the 2012 BiodivERsA call for research proposals, the GAMBAS project funded by the French National Research Agency (ANR-18-CE02-0025) and the project 3DForMod funded by the UE FACCE ERA-GAS consortium (ANR-17-EGAS-0002-01). This study is a contribution to the research program of LMI DYCOFAC (Dynamique des écosystèmes continentaux d'Afrique Centrale en contexte de changements globaux).

Author contributions,

Conceptualization: M.R-M., F.M., R.P. & S.G-F; data curation: G.C. & F.B.; formal analysis: M.R-M. & F.M.; project administration: C.G.; writing – original draft: M.R-M., F.M., R.P. & S.G-F; writing – review & editing: all authors.

Competing interests

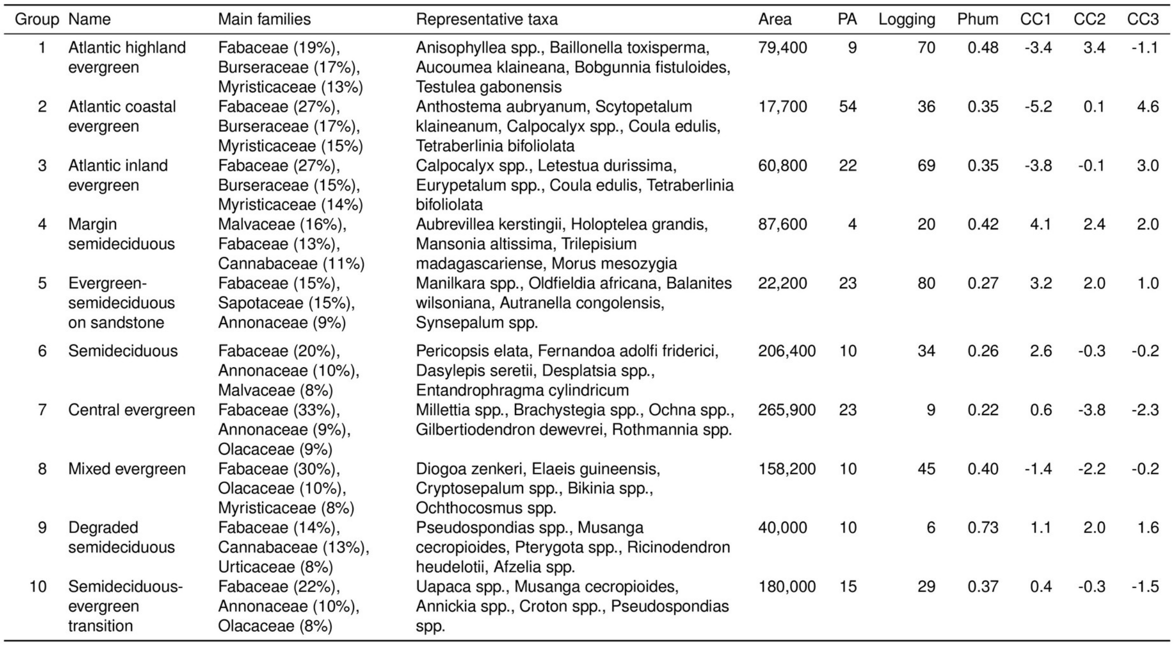
Authors declare no competing interests.

Supplementary Information is available for this paper

Correspondence and requests for materials should be addressed to maxime.rejou@ird.fr

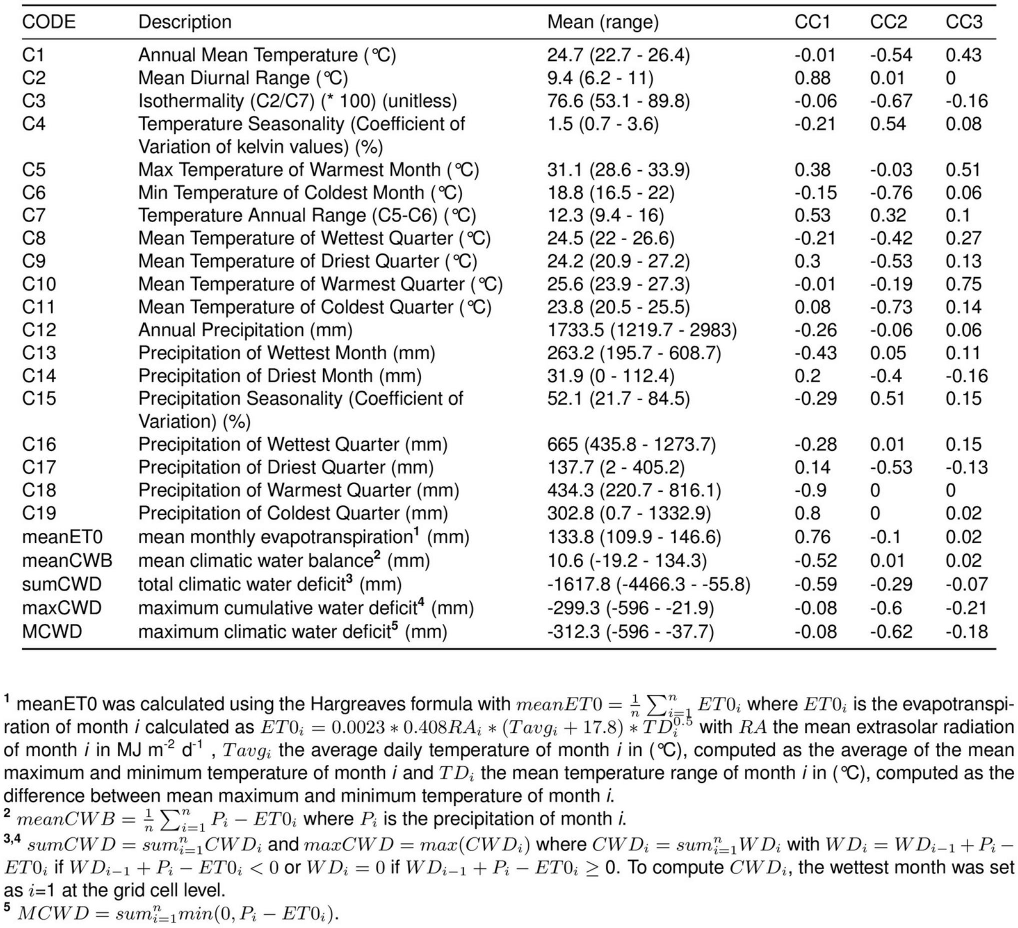
Reprints and permissions information is available at www.nature.com/reprints

