

Research article

Using joint species distribution modelling to identify climatic and non-climatic drivers of Afrotropical ungulate distributions

Alex Cranston^{1,2}, Natalie Cooper² and Jakob Bro-Jørgensen¹

¹Department of Evolution, Ecology and Behaviour, Mammalian Behaviour and Evolution Group, University of Liverpool, Neston, UK

²Science Group, Natural History Museum London, Cromwell Road, London, UK

Correspondence: Alex Cranston (a.cranston@liverpool.ac.uk)

Ecography

2024: e07209

doi: [10.1111/ecog.07209](https://doi.org/10.1111/ecog.07209)

Subject Editor: Douglas A. Kelt

Editor-in-Chief:

Christine N. Meynard

Accepted 28 June 2024



The relative importance of the different processes that determine the distribution of species and the assembly of communities is a key question in ecology. The distribution of any individual species is affected by a wide range of environmental variables as well as through interactions with other species; the resulting distributions determine the pool of species available to form local communities at fine spatial scales. A challenge in community ecology is that these interactions (e.g. competition, facilitation, etc.) often are not directly measurable. Here, we used hierarchical modelling of species communities (HMSC), a recently developed framework for joint species distribution modelling, to estimate the role of biotic effects alongside environmental factors using latent variables. We investigate the role of these factors determining species distributions in communities of Artiodactyla, Perissodactyla and Proboscidea in the Afrotropics, an area of peak species richness for hoofed mammals. We also calculate pairwise trait dissimilarity between these species, from a mixture of morphological and behavioural traits, and investigate the relationship between dissimilarity and estimated residual co-occurrence in the model. We find that while ungulate distributions appear to be predominantly determined (~ 70%) by climatic variables, such as precipitation, a substantial proportion of the variance in ungulate species distributions (~ 30%) can also be attributed to modelled latent variables that likely represent a combination of dispersal barriers and biotic factors. Although we find only a weak relationship between residual co-occurrence and trait dissimilarity, we suggest that our results may show evidence that biotic factors, likely influenced by historical barriers to species dispersal, are important in determining species communities over a continental area. The HMSC framework can be used to provide insight into factors affecting community assembly at broad scales, and to make more powerful predictions about future species distributions as we enter an era of increasing impacts from anthropogenic change.

Keywords: biogeography, community ecology, hierarchical modelling of species communities (HMSC), joint species distribution models (JSDM), ungulates



www.ecography.org

© 2024 The Author(s). Ecography published by John Wiley & Sons Ltd on behalf of Nordic Society Oikos

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Introduction

A central challenge in community ecology is understanding the factors that determine the assembly of species communities. A community, generally defined as the total assemblage of the species that co-occur at the same place at the same time (Krebs 2009), is determined by the individual distributions of the species that comprise it. The factors that influence community assembly are varied, including phylogeographic effects (such as speciation, extinction and dispersal), environmental conditions and biotic interactions among species (Lawton 1999, Hubbell 2001, Vellend 2010, Stegen et al. 2013). In community ecology, these effects are often viewed as a series of filters (Zobel 1997), with historical phylogeographic effects such as speciation and dispersal acting as the first filter to create a regional pool of species that are theoretically able to be part of a community. Local environmental conditions act as a second filter to further limit the regional pool to the subset of species that are adapted to an area's climatic or geophysical conditions. Finally, interactions among species, such as competition among ecologically similar species or plant-pollinator mutualisms, can further affect community assembly by excluding species that would otherwise be able to persist given an area's environmental conditions, or alternatively, by promoting their presence through facilitation. While these factors all influence community composition (Lortie et al. 2004, Sexton et al. 2009, Meier et al. 2010), their relative importance is often unclear.

