

<https://doi.org/10.1038/s44185-026-00138-5>

Climatic preferences of major plant clades define the functional attributes of African savanna types

Vitus Pickelmann  & Steven I. Higgins

Forecasting how the Earth system will respond to global change includes simplifying the functional diversity induced by thousands of plant species into tractable units. The dominant strategy has been to aggregate species into plant functional types (PFTs), assuming that selection leads to unrelated species converging to similar ecological roles. An alternative would be to harness niche conservatism, the tendency of lineages to retain ancestral ecological traits, and assume that phylogenetically related species tend to have similar ecological roles. Using African savannas as case study, we use phylogenetically defined groups to identify phytoclims: climate-defined regions that support distinct sets of plant types. We found that taxonomically-based phytoclims aligned with a respected expert map of Africa. Moreover, this scheme could be reconciled with existing conceptual models used to describe the functional diversity of savannas. This alignment of phylogenetic and functional interpretations suggests that African savannas are dominated by a limited set of pre-adapted taxa, rather than arising through widespread convergence. Our findings suggest that phylogenetic information can provide a parsimonious, functional basis for representing ecosystem diversity in global change models. Harnessing this alignment between phylogenetic and functional groupings offers a promising route for improving the predictive ability of Earth system models.

Predicting the functional attributes and state of ecosystems is a foundational and pressing aim in ecology. To address this challenge while accommodating the complexity of ecosystems early dynamic global vegetation models¹ focused on seed plants and abstracted the circa 350,000 species of seed plants into a small (<50) number of plant functional types (PFTs)². These functional types represented hypotheses of broad categories of response types to abiotic forcing and biotic interactions. The appeal of PFTs lies in their promise that a mechanistic understanding of growth form and leaf economic traits might provide an appropriate foundation for simulating the essential dynamics of ecosystems without needing to parameterize species-level variation.

The PFT approach still underpins the representation of the land surface in current generation Earth system models³ and usage of the biome concept in ecology^{4,5}. Despite its widespread adoption, the PFT paradigm has delivered only moderate predictive success⁶. Understanding this lack of success involves reexamining two assumptions that underpin the use of PFTs in Earth system models. First, it is assumed that evolutionary convergence is sufficiently widespread that unrelated species in similar environments evolve similar PFT syndromes⁴. Second, it assumes that these PFTs

can be meaningfully defined and operationalised for use in models. Neither assumption holds universally.

An alternative approach is to acknowledge that niche conservatism, defined as the tendency of species to retain ancestral ecological traits⁷, may play a dominant role in structuring plant communities. If niche conservatism were dominant, communities would be assembled primarily from pre-adapted lineages, leading to functional similarity driven by shared evolutionary history. In contrast, if niche convergence were dominant, communities would be composed of species that have independently evolved similar functional traits in situ, resulting in functional similarity that is decoupled from phylogeny. Implicit in both cases, is that community assembly is constrained by limiting similarity, which precludes the coexistence of functionally identical species locally. In reality, the traits of species in local communities reflect the combined influence of niche conservatism and niche convergence as community assembly is highly complex⁸. Although this tension between niche conservatism and niche convergence is well studied in phylogenetic community ecology^{9,10}, the consequences for Earth system modelling have not been thoroughly explored. If niche conservatism were a dominant force, molecular phylogenies could be used to

identify functional groups, side-stepping the difficulties associated with defining PFTs using traits.

Although Earth system models have almost exclusively represented vegetation using PFTs under the evolutionary convergence assumption, the fields of vegetation science, conservation biology and biogeography acknowledge niche conservatism by using floristic criteria to make ecosystem maps. These maps, often termed regionalisations¹¹, explicitly acknowledge the importance of niche conservatism, dispersal limitations and other historical contingencies. For example, Daru et al. showed the potential of using phylogenetic information to cluster functionally relevant units in geographic space in a previous study in Africa¹². However, regionalisation studies are usually directed toward identifying centres of endemism, describing evolutionary radiations, and setting conservation priorities, and thus remain largely disconnected from the predictive agendas of functional ecology, global change ecology and Earth system modelling.

In this study, we pose the hypothesis that a priori defined phylogenetic groups can provide more information for the classification and mapping of the distribution of ecosystems than PFT-based methods. Although the use of phylogenetic relatedness as a proxy for functional relatedness comes with caveats, in systems where a few taxa dominate, phylogenetic relatedness may associate with functional attributes¹³. To explore this idea we use the concept of a phytoclimate, a region where climate favours the growth of similar combinations of plant types¹⁴, to examine whether phylogenetically defined plant types can provide useful classifications of ecosystems. We use African savannas as a tractable case study to illustrate the promise of this idea. Informed by the literature on savanna types^{15–19}, we use (paraphyletic) indicator taxonomic groups, that is members of selected families, subfamilies and genera.

The case study

African savannas are defined as ecosystems with a grassy understorey of C4 grass species and a sparse overstorey of shade intolerant trees²⁰. Ecosystems fitting this definition take on many faces: the grass-rich Serengeti, the southern African Miombo woodlands, the Guinean savanna belt and the central Kalahari all fit this savanna definition. Linked to this coherent definition is a solid theoretical understanding of how savannas work²¹; however, there is less consensus on how to classify the functional differences between savannas, manifested by the variety of schemes that have been proposed^{15,17,19,22–28}.

Classifications of African savannas can be split into floristic^{15–17,19} and functional^{23,25,28,29}, but closer inspection reveals that the floristic and functional attributes are often confounded. For example, the sweet- and sourveld scheme, which is a functional concept motivated by agronomic aims²³, uses taxonomically and PFT-based groups. Sweetveld entails fast-open nutrient cycling, high palatability, low plant biomass and is typically associated with a balanced grass taxonomic composition and the taxon *Vachellia* (and the paraphyletic *Senegalia*) dominating the tree layer. Sourveld, in contrast entails slow-closed nutrient cycling, low palatability, high plant biomass, where the typical savanna grass taxa Aristidoideae and Chloridoideae are absent and Combretaceae and Detarioideae dominate the tree layer^{17,25,28}. The distribution of sweetveld or sourveld has been linked to external drivers such as climate and soil fertility^{25,28,30–33}. For example, sweetveld is associated with dry climates and fertile soils, floristically, the grass layer is dominated by Chloridoideae whose usage of the NAD-ME C4 pathway has been linked to high nutrient requirements and a tolerance of aridity^{16,17,28,34}. Typical sweetveld tree taxa often form nitrogen-fixing relationships with rhizobial bacteria (e.g. *Acacia* s.l.: *Vachellia* and *Senegalia*) often interpreted as an adaptation to soils high in base minerals but low in nitrogen^{35,36}. By contrast, sourveld is more common in high rainfall, nutrient poor environments and is characterised by Panicoideae, which predominantly use the NADP-ME C4 pathway. High rainfall is typically associated with leached soils, which selects for tree taxa that enter into ectomycorrhizal relationships (Detarioideae), or in low base mineral situations, with arbuscular mycorrhiza (Combretaceae)^{16,34,36}. Sweetveld - sourveld is not only linked to climate and soils, but also to differences in the internal dynamics associated with fire and

herbivory^{25,37,38}. For example, sweetveld has been interpreted as meso-herbivore controlled (eutrophic savanna) and sourveld as fire-controlled (dystrophic savanna)^{25,37}. Extremely arid environments tend to fall outside the sweetveld-sourveld continuum: here the grass layer is characterized by Aristidoideae, which rely largely on the PCK C4 pathway, which is regarded as an adaptation to extremely arid and nutrient-poor habitats. Characteristic trees in such arid conditions are members of the genus *Commiphora*.

The literature reviewed in the previous paragraph reveals a confounding of functional and phylogenetic attributes of sweet- and sourveld savannas, which suggest that these ecosystems are structured by ecological trait filtering rather than recent convergent evolutionary adaptation. In particular, the dominance of a small number of taxonomic groups with distinctive functional profiles implies that these taxa were pre-adapted to the savanna environment, which only emerged in the late Miocene (11 - 5 million years ago)³⁹, too recent for maladapted taxa to evolve equivalent adaptations or to overcome the priority advantage of the pre-adapted groups^{40,41}. Thus, niche conservatism of the indicator savanna taxa may be the dominant structuring factor in African savannas, which may make phylogenetic information a better proxy for function than traits or life forms. This then leads us to hypothesize that savanna types in Africa can accurately be classified by defining the climatic preferences of the dominant indicator savanna taxa.