Extended data figure/table legends:



Group	Name	Main families	Representative taxa	Area	PA	Logging	Phum	CC1	CC2	CC3
1	Atlantic highland evergreen	Fabaceae (19%), Burseraceae (17%), Myristicaceae (13%)	Anisophyllea spp., Baillonella toxisperma, Aucoumea klaineana, Bobgunnia fistuloides, Testulea gabonensis	79,400	9	70	0.48	-3.4	3.4	-1.1
2	Atlantic coastal evergreen	Fabaceae (27%), Burseraceae (17%), Myristicaceae (15%)	Anthostema aubryanum, Scytopetalum klaineanum, Calpocalyx spp., Coula edulis, Tetraberlinia bifoliolata	17,700	54	36	0.35	-5.2	0.1	4.6
3	Atlantic inland evergreen	Fabaceae (27%), Burseraceae (15%), Myristicaceae (14%)	Calpocalyx spp., Letestua durissima, Eurypetalum spp., Coula edulis, Tetraberlinia bifoliolata	60,800	22	69	0.35	-3.8	-0.1	3.0
4	Margin semideciduous	Malvaceae (16%), Fabaceae (13%), Cannabaceae (11%)	Aubrevillea kerstingii, Holoptelea grandis, Mansonia altissima, Trilepisium madagascariense, Morus mesozygia	87,600	4	20	0.42	4.1	2.4	2.0
5	Evergreen-semideciduous on sandstone	Fabaceae (15%), Sapotaceae (15%), Annonaceae (9%)	Manilkara spp., Oldfieldia africana, Balanites wilsoniana, Autranella congolensis, Synsepalum spp.	22,200	23	80	0.27	3.2	2.0	1.0
6	Semideciduous	Fabaceae (20%), Annonaceae (10%), Malvaceae (8%)	Pericopsis elata, Fernandoa adolfi friderici, Dasylepis seretii, Desplatsia spp., Entandrophragma cylindricum	206,400	10	34	0.26	2.6	-0.3	-0.2
7	Central evergreen	Fabaceae (33%), Annonaceae (9%), Olacaceae (9%)	Millettia spp., Brachystegia spp., Ochna spp., Gilbertiodendron dewevrei, Rothmannia spp.	265,900	23	9	0.22	0.6	-3.8	-2.3
8	Mixed evergreen	Fabaceae (30%), Olacaceae (10%), Myristicaceae (8%)	Diogoa zenkeri, Elaeis guineensis, Cryptosepalum spp., Bikinia spp., Ochthocosmus spp.	158,200	10	45	0.40	-1.4	-2.2	-0.2
9	Degraded semideciduous	Fabaceae (14%), Cannabaceae (13%), Urticaceae (8%)	Pseudospondias spp., Musanga cecropioides, Pterygota spp., Ricinodendron heudelotii, Afzelia spp.	40,000	10	6	0.73	1.1	2.0	1.6
10	Semideciduous-evergreen transition	Fabaceae (22%), Annonaceae (10%), Olacaceae (8%)	Uapaca spp., Musanga cecropioides, Annickia spp., Croton spp., Pseudospondias spp.	180,000	15	29	0.37	0.4	-0.3	-1.5

Extended Data Table 1 | Characteristics of the floristic groups. For each floristic group information is given on the three most abundant families (APGIII classification, except for the subfamilies Caesalpiniaceae and Mimosaceae, which were considered here independently due to their different ecological strategies), the five most representative taxa (i.e., taxa having the highest A score of the Dufrêne and Legendre index), the total area (km 2) covered by each group, the percentage of the area covered by protected areas (PA) and dedicated to logging activities (Logging), the mean probability of being impacted by human activities (phum, this study) and the mean value of the three climatic components (CCs) that best explain the current distribution of central African trees (this study).