The importance of biotic effects in determining species distributions has been a particular source of debate (Wiens 2011, Pigot and Tobias 2013, Louthan et al. 2015). Not only is it unclear to what extent biotic effects might influence distributions, it is also unclear at what scale such effects might be predicted to act. Some suggest that biotic effects are typically likely to be significant only at highly local scales (i.e. $< 10 \text{ km}^2$) rather than on ecoregion or biome-wide scales (i.e. $> 100\,000 \text{ km}^2$; Pearson and Dawson 2003, Soberón and Nakamura 2009). Others, however, stress the importance of biotic effects at broader spatial extents (i.e. $> 10 \text{ km}^2$; Wisz et al. 2013). Empirical evidence remains inconclusive, with some studies suggesting weaker but still present biotic effects at broader spatial scales (Bullock et al. 2000, Belmaker et al. 2015, Staniczenko et al. 2017, Mod et al. 2020, Cosentino et al. 2023), while others have suggested a key role for biotic effects in determining range limits, especially at lower latitudes (Sirén et al. 2021, Freeman et al. 2022), with competition in particular highlighted as the most important biotic effect in determining community composition (Staniczenko et al. 2018). Understanding how and at what scales biotic factors affect species distributions and hence the assembly of species communities remains a key goal in community ecology.

Joint species distribution models (JSDMs) offer a promising approach to investigating questions about the relative importance of environmental and biotic factors in determining species distributions, by allowing the responses of multiple species to their environment and to each other to be

modelled simultaneously (Ovaskainen et al. 2017). These models use commonly collected and readily available ecological data, such as presence–absence data or count data. Although JSDMs are a powerful method for understanding species distributions, biotic effects can still only be inferred with the use of latent variables, rather than being directly measured (Warton et al. 2015). If competition plays a meaningful role in determining ungulate ranges, we would expect ecologically similar species to be less likely to occur together after controlling for the effect of the environment. Environmental filtering will overall lead to similar environments having species with more similar traits (Sutton et al. 2021). However, by using a JSDM with a latent variable approach, we expect to control for the effect of environmental filtering. As species with more similar functional traits are more likely to show niche overlap and therefore competitively exclude each other, this should limit the overlap of distributions relative to that expected from environment alone (Hardin 1960, Abrams 1983, Booth and Murray 2008). Therefore, as part of investigating the relative influence of biotic effects and abiotic effects on community composition, here, we assess the extent to which latent variables can be attributed to biotic effects by testing how residual associations they capture correlate with trait dissimilarity.

We expect the residual associations between species (defined by the latent variables) to fall generally along biogeographic boundaries as outlined by Lorenzen et al. (2012). These boundaries will roughly reflect differences in communities due both to dispersal limitations imposed by geographical barriers and range limits due to biotic effects, therefore we predict that the model will reflect these after controlling for climate variables. Further, we predict that species with more similar traits are less likely to co-occur than would be expected from environmental conditions alone.

The Afrotropics are the most diverse region on Earth for ungulates (Du Toit and Cumming 1999) with close to 30 species occurring in sympatry in areas of peak species richness (IUCN 2022). The importance of biotic factors among ungulate species in determining their distributions is unclear. There is often a high degree of dietary overlap between ungulate species, generally feeding on grass, browse and fruit. This overlap suggests potentially high levels of competition among ecologically similar species. However, African ungulates show significant variation in morphology, both in body size, digestive system and dietary specialisation (i.e. browsing versus grazing), which may reduce interspecific competition (Veldhuis et al. 2019). Furthermore, in comparison to other mammalian clades, wild ungulates rarely exhibit direct interspecific interference (Ferretti and Mori 2020); although we note that this is not the only mechanism by which competition can occur. One reason for the relatively limited evidence of competitive interference in ungulates may be predation pressure favouring the evolution of mutualistic antipredator strategies, such as herding, rather than competitive behaviour (Sinclair 1985, Caro 2005). Other forms of facilitation between ungulate species may also promote co-existence and reduce competition. It is therefore unclear the extent

to which biotic effects will affect range limits in ungulates. Understanding all relevant factors is crucial for predicting the impact of future changes in ungulate distributions resulting from climate change and anthropogenic pressure. This is especially pertinent, given the multiple ungulate species in Africa that are of conservation concern. For example, 61% of the 90 species of antelope currently recognised by IUCN SSC Antelope Specialist Group have decreasing populations (IUCN SSC Antelope Specialist Group 2022). Furthermore, previous work has highlighted that already threatened ungulate species are likely to be disproportionately affected by climate change related range reductions (Payne and Bro-Jørgensen 2016). Here, we use species range data collected from the IUCN Red List together with publicly available climate data to investigate the factors determining ungulate distributions and hence, ungulate community composition over a continental area using JSDBMs.