To achieve this classification we focussed on the seven taxonomic groups identified in the previous paragraphs, namely *Vachellia*, *Commiphora*, Combretaceae, Detarioideae, Panicoideae, Chloridoideae, Aristidoideae. For these groups we compiled a data set of occurrence data for African savanna regions using a combination of presence-only and presence-absence data sources. We then used an existing workflow⁴² to describe the climatic preferences of each of these taxonomic groups. This phytoclimatic workflow involved fitting a species distribution model^{43,44} to the occurrence data for 1255 species-level occurrence data sets, using the average modelled suitability of the species in each of the seven taxonomic groups to generate taxonomic group suitability surfaces, and finally clustering sites with similar values across the seven surfaces to generate ecosystem classes. We refer to these classes as phytoclimates¹⁴ to distinguish the resulting map from biome maps or floristic regionalisations. The phytoclimate map for Africa based on this classification was formally compared to White's vegetation map for Africa¹⁹. In post-hoc analyses we used recursive partitioning to elucidate which climatic and phytoclimatic variables define the savanna types. We further assessed the extent to which the taxon-based phytoclimate map can predict the functional properties of different savanna types.

Results

Environmental suitability mapping

Plant species distribution models were produced for 1255 tree and grass species from seven defining savanna taxonomic groups (Supplementary Table S1). The majority (99.6%) of models displayed acceptable predictive accuracy (AUC > 0.75). The suitability surfaces generated by stacking and averaging the species distribution model suitability projections by taxonomic group displayed high suitability for all seven groups in the greater savanna region of Africa (Fig. 1). This showed that our species and taxa selections yielded models that represent plausible hypotheses of how climatic factors constrain the potential distribution of major savanna tree and grass taxonomic groups.

Savanna classification

The unsupervised classification of the seven suitability surfaces yielded a classification with seven phytoclimates (Fig. 2). A post-hoc recursive partitioning was used to interpret the classes (Fig. 3). The Miombo phytoclimates 4 and 7 were defined by high Detarioideae or Combretaceae suitability, while the arid sweetveld phytoclimates 2 and 3 were connected to a low suitability for Panicoideae (Fig. 3). Regionally, the west African phytoclimates 3 and 5 could be split from the rest according to their low suitability for *Vachellia* (Fig. 3). Phytoclimate 6 was somewhat taxonomically undifferentiated with its

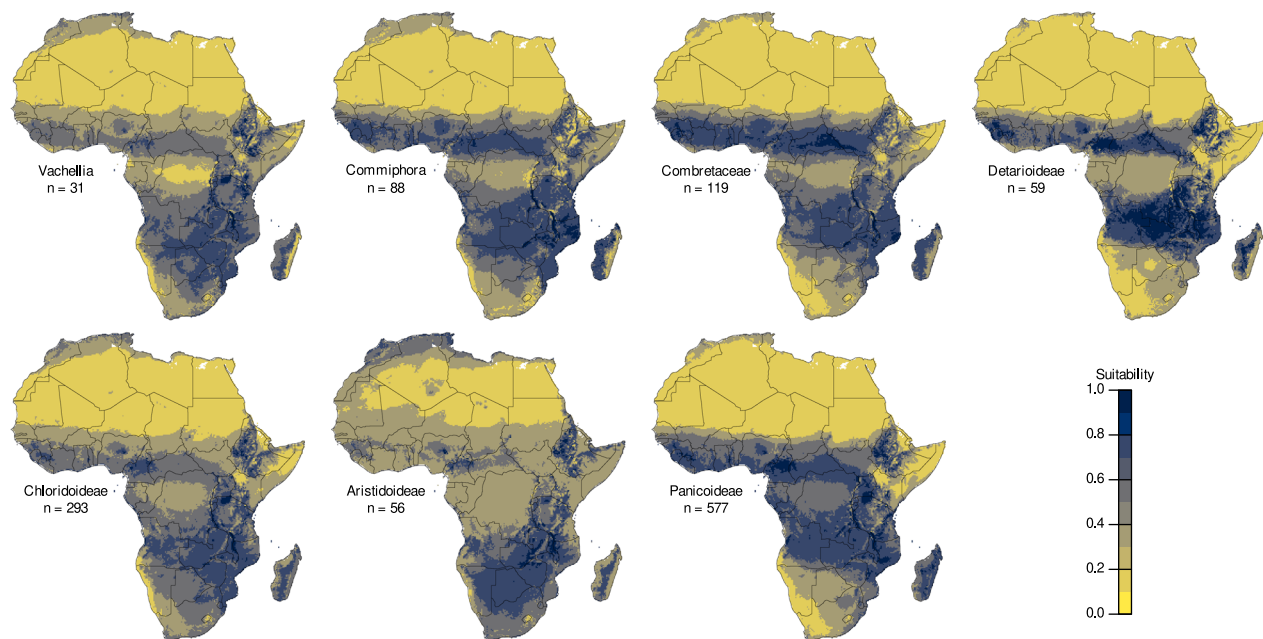


Fig. 1 | Geographic distribution of climatic suitability for seven dominant African savanna plant taxonomic groups. Suitability is calculated as the mean suitability of species belonging to each taxonomic group as estimated by a plant

growth model. A total of 1248 species models were used, *n* indicates the number of species models underlying each taxonomic group's climatic suitability surface.

most distinctive trait being moderate *Vachellia* suitability and high suitability for *Commiphora* (Fig. 3).

Post-hoc recursive partitioning of the phytoclimate classes using effective rainfall and mean annual air temperature revealed that climate could explain a substantial portion of the phytoclimate classification (cross-validated error rate of 0.34 for seven classes, Fig. 3). The major division was between warm and cold phytoclimates (west African and southern-east African), while the remaining divisions were primarily related to rainfall (Fig. 3). Phytoclimates 2 and 3 were associated with very dry-cold and very dry-very hot climates, respectively (Fig. 3). Phytoclimate 4 and 7 were associated with high rainfall irrespective of air temperature (Fig. 3).

Formal map comparison

To formally compare the maps, we used an algorithm that harmonizes the number of classes on the maps and then calculates Cohen's kappa (κ) for each pairing of classes⁴⁵. Overall, 64.2% of pixels in the phytoclimate map agreed better than poor and 54.2% displayed good agreement with the White map ($0.75 > \kappa > 0.4$; Fig. 2). About 10% of the pixels agreed very well (9.9% ; $\kappa \geq 0.75$) with White's map (Fig. 2). Phytoclimate 6 was the only one with low agreement (Fig. 2). In contrast, a growth-form based phytoclimate map of Africa⁴² generated consistently lower κ statistics; in this case 0% of pixels displayed very good agreement, while 22.6% showed good agreement (Supplementary Fig. S2).

Functional analysis of the classes

The functional value of the phytoclimate scheme is further evidenced by differences in their functional attributes (Fig. 4). In particular, the west African savanna phytoclimates (3, 4 and 5) had a lower net primary productivity (NPP) than the other phytoclimates (1, 2, 6 and 7; Fig. 4) and there was a gradual trend towards higher NPP from sweet to sour phytoclimates (Fig. 4). Regarding fire activity, the sweet phytoclimates (2 and 3) had a low burnt area and a high coefficient of variation of burnt area, indicating infrequent fire return intervals (Fig. 4). The phytoclimates intermediate on the sweet- and sourveld gradient showed lower fire activity than the sour phytoclimates (4 and 7), but higher fire activity than the sweet phytoclimates (2

and 3) (Fig. 4). The phytoclimates also differed in the density of the major livestock species (cattle) they supported; however, there was no simple trend of sweetveld phytoclimates supporting higher cattle densities than sourveld phytoclimates (Fig. 4).

A fundamental motivation for classifying ecosystems is to use the classifications to predict ecosystem properties and we therefore tested whether the taxa-based classification could explain more variance in the ecosystem properties shown in Fig. 4 than an existing growth form-based classification⁴². This comparison indicated that the predictive performance of the taxa-based classification was superior across all four ecosystem properties investigated. This pattern was evident using both the WAIC and LOO criteria and the differences detected were well above thresholds typically considered decisive (Supplementary Tables S2, S3). Gelman's R2 measure for absolute goodness-of-fit indicated that our model was able to explain a larger portion of the observed variance (Supplementary Table S4). Further, in a comparison with the floristic map of White¹⁹, both models performed equally well in describing variance in Net Primary Production and cattle density, however the White model described variance in both burnt area and the coefficient of variation of burnt area better than the taxa-based classification (Supplementary Tables S5, S6). The model using White's map also improves the variance explained assessed with Gelman's R2 (Supplementary Table S4). Here, White's map with 26 classes is much more flexible than our map with seven classes, which could explain the slightly improved goodness-of-fit.