CODE	Description	Mean (range)	CC1	CC2	CC3
C1	Annual Mean Temperature (℃)	24.7 (22.7 - 26.4)	-0.01	-0.54	0.43
C2	Mean Diurnal Range (℃)	9.4 (6.2 - 11)	0.88	0.01	0
C3	Isothermality (C2/C7) (* 100) (unitless)	76.6 (53.1 - 89.8)	-0.06	-0.67	-0.16
C4	Temperature Seasonality (Coefficient of Variation of kelvin values) (%)	1.5 (0.7 - 3.6)	-0.21	0.54	0.08
C5	Max Temperature of Warmest Month (℃)	31.1 (28.6 - 33.9)	0.38	-0.03	0.51
C6	Min Temperature of Coldest Month (℃)	18.8 (16.5 - 22)	-0.15	-0.76	0.06
C7	Temperature Annual Range (C5-C6) (℃)	12.3 (9.4 - 16)	0.53	0.32	0.1
C8	Mean Temperature of Wettest Quarter (℃)	24.5 (22 - 26.6)	-0.21	-0.42	0.27
C9	Mean Temperature of Driest Quarter (℃)	24.2 (20.9 - 27.2)	0.3	-0.53	0.13
C10	Mean Temperature of Warmest Quarter (℃)	25.6 (23.9 - 27.3)	-0.01	-0.19	0.75
C11	Mean Temperature of Coldest Quarter (℃)	23.8 (20.5 - 25.5)	0.08	-0.73	0.14
C12	Annual Precipitation (mm)	1733.5 (1219.7 - 2983)	-0.26	-0.06	0.06
C13	Precipitation of Wettest Month (mm)	263.2 (195.7 - 608.7)	-0.43	0.05	0.11
C14	Precipitation of Driest Month (mm)	31.9 (0 - 112.4)	0.2	-0.4	-0.16
C15	Precipitation Seasonality (Coefficient of Variation) (%)	52.1 (21.7 - 84.5)	-0.29	0.51	0.15
C16	Precipitation of Wettest Quarter (mm)	665 (435.8 - 1273.7)	-0.28	0.01	0.15
C17	Precipitation of Driest Quarter (mm)	137.7 (2 - 405.2)	0.14	-0.53	-0.13
C18	Precipitation of Warmest Quarter (mm)	434.3 (220.7 - 816.1)	-0.9	0	0
C19	Precipitation of Coldest Quarter (mm)	302.8 (0.7 - 1332.9)	0.8	0	0.02
meanET0	mean monthly evapotranspiration[1] (mm)	133.8 (109.9 - 146.6)	0.76	-0.1	0.02
meanCWB	mean climatic water balance[2] (mm)	10.6 (-19.2 - 134.3)	-0.52	0.01	0.02
sumCWD	total climatic water deficit[3] (mm)	-1617.8 (-4466.3 - -55.8)	-0.59	-0.29	-0.07
maxCWD	maximum cumulative water deficit[4] (mm)	-299.3 (-596 - -21.9)	-0.08	-0.6	-0.21
MCWD	maximum climatic water deficit[5] (mm)	-312.3 (-596 - -37.7)	-0.08	-0.62	-0.18

[1] meanET0 was calculated using the Hargreaves formula with $meanET0 = \frac{1}{n}\sum_{i=1}^{n} ET0_i$ where $ET0_i$ is the evapotranspiration of month *i* calculated as $ET0_i = 0.0023 * 0.408 RA_i * (Tavg_i + 17.8) * TD_i^{0.5}$ with RA the mean extrasolar radiation of month *i* in MJ m^{-2} d^{-1} , $Tavg_i$ the average daily temperature of month *i* in (℃), computed as the average of the mean maximum and minimum temperature of month *i* and TD_i the mean temperature range of month *i* in (℃), computed as the difference between mean maximum and minimum temperature of month *i*.

[2] $meanCWB = \frac{1}{n}\sum_{i=1}^{n} P_i - ET0_i$ where P_i is the precipitation of month *i*.

[3,4] $sumCWD = sum_{i=1}^{n} CWD_i$ and $maxCWD = max(CWD_i)$ where $CWD_i = sum_{i=1}^{n} WD_i$ with $WD_i = WD_{i-1} + P_i - ET0_i$ if $WD_{i-1} + P_i - ET0_i < 0$ or $WD_i = 0$ if $WD_{i-1} + P_i - ET0_i \geq 0$. To compute CWD_i, the wettest month was set as i=1 at the grid cell level.