Material and methods

Method framework

This study uses JSDBMs to assess the factors that shape species distributions at a continental scale. JSDBMs require data on species distributions to act as the dependent variable, either in the form of presence/absence or abundance data. At a broad spatial scale, the former is widely available for many vertebrate species through the species geographic range polygons provided in the International Union for Conservation of Nature (IUCN) Red List (IUCN 2023). However, using IUCN range data places some limitations on this work. The IUCN range for a species indicates that the species likely occurs within these limits, but does not imply it is distributed evenly across that range or that it occurs at all points within it, and this is particularly problematic at finer spatial scales (IUCN SSC Red List Technical Working Group 2021). Previous studies have also suggested that the IUCN data may sometimes overestimate the extent of species ranges (Ramesh et al. 2017). Using IUCN range data also means that only presence/absence in species distributions can be considered and not abundance, which may vary substantially across a species' range, particularly at range limits (Yancovitch Shalom et al. 2020). Both of these factors mean that negative associations among species may be underestimated due to lack of information on how competition among species impacts abundance and the finer-scale spatial segregation between species that may occur even if their ranges overlap at coarser scales. Because of this latter point, it is important to remember that positive co-occurrences discussed in this paper do not imply that species are necessarily co-occurring at finer resolutions than the one presented here. Despite these limitations, we consider the IUCN range data to be the best source currently available for ungulate ranges in the Afrotropics as it is available for all species and all areas of the region, in contrast to other sources that are biased towards the Global North (e.g. global biodiversity information facility; GBIF).

Data collection

We included all ungulate species (orders Artiodactyla and Perissodactyla) that had at least some of their native range in continental Africa ($n=98$). In addition to ungulates, we included both species of African elephant (*Loxodonta cyclotis* and *Loxodonta africana*) because, although not closely related to other species in the model, as mega-herbivores they are plausible candidates for competition with predominantly herbivorous ungulates (Ferry et al. 2016). Shapefiles for all species ranges (aside from the two African rhino species) were taken from the IUCN Red List (IUCN 2021), selecting only shapefiles for parts of the range where the species was both extant and native. For the rhino species, we used historical range maps produced by Rookmaaker and Antoine (2012), as the IUCN range maps for these species only show countries where these species are present rather than exact distributions due to risks from poachers. Regardless of source, these shapefiles were converted into a presence-absence matrix with a resolution of 10 arcminutes using the R package 'letsR' (Vilela and Villalobos 2015). To focus exclusively on the Afrotropics, we selected only points south of 15°N.

To control for the effect of human activities on ungulate distributions, we selected only points that fell within highly protected areas, reasoning that these areas were most likely to have 'natural' ungulate communities, with ungulate presence/absence relatively unaffected by human activity. We downloaded shapefiles for African protected areas from the World Database on Protected Areas (WDPA; UNEP-WCMC and IUCN 2022). These were filtered to select only IUCN category Ia (Strict Nature Reserve), Ib (Wilderness Area) and II (National Park) protected areas. Data on the IUCN category was not available for all protected areas in WDPA. Therefore, we also selected protected areas with the self-described designation 'national park'. To further control for the effect of human influence, we used the FAO's Global Land Cover-SHARE database (Latham et al. 2014) to remove points where the dominant land cover was cropland or urban surface, and removed points where land cover was not recorded or where the dominant cover was open water, i.e. lakes.