Discussion

Our phytoclimate mapping of African savannas based on a set of a priori identified savanna plant taxonomic groups showed strong agreement with established vegetation mappings^{19,46,47}. It also aligns with a recent study that used tree composition data to show that the composition of African savannas are strongly structured by phylogeny⁴⁸. This suggests that phylogenetically based groups are a viable alternative to plant functional trait-based groups for ecosystem classification. This study used dominant indicator taxa as a simplified proxy for phylogenetic relatedness to illustrate the promise of this approach. Furthermore, recursive partitioning elucidated that the climatic variables, effective

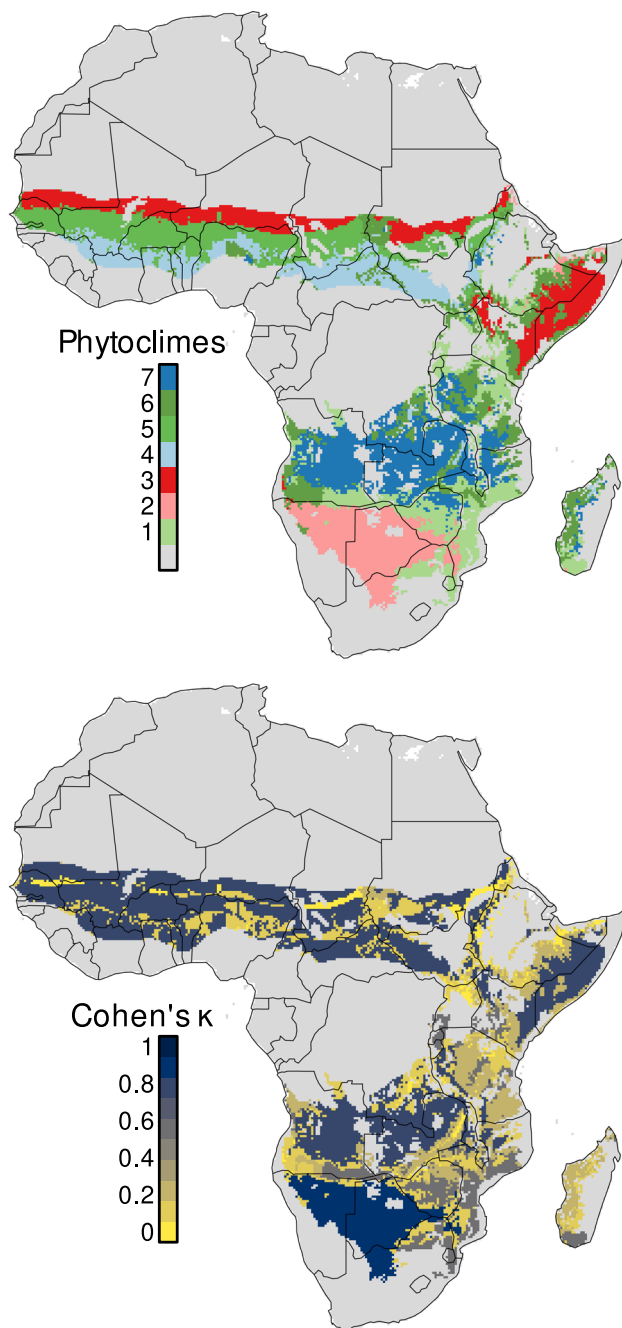


Fig. 2 | African savanna phytoclimes defined using an unsupervised classification of the taxonomic group suitability surfaces shown in Fig. 1 and the agreement between the phytoclime map and White's vegetation map of Africa (Supplementary Fig. S1) estimated using Cohen's κ . Existing conventions⁴⁶ interpret $\kappa \leq 0.40$ as poor agreement, $0.40 < \kappa < 0.75$ as good agreement and $\kappa \geq 0.75$ as very good agreement. Grey areas on both maps indicate non-savanna areas. Supplementary Fig. S2 shows the κ statistics when contrasting the White map to a published growth-form based phytoclime map⁴². For maps with different number of phytoclimes (2–10) see Supplementary Fig. S8.

rainfall and air temperature, were sufficient to explain the majority of the variation in the phytoclime classification. Our findings are supported by previous work^{17,28,48} that emphasized climatic factors as determinants of savanna types. Further, the value of our map is that it is entirely based on species distribution data and climatic forcing, meaning that the models can be used to project shifts in these phytoclimes and their associated functional properties with changing climates.

The suitability surfaces that underlie the phytoclime map revealed that of the sourveld taxonomic groups (Combretaceae and Detarioideae), Combretaceae had a higher suitability in the Kalahari, and a much higher suitability in the South African lowveld and the non-Miombo savannas of east Africa whereas the Detarioideae suitability was highest in the core Miombo regions, which is in agreement with expectations from the literature^{18,19,23,47–51}. *Commiphora* suitability was high in the Sahel zone and horn of Africa, matching descriptions in previous work^{17,19,47,48}. In east Africa, *Vachellia* suitability was highest in areas where on volcanic geology from the great African rift system (cf. ref. 25), matching³².

For the grasses, the suitability surfaces for Panicoideae, Aristidoideae and Chloridoideae aligned with a bioclimatic scheme proposed by Johnson and Tothill¹⁶. Our profiles also agreed with empirical climate-abundance relationships revealed by Forrester et al.⁵² for South Africa. Aristidoideae and Chloridoideae displayed lower suitability in west Africa than southern-east Africa, which is consistent with Rattray's African grass cover maps⁵³.

Our phytoclime mapping showed substantial agreement with the established UNESCO vegetation map of Africa by White¹⁹, where 54.2% of pixels displayed good agreement with the White map ($0.75 > \kappa > 0.4$). In comparison to a purely functionally based mapping, our phylogenetically based approach showed better agreement with the White map when assessed using κ statistics. This suggested that the climatic niches of the dominant indicator taxa better explain the variation of savanna types than a growth form based approach⁴². It also added to the discussion that climate is likely a determinant of savanna distribution^{25,32,36,46,54}. To put the agreement between the phytoclime map and the White map into context, it is illustrative to benchmark our κ statistics against those reported in a comparison of global biome maps⁴⁵. When comparing global maps one might expect high consensus, since there is, in general, agreement on what a tropical forest, tundra, desert or temperate forest is. Nonetheless, this benchmark study⁴⁵ reported very good agreement ($\kappa \geq 0.75$) in 0% of the pixels and good agreement ($0.40 < \kappa < 0.75$) in only 19.3% of the map pixels.

The idea that soil nutrient status directly determines savanna types, or indirectly through its influence on herbivory^{25,28,30,32,34,36,46,51,54} or on fire^{24,37,54,55}, has been widely discussed. Indeed, the importance of fire and herbivores as determinants of savanna types is now embedded in the functionally based IUCN ecosystem typology⁵⁶ which recognizes trophic (herbivore controlled) and pyric (fire controlled) types. In this typology fire dominates in high rainfall, low nutrient, high grass biomass savannas, while herbivore impacts dominate in higher nutrient systems where canopy-openness limits predator success^{25,37}. We propose that we can reconcile this apparent contradiction between climatic, consumer (fire, herbivory) and edaphic (soil fertility) determinants by accepting that each of this studies climatically determined savanna types (savanna phytoclimes) support distinct ecologies, each characterised by distinct pathways of biomass production and consumption. Acocks's²³ agronomically motivated sweet- and sourveld provides a useful framework for this reconciliation allowing us to arrange the seven African savanna phytoclimes along a sweetveld-sourveld gradient (Fig. 3). The Miombo-adjacent phytoclimes 4 and 7 can be interpreted as sourveld, phytoclimes 1, 5 and 6 as mixed veld, while phytoclimes 2 and 3 can be interpreted as sweetveld. Along this functional gradient, relative *Vachellia*, Chloridoideae and Aristidoideae suitability increased towards the sweet end, while relative Detarioideae and Panicoideae suitability increased towards the sour end.