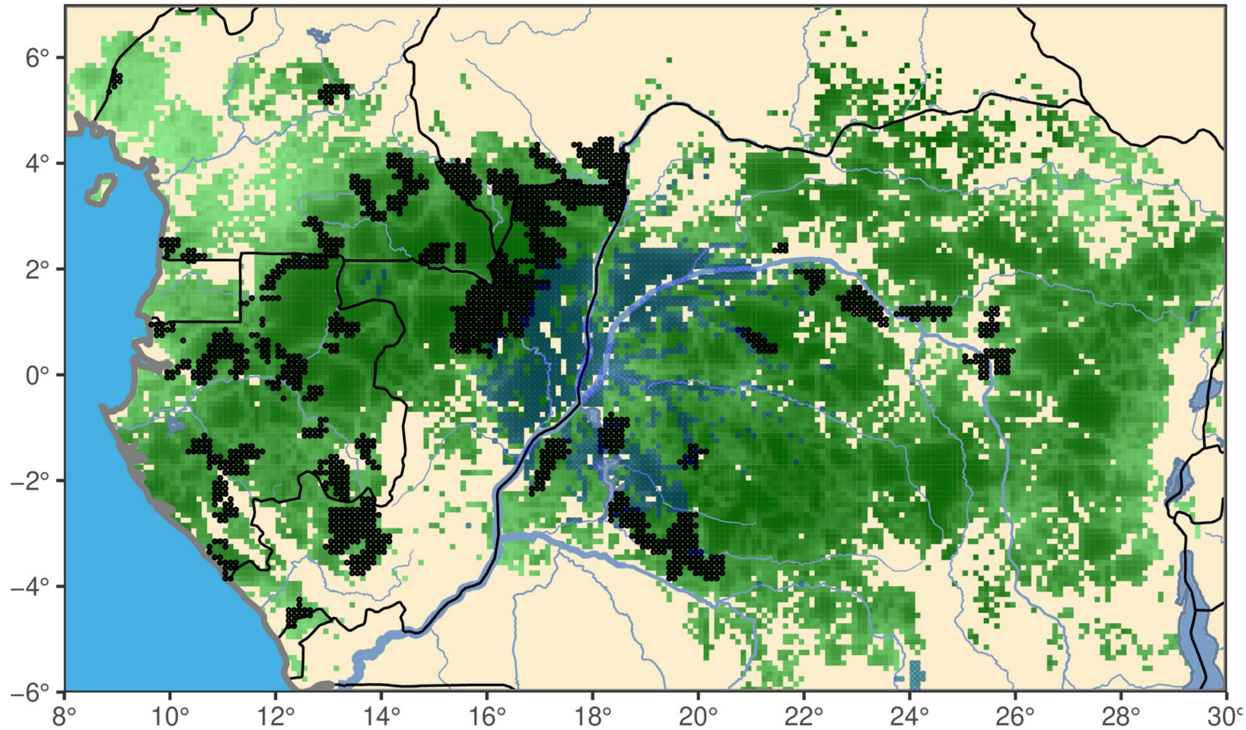
[5] $MCWD = sum_{i=1}^{n} min(0, P_i - ET0_i)$.

Extended Data Table 2 | Climatic predictors. Correlations with the three climatic components (CCs) are given in the last three columns (see also Extended Data Fig. 4).

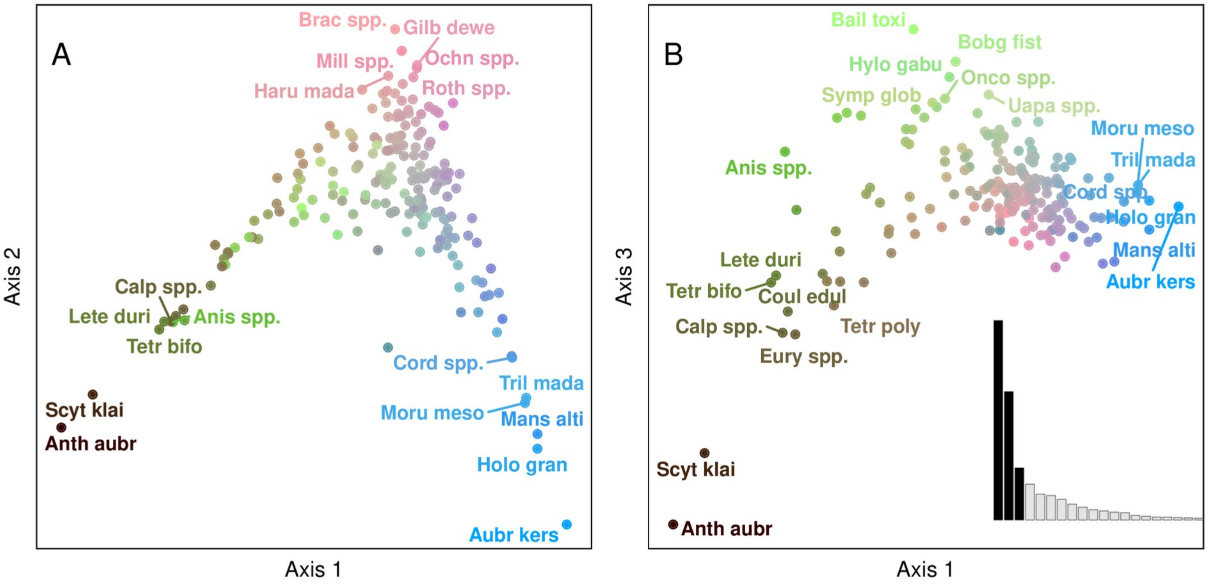
[1]meanET0 was calculated using the Hargreaves formula with $meanET0 = \frac{1}{n}\sum_{i=1}^{n} ET0_i$ where *ET0*$ET0_i$ is the evapotranspiration of month *i* calculated as $ET0_i$= 0.0023*0.408RA_i * ($Tavg_i$ + 17.8)*$TD_i^{0.5}$ with RA_i the mean extrasolar radiation of month *i* in MJ m^{-2} d^{-1}, $Tavg_i$ the average daily temperature of month *i* in °C, computed as the average of the mean maximum and minimum temperature of month *i*, and TD_i the mean temperature range of month *i* in °C, computed as the difference between mean maximum and minimum temperature of month *i*. [2]$meanCWB = \frac{1}{n}\sum_{i=1}^{n} P_i - ET0_i$where P_i is the precipitation