All bioclimatic variables used as fixed effects in the model were taken from WorldClim (Fick and Hijmans 2017). Previous work (Payne 2015) used a variable importance analysis on 30+ variables and found that the most important variables for use in predicting distributions of African antelope were hottest annual temperature, coldest annual temperature, mean annual temperature, range of temperature and log of mean annual precipitation. Because hottest annual temperature and coldest annual temperature both strongly correlate with mean annual temperature, for simplicity we selected this single variable. Likewise, within-year range in temperature correlates almost perfectly with mean diurnal range, hence we chose this latter variable. Additionally, we thought it important to capture variability with respect to precipitation, so we initially included both precipitation seasonality (coefficient of variation), i.e. variation in precipitation from year to year, and precipitation range within a year (precipitation

of wettest quarter – precipitation of driest quarter). However, in our checks for multicollinearity the within-year precipitation range had a high variance inflation factor, so this variable was removed from the analysis. The final variables used in the model therefore were mean annual temperature (°C), annual precipitation (mm), mean diurnal range (°C) and precipitation seasonality (coefficient of variation; mm). We calculated the variance inflation factor for each of these variables; all were between 1 and 1.4, indicating no multicollinearity in the model.

At a simplified level, the Afrotropics can be divided into two broad vegetation zones, savannahs and tropical rainforests (Lorenzen et al. 2012), with the difference between the two largely determined by level of precipitation and wildfire (Staver et al. 2011). We reasoned that biotic interactions would be especially prominent in determining the distribution of species within a biome, rather than between biomes (as here the climatic differences between the highly different habitats would be expected to be pre-eminent in determining the distributions). Therefore, for the JSDMs, we fitted models for different Afrotropical biomes separately, fitting two models. One used only points from within the Tropical and Subtropical Moist Broadleaf Forests biome (hereafter, the tropical forests model), including 30 species, and one that contained all points not in this biome, covering a range of open and semi-open habitats (hereafter, the open habitats model), which included 57 species. We obtained data on biomes using shapefiles taken from the RESOLVE Ecoregions map (Dinerstein et al. 2017).

We collected data on eight different traits for each species: adult body mass (g), adult body length (m), percentage of diet consisting of fruit, percentage of diet consisting of dicots, percentage of diet consisting of monocots, whether diet showed geographic or seasonal variability (yes/no), the digestive system for each species (monogastric, ruminant or pseudoruminant) and lastly, whether the species was migratory (yes/no). Trait data for body mass and body length were taken from COMBINE, a database of intrinsic and extrinsic mammalian traits (Soria et al. 2021). This database did not recognise the African forest elephant *Loxodonta cyclotis* as a distinct species from the African bush elephant (*Loxodonta africana*), so this species was excluded from dissimilarity analyses (below). Data on diet for most species was taken from Gagnon and Chew (2000). From this paper, we recorded the proportion of fruit, dicotyledon plants, and monocotyledon plants in the diets of all extant African Bovidae (aside from Caprinae), which made up the majority of species in our models. From this paper, we also recorded whether diet showed seasonal and/or geographic variability (a binary yes/no). For species not covered by Gagnon and Chew (2000), i.e. all non-bovid species covered by the models, we estimated these variables with reference to Kingdon et al. (2013). For all species, we recorded the digestive system (monogastric, ruminant or pseudoruminant) using Kingdon et al. (2013). Finally, we recorded whether each species was migratory using Abraham et al. (2022), although we were only able to record data for 75 species in our dataset. Overall, for the

open habitats model we were able to record complete trait data for 48 out of 57 species included in the model and for the tropical forests model, we had complete trait data for 24 out of 30.

Analyses

All analyses used R ver. 4.1.1 (www.r-project.org). All data used is available on Dryad (Cranston et al. 2024). All code used is available at <https://zenodo.org/doi/10.5281/zenodo.12739431>. The core analyses in HMSC (below) were conducted using a modified version of the HMSC pipeline made publicly available by the package authors and available at www.helsinki.fi/en/researchgroups/statistical-ecology/software/hmsc.