Despite the phytoclime map being able to reproduce important distinctions apparent in the White¹⁹ map, some deficiencies were apparent. First, the mopaneveld category on the White map fell into two phytoclimes, causing a low agreement statistic in one of the two. This disagreement is because phytoclimes are climatic units (zonobiomes to use Walter's terminology⁵⁷), whereas White's vegetation types used biogeographic and floristic classification criteria. Mopaneveld is in Walter's terminology azonal, covering an intersection between alluvial soils and frost free climates^{19,58}. Second, in east Africa White's *Acacia-Commiphora* bushveld¹⁹ did not align with our Phytoclimes. White noted that this formation included substantial functional, taxonomic and environmental variation in addition to azonal

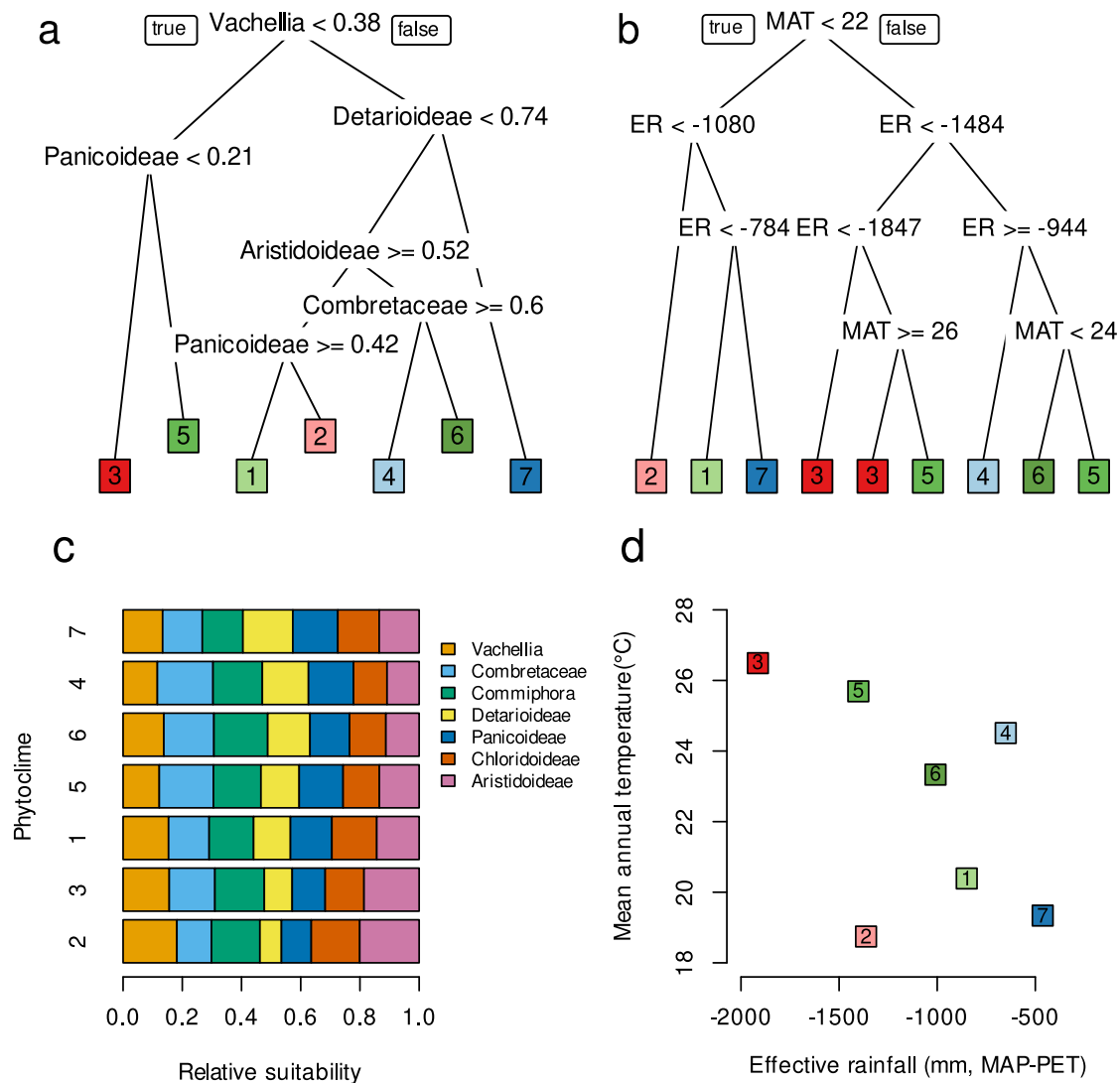


Fig. 3 | Decision trees based on recursive partitioning revealing how (a) plant taxonomic group and (b) climatic variables determine phytoclimate type. MAT indicates mean annual temperature (°C) and ER indicates effective rainfall (mm, ER = mean annual precipitation - Penman potential evapotranspiration). Displayed

below are (c) the distribution of relative taxonomic group suitability in each phytoclimate and (d) the mean position of each phytoclimate in a mean annual temperature and effective rainfall biplot.

vegetation caused by land-use¹⁹. White assigned southern African savannas to Miombo-, Mopane-, or “all other savannas”¹⁹ suggesting units that disagree with the phytoclimate map. Other authors suggest that White’s “all other savannas” category could be split into *Combretaceae* dominated and *Acacia* s.l. dominated savanna^{17,18,34}. The disagreement in east and southern Africa may be related to the complexity, diversity and lack of clear indicator taxonomic groups for trees and grasses in these regions^{47,48}. East Africa, for instance, seems to be differentiated from other savannas by their abundance of *Euphorbiaceae* succulents and *Rutaceae* and *Primulaceae* shrubs⁴⁸, which has not been captured by the historical classifications motivated by tree- and grass-differences in savannas^{17,19,50}.

The taxa-based phytoclimate classes developed here were able to partition significant components of the variance in the functional attributes of the study area. The partitioning of variance in fire activity with phytoclimates parallels distinctions made in a previous study²⁴, which classified large parts of phytoclimate 4 and 7 into a “frequent-cool-small” fire activity class, while the other phytoclimates generally fell into the “rare” fire activity class. The picture was more complex when predicting supported herbivore biomass, which we attempted by using cattle density data provided by the FAO⁵⁹. The caveat here is that cattle density is not a perfect proxy of the forage value of

an ecosystem, since herbivores with high body mass can tolerate lower quality forage^{25,34}. Also, supplemental feeding can lead to higher cattle numbers than what would be supported normally. Wild herbivore data products suggest that our sourveld phytoclimates (phytoclimates 4 and 7) were associated with a large bodied (bulk-feeder) herbivore community and our sweetveld phytoclimates (phytoclimates 2 and 3) were associated with an arid gazelle herbivore community²⁶. However, this alignment is circular since these wild herbivore data products²⁶ used both rainfall and White’s vegetation types to infer herbivore biomass.

Of broad significance is that this phytoclimate classification is underpinned by indicator taxa, suggesting that the phylogenetic and functional groups are confounded in African savannas. Indeed, floristic studies of African savannas suggest that tree community compositions are the outcome of phylogenetic filtering and not dispersal limitation⁴⁸. This leads us to speculate that African savannas are dominated by a small set of taxonomic groups that were pre-adapted to savannas and radiated within the savanna environment to establish a priority effect that has slowed radiation and adaptation of other lineages in the savanna climate space. Analogous results were found by Ackerly in a more local study in the Californian chaparral⁹. This contrasts with evidence from other savanna systems such as the South

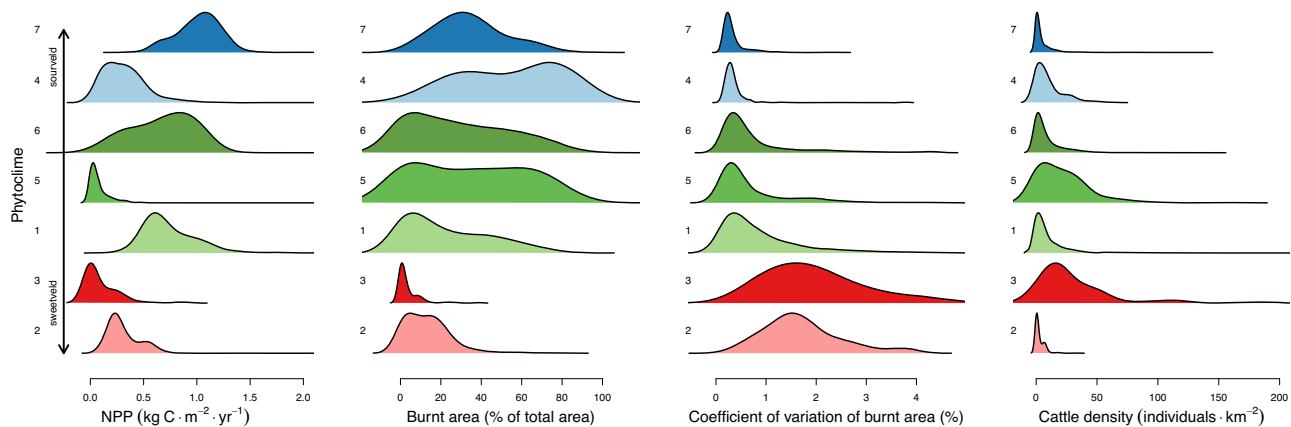


Fig. 4 | Kernel density estimates of functional attributes of each phytoclimate. The phytoclimates are arranged on a sweetveld to sourveld gradient. Supplementary Fig. S3 shows an analogous plot using a published growth - form based phytoclimate map⁴².