of month *i*. [3,4] $sumCWD=\sum_{i=1}^{n} CWD_i$ and $maxCWD=max\left(CWD_i\right)$ where $sumCWB=\sum_{i=1}^{n} CWD_i$ with $WD_i=WD_{i-1}+P_i-ET\,0_i$ if $WD_i=WD_{i-1}+P_i-ET\,0_i$ or $WD_i=0$ if $\left(WD_{i-1}+P_i-ET\,0_i\right)\geq 0$. To compute CWD_i, the wettest month was set as i=1 at the grid cell level. [5]

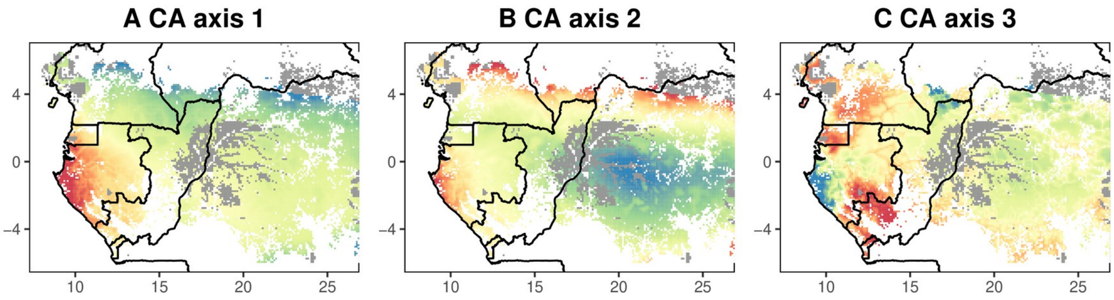
$$MCWD=\sum_{i=1}^{n} min\left(0, P_i-ET\,0_i\right).$$



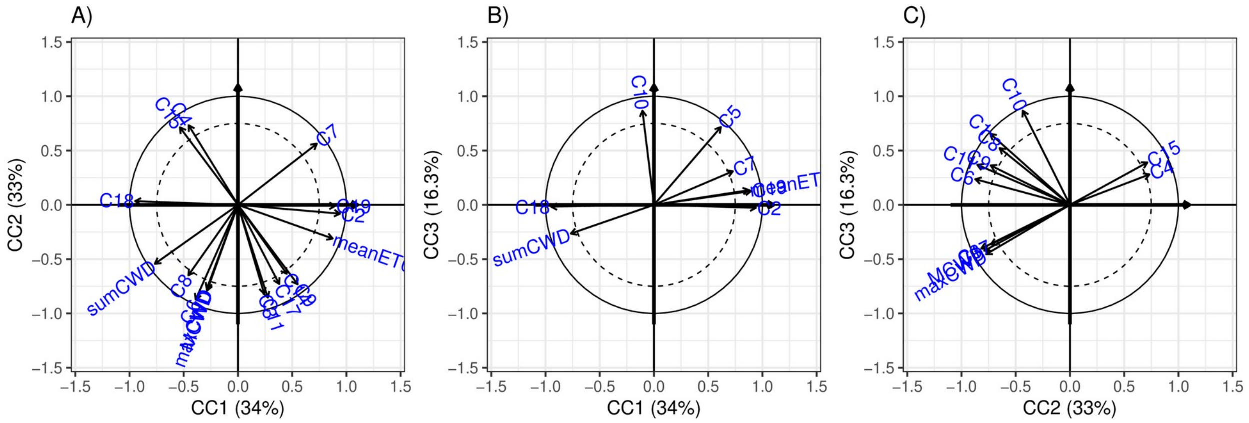
Extended Data Figure 1 | Study area and sampling plots. In green, the current distribution of tropical forests following the European Space Agency Climate Change Initiative (ESA-CCI) landcover (V.1.6) with a dark green-to-white gradient representing anthropogenic pressure (see methods) and non-forested areas represented in beige; the sampling grid cells (n=1,571 10x10-km^2 grid cells) are in black and the flooding forests, as proposed by the ESA-CCI landcover, are in blue.