Joint species distribution models

We used joint species distribution models (JSDMs) to evaluate the relative impact of biotic and abiotic factors on ungulate species distributions, specifically using the hierarchical modelling of species communities (HMSC; Ovaskainen et al. 2017) approach. We chose HMSC due to its relative ease of use and excellent predictive power in comparison to other comparable methods (Norberg et al. 2019). An overview of the basic workflow and key outputs from HMSC is presented in Fig. 1, but to summarise HMSC uses generalised linear mixed models (GLMMs), fitted using Bayesian inference, where latent variables are estimated by HMSC to explain the residual variation in species' occurrences, i.e. the variation not explained by the fixed effects of the model (Tikhonov et al. 2020). Typically, these fixed effects are directly measurable environmental variables, while the latent variables may estimate variables that are difficult or impossible to measure directly, such as biotic factors like competition or facilitation among species. However, it is important to note that these latent factors may also capture unexplained environmental variation and it is not possible to interpret them as representing a specific variable with 100% confidence. Mathematically, these latent factors consist of two parameters, η (eta), and λ (lambda). Eta is the spatially explicit part of the latent variable with a different value at each site in the model; it is mathematically equivalent to the values of the various fixed effects provided to the model, in that each sample site has its own value of eta and this value varies over space. It is best interpreted as reflecting differences in species communities after controlling for the fixed effects. Lambda is the response of the species to the random effect; it is mathematically equivalent to the response of the species to the various fixed effects provided to the model. Species with more similar values of lambda will have higher residual association, i.e. they will associate more than expected given the fixed effects alone. HMSC derives an additional parameter, Ω (omega), by comparing pairwise values of lambda between species to give a value for residual species association. Alongside these parameters, HMSC also estimates a final parameter, β (beta), which describes the response of each species to each fixed effect.

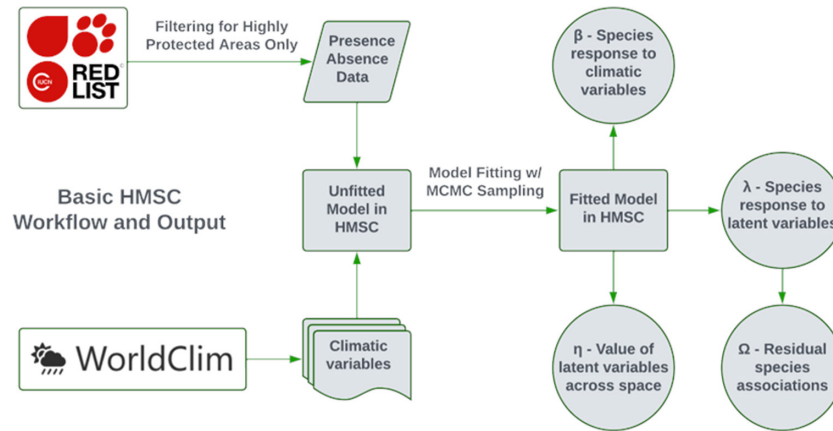


Figure 1. A general overview of our workflow in HMSC with all relevant inputs and outputs. Presence–absence data derived from the IUCN species range maps act as the dependent variables. Climatic variables sourced from WorldClim (annual mean temperature, annual precipitation, mean diurnal range, and precipitation seasonality) act as our independent variables (fixed effects). HMSC then fits the model, estimating values for all parameters using MCMC sampling, and crucially, estimating the latent variables that make up the random effect part of the model. The key outputs of HMSC for our purposes in this study are: β , or beta, the set of parameters estimating the response of each species to each fixed effect; η , or eta, the spatially explicit part of the latent variable or random effect, with a different value at each site in the model, λ , or lambda, the set of parameters estimating the response of each species to each random effect; Ω , or omega, the residual species associations, i.e. the predicted association between species after controlling for the fixed effects (this is derived directly from lambda).