American Cerrado, which showed rapid, in situ, convergent evolution of fire-related traits⁶⁰. This study uses a simplified phylogenetic proxy for functionally relevant plant types and although it illustrates the potential of this approach, we acknowledge that further work is needed to expand this framework by potentially including other taxonomic groups that have been recognized to be indicators of different savannas^{47,48}. Future studies should also explore phylogenetic relatedness explicitly by considering phylogenetic distance when fitting the species distribution models and building ecosystem-type classifications.

A more general challenge is to develop functional classifications that explicitly accommodate the tension between phylogenetic conservatism and evolutionary convergence. Currently, plant functional type and biome concepts assume that evolutionary convergence overwhelms phylogenetic niche conservatism and therefore that phylogenetic and biogeographical context can be ignored when forecasting Earth system dynamics. However, the idea of biome level convergence as an organizing principle for predicting global ecosystem properties has been questioned before, both in general⁶¹ and specifically for savannas³⁵. As ecosystems reassemble under climate change⁶², it is likely that their near-term functional properties will be determined more by the traits of the taxonomic groups pre-adapted to these new climatic regimes than by the in situ convergent evolution of new adaptations. Our findings reinforce the view that biome-level convergence is not a universal organizing principle, and highlight the need to integrate phylogenetic context when predicting ecosystem properties under global change.

Methods

Approach

We identified phytoclimates - regions where climate favours the growth of similar combinations of plant types¹⁴, using plant types that are taxonomically defined. The foundation of this approach is laid by the assumption that we can use species distribution data to reveal the environmental preferences of different plant types⁴². We used the existing literature^{15–19} to identify meaningful plant types. As motivated in the main text we used seven savanna taxa to define the types. We use the term taxa here to refer to paraphyletic assemblages of species that belong to a specific taxon (family, subfamily or genus) and occur in African savannas. The taxa were *Vachellia* (Fabaceae); the genus *Commiphora* (Burseraceae); the family Combretaceae; the subfamily Detarioideae (Fabaceae); the subfamilies Panicoideae, Chloridoideae and Aristidoideae of the Poaceae. The workflow involved first fitting a species distribution model to all species with sufficient data of each group and then averaging the suitability scores of all member species at each location. The seven resulting taxonomic group suitability surfaces were then clustered into classes, and these classes represent the

phytoclimates. In the next sections we describe the details. All analyses were performed using the statistical software environment R⁶³, while maps were created using the R package terra⁶⁴.

Plant growth model for species distribution modelling

In principle, any species distribution model can be used to estimate the climatic suitability of geographic locations for the plant groups of interest. We have chosen a process-based plant growth model⁶⁵ as implemented in the R package TTR . PGM⁴⁴ that infers how climatic factors define the physiological limits of plant distributions. The underlying physiological model⁴³ simulates Carbon (C) and Nitrogen (N) allocation between shoot and root and how these substrate pools determine biomass growth. Physiological rates in the model (growth, C and N assimilation) are constrained by the environment; and it is assumed that these constraints are species specific⁶⁵. We used a 15 parameter version of the OAK variant of the model⁴⁴.

Environmental data for the plant growth model

The forcing data for fitting the physiological model was derived from a range of 1 km² data products. The CHELSA 2.1 ambient dataset (1981–2010)⁶⁶ was used for monthly averages of daily mean, minimum and maximum air temperature, solar radiation, potential evapotranspiration and precipitation. A soil water balance model included in the TTR . PGM package was used to calculate the monthly soil water content using the CHELSA 2.1 precipitation, radiation and potential evapotranspiration data along with soil wilting point and field capacity data⁶⁷. A Farquhar type photosynthesis model⁶⁸ included in the TTR . PGM package was used to calculate C3 and C4 photosynthesis rates as forced by atmospheric CO₂, air temperature, and solar radiation. The atmospheric CO₂ concentration was assumed to be a constant 338 ppm.

Species data for the plant growth model

We used presence- and absence geographical data of 1255 savanna species to estimate species-level physiological parameters in the model. The species data covered the African continent with small gaps (Supplementary Fig. S4). Species were chosen as follows: A species list provided for the African continent by BIEN⁶⁹ was filtered for those species that belonged to the selected African savanna taxonomic groups (i.e., *Vachellia*; Detarioideae; Combretaceae; *Commiphora*; Panicoideae; Chloridoideae; Aristidoideae). Further, species had to have at least two occurrences in the savanna region of the White map¹⁹ (Supplementary Fig. S1). The savanna region in Africa as outlined by White¹⁹ includes, in White's terminology, woodland, wooded grassland and bushland (map units 22, 25–32, 35–37 & 40–48). True forest species in Detarioideae and Combretaceae were further filtered out using the AFROTROPTREE dataset⁷⁰ if they were present in a plot delimited in that

dataset as “forest” as our goal was to specifically determine savanna taxonomic group suitability. We acknowledge that the taxonomic groups used in our study do not exclusively occur in African savannas. Some heavily cultivated species were preemptively removed from the species list (*Zea mays*, *Sorghum* sp., *Tamarindus indica*). Climbers in Combretaceae were excluded using the BIEN trait database^{69,71} and the RAINBIO database⁷². Taxonomic resolution was performed using the taxonomic name resolution service (TNRS) provided by BIEN⁷³. We originally considered *Acacia* s.l. (including *Senegalia* next to *Vachellia*) but in the final analyses use only *Vachellia* as results were qualitatively similar and *Acacia* s.l. is not monophyletic.

The occurrence data for each model fit was a geographically homogeneous sample of ca. 100 presence, ca. 100 true absence and ca. 100 pseudo-absence points to inversely estimate physiological response curves for each species. The three different data types were handled as follows: Presence locations for each species were derived from geo-referenced vouchers specimens from GBIF^{74,75}, and the RAINBIO database for Africa⁷². Records were filtered using the R Package *CoordinateCleaner*⁷⁶, removing records close to country centroids, capitals, biological institutions or botanical gardens. Presence data was spatially thinned to a target number of 100 using the R package *SpThin*⁷⁷. Next to spatial thinning, we also performed a post-hoc analysis to ensure our model residuals are not substantially autocorrelated (Supplementary Fig. S5). For this, we calculated Moran's I with the R package *terra*⁶⁴ for the Pearson residuals of all 1255 model fits, where Pearson residuals were calculated using Eq. (1):

$$r_i = \frac{y - \hat{y}}{\sqrt{\hat{y} \times (1 - \hat{y})}} \quad (1)$$

where y was the observed presence-absence data for a fit i and $\hat{y}$ was the predicted probability of occurrence.

Plot data was used to further add presence locations to the presence-only data. Additionally, plot data was used as true absence data for a species in locations where it was missing. Plot data used in this study was: the AFROTROPIC database⁷⁰, which is a collection of species lists from 753 afro-tropical forest and savanna sites; National Vegetation Database of South Africa (NVD) and associated ACKDAT database, which is a digitization of South African vegetation mapping efforts spearheaded by John Acocks^{23,78,79}, the West African Vegetation database⁸⁰ by Senckenberg BiK-F that provided around 1200 locations in west Africa, originally stemming from the SUN and BIOTA projects; the FLOTROP database⁸¹, which contributed around 11,000 locations from the Sudano-Sahel area of Africa. The West African Vegetation and FLOTROP databases were accessed via GBIF⁷⁴, these occurrences were filtered for full inventory surveys and species lists created for each unique geolocation. The providers of both databases noted that some surveys only included woody species, herbaceous species or both. This information was not explicitly given in the database survey details. Therefore, plots from both databases were classified as “tree only”, “herb only” or “both” using life form data from the BIEN^{69,71} and RAINBIO databases⁷². This classification was then used to reliably derive absence data for each life form, e.g., omitting “herb only” locations for tree species absence sampling. This pool of absence locations was then used to sample true absence locations for each species. Additionally, pseudo-absence points were created. The creation of pseudo-absence was stratified such that Koepfen-Geiger climate zones not yet represented in true presence and absence points were sampled. This ensured that the entire global environmental range was represented in the sample. The sampling of absence points was structured so that the final data set for a species consisted of 1/3 of each presence, true-absence, and pseudo-absence data points.