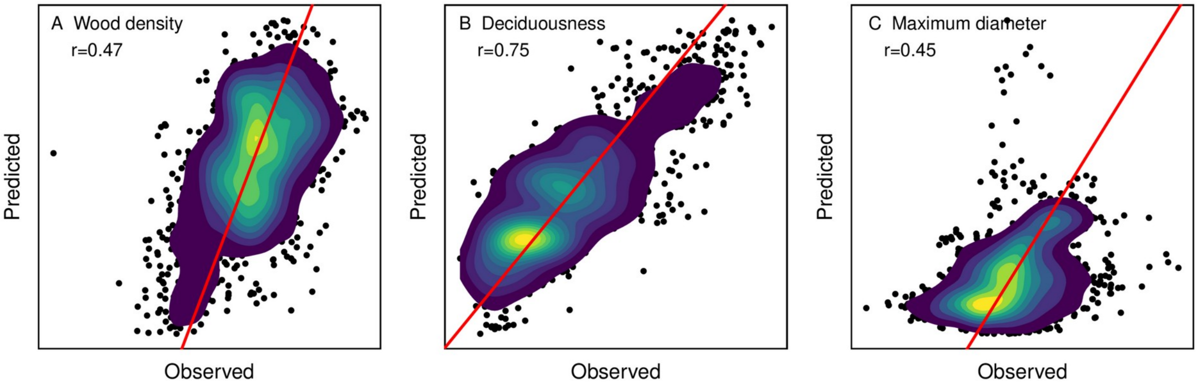
Extended Data Figure 2 | Taxon CA planes 1-2 (A) and 1-3 (B) with labels for the 12 most representative taxa on each axis. Color code corresponds to that reported in Fig. 1. The first eigenvalues are reported in the B panel, highligthing in black the first three axes. Taxon codes and scores of the 193 taxa are given in Table S2.



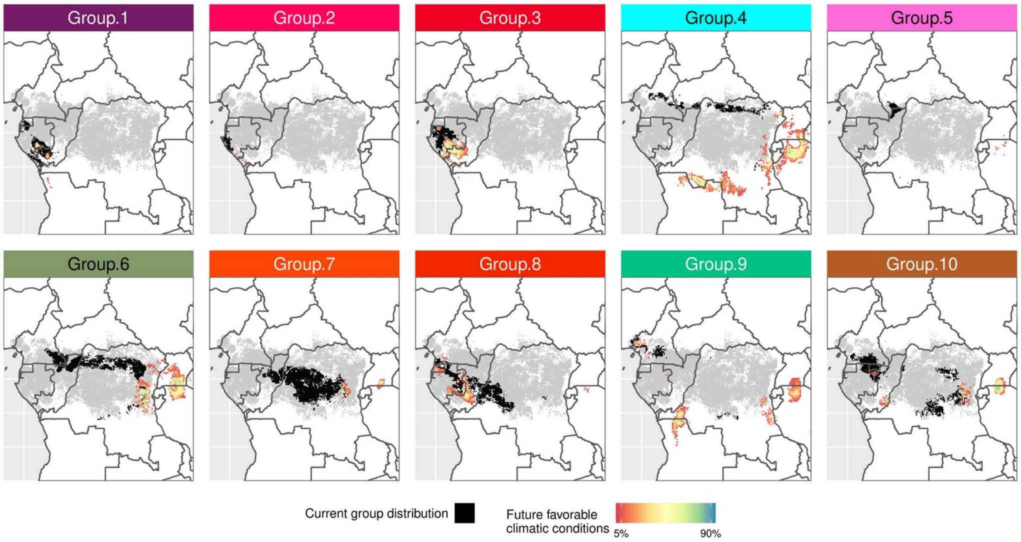
Extended Data Figure 3 | Individual predicted floristic gradients illustrated by the three first axes of the correspondence analysis (CA) performed on predicted taxon abundances. A composite map of these three axes is given in Fig. 1 and the corresponding taxon CA planes are provided in Extended Data Fig. 2.