Using HMSC, we fitted a probit model to estimate the probability of presence for each species at each point in our dataset. Each model used a random sample of 300 points for each model from our overall dataset to make running the models computationally feasible. HMSC has lower accuracy in explaining the distribution of very rare species and including such species can increase computational times and complicates MCMC convergence (Ovaskainen and Abrego 2020). Therefore, for each model, we included only species that occurred in at least 2.5% of sites in the sample of data used. After removing species with low prevalence, we noted that for the tropical forests model, 24 species typically regarded as savannah or open habitat specialists, were still included in the model, e.g. common eland *Tragelaphus oryx* and plains zebra (*Equus quagga*; see the Supporting information for full species list). To a lesser extent, the same problem occurred for the open habitats model with inclusion of closed forest specialists, albeit with a much smaller number of species ($n=4$). Inspection of maps revealed that this was probably due to the spatial resolution of both the IUCN and RESOLVE Ecoregions map which – combined with a degree of error at range/biome boundaries due to their broadscale nature – led to an artefactual degree of overlap of these savannah species’ ranges with the tropical and subtropical moist broadleaf forests biome in the boundary area between open and closed habitats, and vice versa for the tropical forest species. Therefore, we manually removed all species regarded as inhabitants of savannahs rather than forests, based on Kingdon et al. (2013), from the tropical forest model ($n=24$), and removed four closed canopy specialists from the open habitats model (retaining those considered to inhabit the forest–savannah ecotone), again consulting Kingdon et al. (2013). We opted not to increase the prevalence threshold, as setting this threshold any higher in an effort to

remove all unwanted species would lead to also removing species of interest, such as Jentink’s duiker *Cephalophus jentinki* and zebra duiker *Cephalophus zebra*, which have very limited ranges. For a full list of species removed from both models, please see the Supporting information.

In addition to our fixed effects (mean annual temperature, annual precipitation, mean diurnal range and precipitation seasonality), we used HMSC’s ability to estimate spatially explicit latent variables in the random effect part of the model to model the residual associations between ungulate species. HMSC uses a Markov chain Monte Carlo (MCMC) approach to sample the posterior distribution for each parameter estimated by the model. For our models, we used four chains, and each was run for 37 500 iterations. We discarded the first 12 500 iterations as burn-in and then thinned by 100 to yield 250 posterior samples for each chain. We checked MCMC convergence for each model using a potential scale reduction factor. In each model for each parameter inspected, the maximum value of this factor was close to 1 and MCMC convergence was satisfactory, with the sole exception of the omega parameter in the tropical forests model, where the maximum potential scale reduction factor was 3.054, indicating these parameter estimates representing residual species associations may be less reliable than other parameters that we consider here. However, as our models were already taking around 100 h to run, a longer run time to achieve full convergence was unfeasible in the time available.

We evaluated model fit for each species in the model using root mean squared error (RMSE), and Tjur’s R^2 . After running the models, we estimated the relative importance of environmental versus biotic factors in determining species distributions by partitioning the variance explained by each of the fixed effects and the latent variables. We also examined the residual associations among ungulate species captured

by the latent variables, plotting these as a correlation matrix using the `gplots` package (Warnes et al. 2005), with positive values indicating a higher degree of co-occurrence than would be expected from environmental conditions alone (and vice versa), with only values with over 95% support considered statistically significant.

Trait dissimilarity analysis

We calculated pairwise dissimilarity in our chosen traits between species using Gower's distance (Gower 1971), implemented in R using the package `gower` (van der Loo 2022). Gower's distance gives a value of 1 for species that had maximally dissimilar traits and 0 for species where the traits were identical. We then ran a linear regression of omega values, estimated by HMSC, against these Gower distance values. As omega values are pairwise estimates of the residual association between species (i.e. the extent to which species are expected to co-occur after controlling for the fixed effects), with higher values of omega indicating higher association, a positive association would support our hypothesis that more dissimilar species were more likely to associate with each other. For this analysis we excluded omega values with less than 95% support.

Results

Joint species distribution models

The explanatory and predictive power of both models was relatively high (explanatory power: open habitats: RMSE=0.19, mean Tjur's $R^2 = 0.65$; tropical forests: RMSE=0.12, mean Tjur's $R^2 = 0.80$; predictive power: open habitats model: mean Tjur's $R^2=0.54$; tropical forests model: mean Tjur's $R^2=0.66$) although this varied significantly across species; for example, in the open habitats model, Tjur's R^2 for the explanatory power ranged from 0.22 for the lechwe *Kobus lechwe* to 0.95 for the East African oryx *Oryx beisa*. There was no relationship between extent of occupancy and Tjur's R^2 for the explanatory power in either model (open habitats: $R^2 = 0.0001$, $F_{1,55} = 0.008$, $p = 0.93$; tropical forests: $R^2 = 0.018$, $F_{1,28} = 0.50$, $p = 0.48$).