Parametrization of the plant growth model

Models were only attempted if the previous step produced data sets with 5 or more presence records (i.e. 15 or more presence-absence points). 97.4% of species fits had more than 5 and 89.0% of fits more than 10 presence data points (and thus more than 15 or 30 total data points, respectively). Species data sources are documented in Supplementary Table S7. Photosynthetic

parameters of the Farquhar-type submodel were assumed to be constant across species save for photosynthetic type (C_3 vs. C_4). All tree species were assumed to be C_3 , while literature was consulted to determine photosynthetic type for grasses⁸².

We generated initial parameter estimates using the differential evolution genetic algorithm (as implemented in the R package *DEoptim*⁸³). These parameter estimates were used to initiate an MCMC-based algorithm (random walk Metropolis) using the R package *LaplacesDemon*⁸⁴. The posterior parameter estimates were then used for a projection of the model-estimated probability of occurrence for a given species in geographic space. We generated species distribution models for 1255 species. Modelled geographic distributions were evaluated with measures of predictive accuracy, the area under curve (AUC) and the true skill statistic (TSS; Supplementary Table S1). Typical threshold values for identifying poor fits from studies using the TTR . PGM for phytoclimate identification were adapted from previous work^{14,42}. These studies used AUC thresholds of >0.75 and TSS thresholds of >0.7. Using the AUC threshold, 99.6% of attempted fits were retained, whereas only 78.9% are retained under the TSS threshold. We used the AUC threshold because our subsequent steps in the analysis were not sensitive to this decision (Supplementary Fig. S6).

Suitability and phytoclimate determination

The remaining 1248 range maps were overlaid and averaged to create one geographic suitability surface per phylogenetic group. The suitability surfaces were then normalized by dividing the values of all locations by the maximum suitability value, thus creating a relative suitability surface for each group ranging between 0 and 1. Phytoclimates were identified by applying an unsupervised model-based clustering algorithm implemented in the R package *mcclust*⁸⁵ to the seven suitability surfaces, i.e., clustering sites according to the site $\times$ taxonomic group suitability matrix for the seven taxonomic groups. The geographic region used in this classification was restricted to the savanna region delimited by White¹⁹. A moderate number of clusters, seven, was selected to balance goodness-of-fit of the clustering model and ecological interpretability (see Supplementary Fig. S7 for goodness-of-fit for 1–30 clusters). For transparency, we provide phytoclimate maps for cluster numbers 2–10, as there is no substantial improvement of model fit for more than 10 clusters (Supplementary Fig. S8).

Formal map comparison

The phytoclimate map was quantitatively compared to the White map (ref. 19, Supplementary Fig. S1). We achieved this by following a protocol for formal comparison of two maps with differing categories as described in ref. 45. The “major vegetation types and mosaics”-categories from the savanna region of White's map (cf. Supplementary Table S8) were used. What constituted the savanna region on White's map for this analysis is described in the section “Species data”. The number of categories on the maps were harmonized by computing the overlap of mapping categories on a pixel level and merging categories that shared the same majority of overlap (see Supplementary Table S8 and ref. 45 for details). A contingency table was then computed (see ref. 45). Cohen's kappa (κ), a measure of agreement, was calculated for each pair of map categories using Eq. (2):

$$\kappa = \frac{p_{ij} - p_i p_j}{\frac{p_i + p_j}{2} - p_i p_j} \quad (2)$$

where p_{ij} is the observed fraction of pixels in which category i (from map 1) and category j (from map 2) coincide. The terms p_i and p_j are the proportions of the respective maps occupied by categories i and j . In words, the numerator expresses how much the observed coincidence (p_{ij}) exceeds the coincidence expected by chance ($p_i p_j$). The denominator scales this difference by the maximum possible coincidence given the category prevalences ($\frac{p_i + p_j}{2}$), adjusted for the chance expectation. Then agreement was mapped by displaying Cohen's κ in geographic space for each pairing of categories at a given pixel. Cohen's κ gives a value ranging between -1 and 1 where, according to existing conventions⁸⁶, $\kappa \leq 0.40$ indicates poor agreement,

$0.40 < \kappa < 0.75$ indicates good agreement, and $\kappa \geq 0.75$ indicates very good agreement. The same workflow was also applied to compare the agreement of a growth-form-defined mapping of African biomes⁴² with the White map (see Supplementary Table S9).

Post-hoc analyses

We analysed the phytoclimate outputs using recursive partitioning as implemented in the R package `rpart`⁸⁷. We produced a classification tree that predicted phytoclimates from mean annual temperature (°C) and effective rainfall (mm) and a second tree that predicted phytoclimates from the savanna taxonomic group suitability surfaces generated from the species distribution models (Fig. 1). The recursive partitioning produces a classification tree for the given dataset and finds a solution in the predictor matrix (site $\times$ climate or site $\times$ taxonomic suitability matrix) for each split that best reduces the Gini impurity⁸⁷. Effective rainfall was defined as mean annual rainfall - evapotranspiration. For this analysis, evapotranspiration and temperature data were taken from CHELSA 2.1⁶⁶. The control parameter in `rpart` was manually adjusted for each fit to yield acceptable cross-validated error rates. The R package `rpart.plot`⁸⁸ was used to plot the resulting regression trees.

Functional attributes of the phytoclimates

To explore functional differences between the phytoclimates we examined their net primary production, fire activity and ability to support herbivores. Net primary productivity (NPP) was extracted from the annual MODIS NPP data product (MOD17A3HGF⁸⁹) using data from 2001 to 2023. For NPP, only protected areas were considered to exclude effects of land-use. Information on protected areas in Africa was taken from the World protected area database⁹⁰. NPP values above 3.0 were excluded, as they were considered ecologically unrealistic for savannas and to be artifacts of land-use or the remote sensing analysis. The density of the major livestock species in Africa (cattle, *Bos taurus*) was extracted from the UN-FAO gridded livestock census data product of 2020⁵⁹. This was taken as a proxy of how much livestock herbivory is supported by the different savanna types, which is expected to be a function of forage amount and forage quality. Fire activity was compared using mean annual burnt area and the coefficient of variation of annual burnt area (as an indicator of return interval) from the Global Fire Emissions dataset 4.1⁹¹. This dataset offers burnt area [% of pixel] on an annual temporal resolution with coverage of 19 years between 1997 and 2015. These data products were sampled at the resolution of the phytoclimate map (0.25×0.25 degrees) for each phytoclimate.

The ability to predict variation in these ecosystem properties using the taxa-based phytoclimate classification was formally analysed and compared to an existing growth-form-based phytoclimate classification⁴², as well as the floristic map published by White¹⁹ (Supplementary Fig. S3, Supplementary Tables S2, S3, S5, S6). We fitted Bayesian linear models in JAGS⁹², with the ecosystem metrics as response variables and the phytoclimate classes as predictor variables. Priors were uninformed uniform. The predictive performance of the models was evaluated using the Widely Applicable Information Criterion (WAIC) and leave-one-out cross-validation (LOO) as implemented in the R package `loo`⁹³. Absolute goodness-of-fit was further assessed using the Gelman's R2 statistic⁹⁴. Both WAIC and LOO provide measures of out-of-sample predictive accuracy that account for model complexity. Lower WAIC values and higher (less negative) LOO values indicate better predictive performance.

Data availability

The study was conducted using publicly available datasets as outlined in the methods sections. A comprehensive list of the used species is provided in Supplementary Table S1. We will provide the presence-absence dataset derived from Supplementary Table S7 used for species distribution modelling for transparency as data sources might have slightly changed or public availability might have become restricted in a public repository (<https://doi.org/10.5281/zenodo.20269856>).