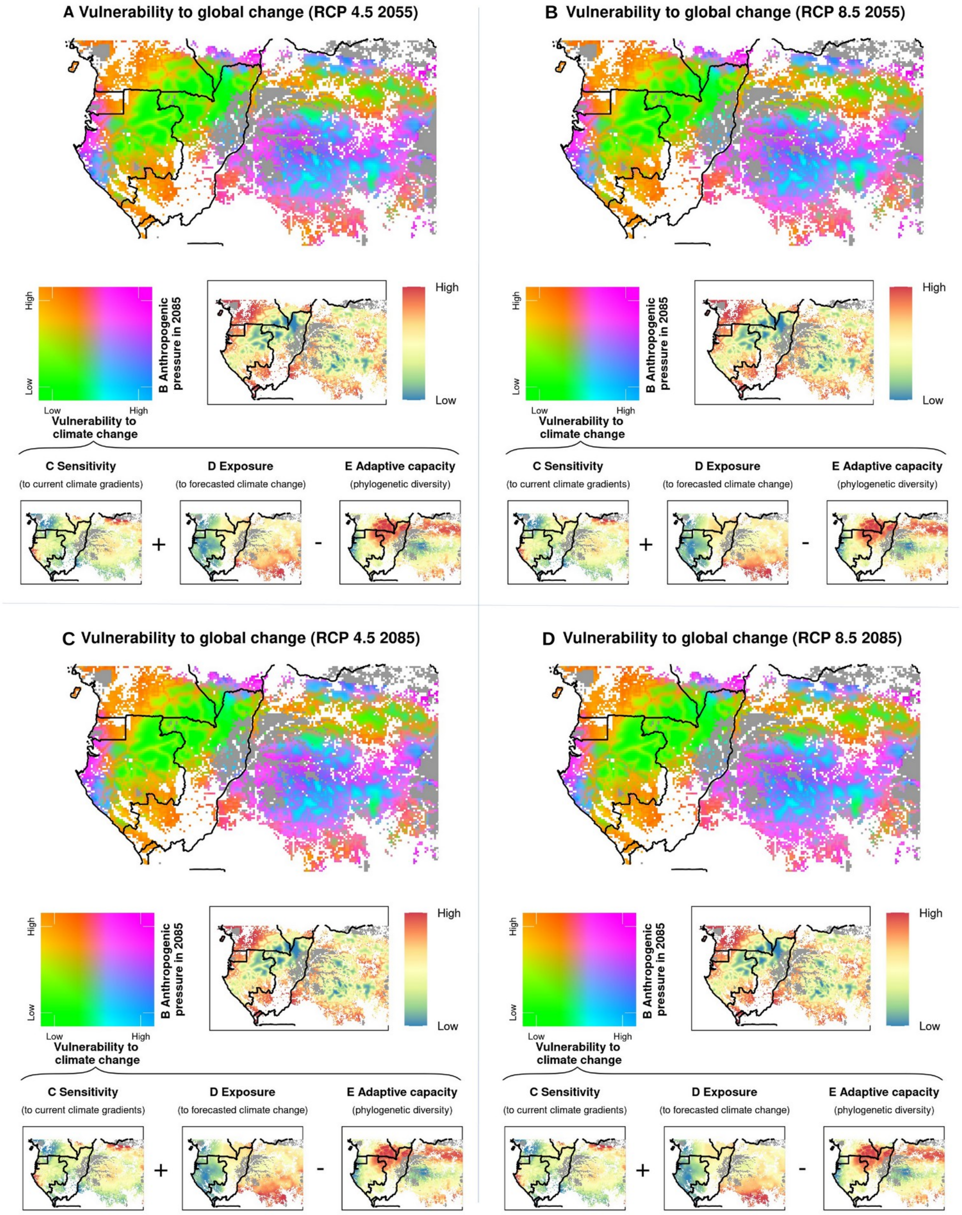
Extended Data Figure 4 | Plans 1-2 (A), 1-3 (B) and 2-3 (C) of the SCGLR climatic components (CCs). All climatic variables having a correlation < 0.75 with the two components (dashed circle) were excluded for the sake of clarity. For abbreviations, see Extended Data Table 2.



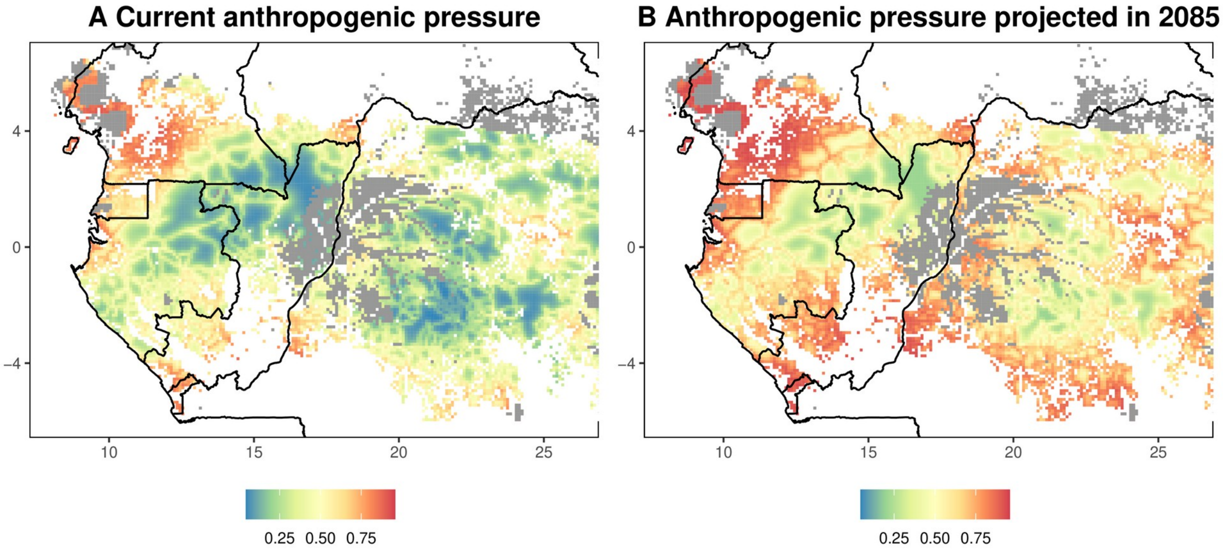
Extended Data Figure 5 | Spatial cross-validation results of the predictions of functional assemblages. The observed and predicted community weighted mean trait values within the 1,571 10x10-km^2 grid cells are given for (**A**) wood density; (**B**) deciduousness and (**C**) maximum diameter. The 1:1 line is displayed in red.