Code availability

Analyses were conducted in R⁶³, an openly accessible statistical environment using publicly available extensions, R packages, as outlined in the methods section. Code for fitting the species distribution model, as well as analysis and figure generation will be publicly repositied (<https://doi.org/10.5281/zenodo.20269856>).

Received: 15 January 2026; Accepted: 11 May 2026;

Published online: 23 May 2026

References

1. Prentice, I. C. et al. Dynamic global vegetation modeling: quantifying terrestrial ecosystem responses to large-scale environmental change. In *Terrestrial Ecosystems in a Changing World* (eds Canadell, J. G., Pataki, D. E. & Pitelka, L. F.) 175–192 (Springer, 2007).
2. Diaz, S. & Cabido, M. Plant functional types and ecosystem function in relation to global change. *J. Veg. Sci.* **8**, 463–474 (1997).
3. Fisher, R. & Koven, C. Perspectives on the future of land surface models and the challenges of representing complex terrestrial systems. *J. Adv. Model. Earth Syst.* **12**, e2018MS001453 (2020).
4. Moncrieff, G., Bond, W. & Higgins, S. Revising the biome concept for understanding and predicting global change impacts. *J. Biogeogr.* **43**, 863–873 (2016).
5. Huntley, B. et al. Global biome patterns of the middle and late pleistocene. *J. Biogeogr.* **50**, 1352–1372 (2023).
6. de Bello, F. et al. Raunkiaeran shortfalls: challenges and perspectives in trait-based ecology. *Ecol. Monogr.* **95**, e70018 (2025).
7. Wiens, J. et al. Niche conservatism as an emerging principle in ecology and conservation biology. *Ecol. Lett.* **13**, 1310–1324 (2010).
8. Mayfield, M. & Levine, J. Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecol. Lett.* **13**, 1085–1093 (2010).
9. Ackerly, D. Adaptation, niche conservatism, and convergence: comparative studies of leaf evolution in the California Chaparral. *Am. Naturalist* **163**, 654–671 (2004).
10. Cavender-Bares, J., Ackerly, D., Baum, D. & Bazzaz, F. Phylogenetic overdispersion in Floridian Oak communities. *Am. Naturalist* **163**, 823–843 (2004).
11. Morrone, J. The spectre of biogeographical regionalization. *J. Biogeogr.* **45**, 282–288 (2018).
12. Daru, B. et al. A novel phylogenetic regionalization of phytogeographical zones of southern Africa reveals their hidden evolutionary affinities. *J. Biogeogr.* **43**, 155–166 (2016).
13. Swenson, N. The assembly of tropical tree communities - the advances and shortcomings of phylogenetic and functional trait analyses. *Ecography* **36**, 264–276 (2013).
14. Higgins, S. et al. Changes in how climate forces the vegetation of southern Africa. *Ecosystems* **26**, 1716–1733 (2023).
15. Cole, M. *The Savannas. Biogeography and Geobotany* (Academic press, 1986).
16. Johnson, R. & Tothill, J. Definition and broad geographic outline of savanna lands. In *Ecology and Management of the World's Savannas* (eds Tothill, J. & Mott, J.) 1–13 (Australian Academy of Science, 1985).
17. Scholes, R. Savanna. In *Vegetation of Southern Africa* (eds Cowling, R., Richardson, D. & Pierce, S.) 258–274 (Cambridge University Press, 1997).
18. Scholes, R. Woodlands of South Africa. In *Indigenous Forests and Woodlands in South Africa* (eds Lawles, M., Eeley, H., Shackleton, C. & Geach, B.) 3–29 (University of KwaZulu-Natal Press, 2004).
19. White, F. *The Vegetation of Africa: A Descriptive Memoir to Accompany the Unesco/aetfat/unso Vegetation Map of Africa* (United Nations, 1983).
20. Ratnam, J. et al. When is a 'forest' a savanna, and why does it matter? *Glob. Ecol. Biogeogr.* **20**, 653–660 (2011).

21. Holdo, R. & Nippert, J. Linking resource- and disturbance-based models to explain tree-grass coexistence in savannas. *N. Phytologist* **237**, 1966–1979 (2023).
22. Frost, P. et al. *Responses of Savannas to Stress and Disturbance: A Proposal for a Collaborative Programme of Research* (Biological International, Special Issue 10 (International Union of Biological Sciences (IUBS), 1986).
23. Acocks, J. *Veld types of South Africa* (Botanical Research Institute, 1988).
24. Archibald, S., Lehmann, C., Gomez-Dans, J. & Bradstock, R. Defining pyromes and global syndromes of fire regimes. *PNAS* **110**, 6442–6447 (2013).
25. Bell, R. The effect of soil nutrient availability on community structure in African ecosystems. In *Ecology of Tropical Savannas* (eds Huntley, B. & Walker, B.) 193–216 (Springer, 1982).
26. Hempson, G., Archibald, S. & Bond, W. A continent-wide assessment of the form and intensity of large mammal herbivory in Africa. *Science* **350**, 1056–1060 (2015).
27. Sankaran, M. et al. Determinants of woody cover in African savannas. *Nature* **438**, 846–849 (2005).
28. Scholes, R. & Walker, B. H. *An African Savanna: Synthesis of the Nylsvley Study*, chap. Rich savanna, poor savanna, 190–207 (Cambridge University Press, 1993).
29. Huntley, B. Southern African savannas. In *Ecology of Tropical Savannas* (eds Huntley, B. & Walker, B.) 101–120 (Springer, 1982).
30. Carter, G. *A Functional Classification of a Range of Southern African Savanna Types*. Master's thesis (University of the Witwatersrand, 1993).
31. Walker, B. & Noy-Meir, I. Aspects of the stability and resilience of savanna ecosystems. In *Ecology of Tropical Savannas* (eds Huntley, B. & Walker, B.) 556–590 (Springer, 1982).
32. Frost, P. et al. *Responses of Savannas to Stress and Disturbance* (International Union of Biological Sciences, 1986).
33. Scholes, R. Nutrient cycling in semi-arid grasslands and savannas: its influence on pattern, productivity and stability. In *Grasslands for Our World* (ed. Baker, M.) 484–487 (SIR publishing, 1993).
34. Scholes, R. The influence of soil fertility on the ecology of southern African dry savannas. *J. Biogeogr.* **17**, 415–419 (1990).
35. Cramer, M., Chimphango, S., van Cauter, A., Waldram, S. & Bond, W. Grass competition induces N₂ fixation in some species of African Acacia. *J. Ecol.* **95**, 1123–1133 (2007).
36. Hogberg, P. Root symbioses of trees in savannas. In *Mineral Nutrients in Tropical Forest and Savanna Ecosystems* (ed. Proctor, J.) 121–136 (British Ecological Society, 1989).
37. Archibald, S. & Hempson, G. Competing consumers: contrasting the patterns and impacts of fire and mammalian herbivory in Africa. *Philos. Trans. B.* **371**, 20150309 (2016).
38. Walker, B. *Determinants of Tropical Savannas*, chap. A general model of savanna structure and function, 1–12 (IRL press limited, 1987).
39. Beerling, D. & Osborne, C. The origin of the savanna biome. *Glob. Change Biol.* **12**, 2023–2031 (2006).
40. Gorel, A. et al. Climatic niche lability but growth form conservatism in the African woody flora. *Ecol. Lett.* **25**, 1164–1176 (2022).
41. Jacobs, B. Palaeobotanical studies from tropical Africa: relevance to the evolution of forest, woodland and savannah biomes. *Philos. Trans. B.* **359**, 1573–1583 (2004).
42. Conradi, T. et al. An operational definition of the biome for global change research. *N. Phytologist* **227**, 1294–1306 (2020).
43. Thornley, J. Modelling shoot: root relations: the only way forward? *Ann. Bot.* **81**, 165–171 (1998).
44. Higgins, S., Burkhardt, A., Conradi, T. & Muhoko, E. TTR.PGM: an R package for modelling the distributions and dynamics of plants using the Thornley transport resistance plant growth model. *Methods Ecol. Evol.* **16**, 1599–1608 (2025).
45. Champreux, A., Saltre, F., Traylor, W., Hickler, T. & Bradshaw, C. How to map biomes: quantitative comparison and review of biome-mapping methods. *Ecol. Monographs* **94**, e1615 (2024).
46. Osborne, C. et al. Human impacts in African savannas are mediated by plant functional traits. *N. Phytologist* **220**, 10–24 (2018).
47. Fayolle, A. et al. A sharp floristic discontinuity revealed by the biogeographic regionalization of African savannas. *J. Biogeogr.* **46**, 454–465 (2019).
48. Rees, M. et al. Patterns and drivers of phylogenetic beta diversity in the forests and savannas of Africa. *J. Biogeogr.* **52**, e15140 (2025).
49. Langdale-Brown, I. & Trapnell, C. East Africa Vegetation. In *East Africa: The people and its resources* (ed. Morgan, W. T. W.) (Oxford University Press, 1967).
50. Belsky, A. Tree/grass ratios in East African savannas: a comparison of existing models. *J. Biogeogr.* **17**, 483–489 (1990).
51. Scholes, R. et al. Trends in savanna structure and composition along an aridity gradient in the Kalahari. *J. Vegetation Sci.* **13**, 419–428 (2002).
52. Forrestel, E. et al. Different clades and traits yield similar grassland functional responses. *Proc. Natl. Acad. Sci. USA* **114**, 705–710 (2017).
53. Rattray, J. *The grass cover of Africa* (FAO, 1978).
54. Staver, A., Botha, J. & Hedin, L. Soils and fire jointly determine vegetation structure in an African savanna. *N. Phytologist* **2016**, 1151–1160 (2017).
55. Lehmann, C. et al. Savanna vegetation-fire-climate relationships differ among continents. *Science* **343**, 548–552 (2014).
56. Keith, D., Ferrer-Paris, J., Nicholson, E. & Kingsford, R. *IUCN Global Ecosystem Typology 2.0* (IUCN, 2020).
57. Walter, H. & Breckle, S. *Vegetation und Klimazonen* (Ulmer, 1999).
58. Huntley, B. *Ecology of Angola*, chap. Arid savanna biome, 59–62 (Springer, 2023).
59. Gridded livestock Density global data product (GLW 4) <https://data.apps.fao.org/catalog/iso/9d1e149b-d63f-4213-978b-317a8eb42d02> (2024).
60. Simon, M. F. et al. Recent assembly of the Cerrado, a neotropical plant diversity hotspot, by in situ evolution of adaptations to fire. *Proc. Natl. Acad. Sci. USA* **106**, 20359–20364 (2009).
61. Moncrieff, G., Hickler, T. & Higgins, S. Intercontinental divergence in the climate envelope of major plant biomes. *Glob. Ecol. Biogeogr.* **24**, 324–334 (2015).
62. Conradi, T. et al. Reassessment of the risks of climate change for terrestrial ecosystems. *Nat. Ecol. Evol.* **8**, 888–900 (2024).
63. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria <https://www.R-project.org/> (2024).
64. Hijmans, R.terra: *Spatial Data Analysis*, R package version 1.7-80, <https://github.com/rspatial/terra> (2024).
65. Higgins, S. et al. A physiological analogy of the niche for projecting the potential distribution of plants. *J. Biogeogr.* **39**, 2132–2145 (2012).
66. Karger, D. et al. Climatologies at high resolution for the earth's land surface areas Envidat. <https://doi.org/10.16904/envidat.228> (2021).
67. Dai, Y. et al. A global high-resolution data set of soil hydraulic and thermal properties for land surface modeling. *J. Adv. Model. Earth Syst.* **11**, 2996–3023 (2019).
68. von Caemmerer, S. *Biochemical Models of Leaf Photosynthesis*. Techniques in plant sciences (CSIRO publishing, 2000).
69. Botanical Information and Ecology Network <https://bien.nceas.ucsb.edu/bien/> (2024).
70. Aleman, J. et al. Floristic evidence for alternative biome states in tropical Africa. *PNAS* **117**, 28183–28190 (2020).
71. Maitner, B. BIEN: Tools for Accessing the Botanical Information and Ecology Network Database, R package version 1.2.7, <https://CRAN.R-project.org/package=BIEN> (2025).
72. Dauby, G. et al. RAINBIO: a mega-database of tropical African vascular plants distributions. *Phytokeys* **74**, 1–18 (2016).

73. Boyle, B. et al. Taxonomic Name Resolution Service, version 5.1 <https://tnrs.biendata.org/> (2021).
74. GBIF: The global Biodiversity Information Facility. What is GBIF? Available from <https://www.gbif.org/what-is-gbif> (2024).
75. Chamberlain, S. et al. rgbif: Interface to the Global Biodiversity Information Facility API. R package version 3.8.4, <https://CRAN.R-project.org/package=rgbif> (2026).
76. Zizka, A. et al. CoordinateCleaner: standardized cleaning of occurrence records from biological collection databases, R package version 3.0.1, <https://github.com/ropensci/CoordinateCleaner> (2019).
77. Aiello-Lammens, M., Boria, R., Radosavljevic, A. & Vilela, B. spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography* **38**, 541–545 (2015).
78. Rutherford, M., Powrie, L. & Midgley, G. ACKDAT: a digital spatial database of distributions of South African plant species and species assemblages. *South Afr. J. Bot.* **69**, 99–104 (2003).
79. Rutherford, M., Mucina, L. & Powrie, L. The South African National Vegetation Database: history, development, applications, problems and future. *South Afr. J. Sci.* **108**, 8 (2012).
80. Schmidt, M. et al. The West African Vegetation database. *Biodivers. Ecol.* **4**, 105–110 (2012).
81. Taugourdeau, S. et al. Flotrop, a massive contribution to plant diversity data for open ecosystems in northern tropical Africa. *Nat. Sci. Data* **6**, 118 (2019).
82. Osborne, C. et al. A global database of C₄ photosynthesis in grasses. *N. Phytologist* **204**, 441–446 (2014).
83. Ardia, D., Boudt, K., Carl, P., Mullen, K. & Peterson, B. Differential evolution with DEoptim: an application to non-convex portfolio optimization. *R. J.* **3**, 27–34 (2011).
84. Statisticat & LLC. LaplacesDemon: Complete Environment for Bayesian Inference, R package version 16.1.6, <https://web.archive.org/web/20150206004624/http://www.bayesian-inference.com/software> (2021).
85. Scrucca, L., Fraley, C., Murphy, T. & Raftery, A. *Model-Based Clustering, Classification, and Density Estimation Using mclust in R* (Chapman and Hall/CRC, 2023). <https://mclust-org.github.io/book/>.
86. Monserud, R. & Leemans, R. Comparing global vegetation maps. *Ecol. Model.* **62**, 275–293 (1992).
87. Therneau, T. & Atkinson, B. *rpart: Recursive Partitioning and Regression Trees*, R package version 4.1.23. <https://CRAN.R-project.org/package=rpart> (2023).
88. Milborrow, S. *rpart.plot: Plot 'rpart' Models: An Enhanced Version of 'plot.rpart'* R package version 3.1.2, <https://CRAN.R-project.org/package=rpart.plot> (2024).
89. Running, S. & Zhao, M. MODIS/Terra Net Primary Production Gap-Filled Yearly L4 Global 500m SIN Grid V061. NASA EOSDIS Land Processes Distributed Active Archive Center. Accessed 2024-10-23 from <https://doi.org/10.5067/MODIS/MOD17A3HGF.061> (2018).
90. UNEP-WCMC & IUCN. Protected Planet: The World Database on Protected Areas (WDPA), www.protectedplanet.net (2024).
91. Randerson, J., van der Werf, G., Giglio, L., Collatz, G. & Kasibhatla, P. Global fire emissions database (GFED 4.1), <https://www.globalfiredata.org/data.html> (2018).
92. Plummer, M. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*, 1–10 (Technische Universität Wien, Vienna, Austria, 2003). <https://www.r-project.org/conferences/DSC-2003/Proceedings/Plummer.pdf>.
93. Vehtari, A. et al. loo: Efficient leave-one-out cross-validation and waic for Bayesian models, R package version 2.8.0, <https://mc-stan.org/loo/> (2024).
94. Gelman, A., Goodrich, B., Gabry, J. & Vehtari, A. R-squared for Bayesian regression models. *Am. Statistician* **73**, 307–309 (2019).

Acknowledgements

We dedicate this manuscript to the memory of William Bond: mentor, friend and provocateur. We thank Norman Owen-Smith for stimulating us to revisit the question, what are the determinants of the different African savannas types? Open Access funding supported by the University of Bayreuth Open Access publishing fund.

Author contributions

S.I.H. conceived the study, V.P. and S.I.H. designed the study and wrote the paper, V.P. performed the research and undertook the analyses.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s44185-026-00138-5>.

Correspondence and requests for materials should be addressed to Vitus Pickelmann.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2026