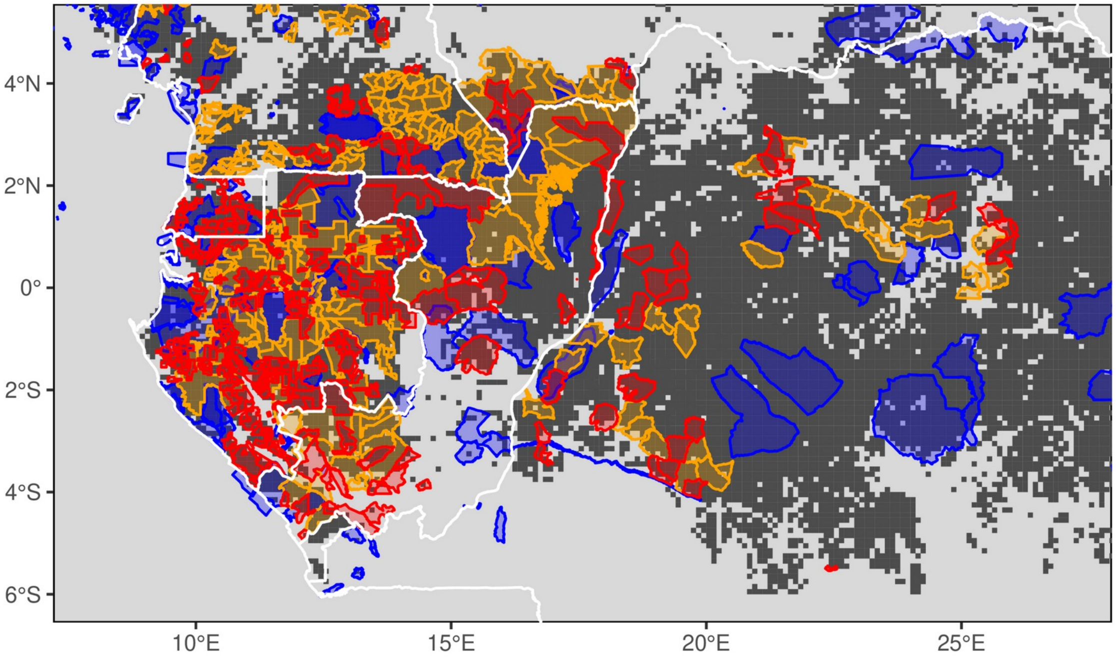
Extended Data Figure 6 | Projected changes under RCP scenario 4.5 in 2085 of the climatic conditions of the ten forest types. Areas where climate models predict similar climatic components (CCs) values as those currently found within forest types (in black) are illustrated with a color gradient indicating the level of agreement amongst the 18 climate models (in %; no color indicates that none of the original 18 climate models predicted similar conditions). More specifically, we used 3D concave hull (alpha shape) models to assess where the combinations of current Ccs corresponding to each forest type are predicted to be represented in 2085.



Extended Data Figure 7 | The vulnerability map under two different RCP scenarios, RCP 4.5 and RCP 8.5, and for two years, year 2055 and year 2085. As can be seen, the predicted vulnerability is little impacted by the IPCC scenario chosen because it expresses a relative vulnerability over the study area and, if different scenarios predict different amplitudes of climate change, spatial patterns of climate exposure remains similar (see Methods).



Extended Data Figure 8 | Current and projected anthropogenic pressure over central Africa predicted from our index of human-induced forest disturbance intensity.



Extended Data Figure 9 | Protected area network (blue) and areas dedicated to logging activities (orange and red) in central Africa. Data on protected areas were obtained from the World Database on Protected Areas (last access: 14/08/2018) excluding marine, hunting and game-oriented areas except for the Democratic Republic of Congo where data from the World Resource Institute were used and downloaded from ArcGIS hub (last access: 01/06/2019). Logging activity data were kindly provided by the Observatoire des Forêts d'Afrique Centrale based on an unpublished work completed in June 2018, except for DRC where more updated data (June 2019) were provided by the AGEDUFOR national project. Areas in orange illustrate forest concessions that are known to have, or to be in the process of having, an officially validated sustainable forest management plan. Red areas illustrate forest areas that are currently dedicated to logging but that either do not have an official management plan or have an uncertain status.