In both models, the majority of the explained variance was attributable to the climatic fixed effects. In the open habitats model, 71% was explained by the climate factors versus 29% by spatial random effects, while in the tropical forests model, the proportion was 70–30%. Again, in both models there was a great deal of variation across species, with some species in some models having the majority of their explained variance attributed to the spatial random effects. For example, the impala *Aepyceros melampus* in the open habitats model and Peters's duiker *Cephalophus callipygus* in the tropical forests model have 86 and 72% respectively of their explained variance attributable to the spatial random effects (Supporting information)

The residual associations among species were plotted in a correlation plot for both models (Fig. 2–3). These plots were arranged using hierarchical clustering, with more similar rows and columns (each one representing a species in the model)

grouped together. This helped us identify patterns in the data as species with more similar associations group together in the plot. The dendrogram at the top and left of the plot shows this hierarchical clustering, although we note that our interpretation of the number of clusters is based on our biological interpretation of the plots rather than any formal analysis of the optimal number of clusters, i.e. k-means clustering.

The open habitats model (Fig. 2) had four clusters (plus six species that did not appear to belong to any cluster). Working through the clusters from the bottom left to top right of the plot, the first cluster (A) consisted largely of species with most of their distribution occurring in West Africa and across the Sudanian savannah belt, such as the kob *Kobus kob*, giant eland *Tragelaphus derbianus* and the red-flanked duiker *Cephalophus rufilatus*. The second (B) consisted largely of species with most of their distribution occurring in arid regions of the Horn of Africa (Ethiopia, Somalia, north-east Kenya), such as Günther's dik-dik *Madoqua guentheri*, lesser kudu *Tragelaphus imberbis* and East African oryx *Oryx beisa*. The third (C) consisted largely of species with their distribution in southern African savannahs, such as the puku *Kobus vardonii* and southern reedbeek *Redunca arundinum*, but also including some with a more widespread distribution such as the roan antelope *Hippotragus equinus*. Finally, the fourth (D) was made up primarily of species with their distribution in eastern, and most also extending into southern, African savannahs, including the Kirk's dikdik *Madoqua kirki*, Thomson's gazelle *Eudorcas thomsoni* and Grant's gazelle *Nanger granti*. We note that despite these general patterns, some of the species in clusters A, B, C and D showed positive associations with each other – we attribute this to the fact that, although these clusters were made of species with most or all of their ranges in the discrete parts of the biome, they also contained species with more widespread distributions such as the common duiker *Sylvicapra grimmia* and the giraffe *Giraffa camelopardalis* that may drive these associations.

In the tropical forests model (Fig. 3) there were four clusters (plus nine species that did not appear to cluster with any other). The first cluster (A) consisted largely of species with a significant part of their distribution in the Guinean forests of West Africa, such as Jentink's duiker *Cephalophus jentinki*, royal antelope *Neotragus pygmaeus* and Maxwell's duiker *Philantomba maxwellii*. The second (B) consisted largely of species with the bulk of their distribution in the East African montane or coastal tropical forests, such as Harvey's duiker *Cephalophus harveyi*, suni *Nesotragus moschatus* and red forest duiker *Cephalophus natalensis*. The third and fourth clusters (C and D) consisted largely of species with a significant part of their distribution occurring in the Congolian rainforest region, such as Peters's duiker *Cephalophus callipygus*, white-bellied duiker *Cephalophus leucogaster* and Bates' pygmy antelope *Nesotragus batesi*. We also note that many of the species in these last two clusters showed positive associations with each other. It is worth restating at this point that these residual associations correlation plots are based on the omega parameter estimated by HMSC and, in the case of this model, the maximum potential scale reduction factor for this parameter was well above the ideal value of 1, indicating the

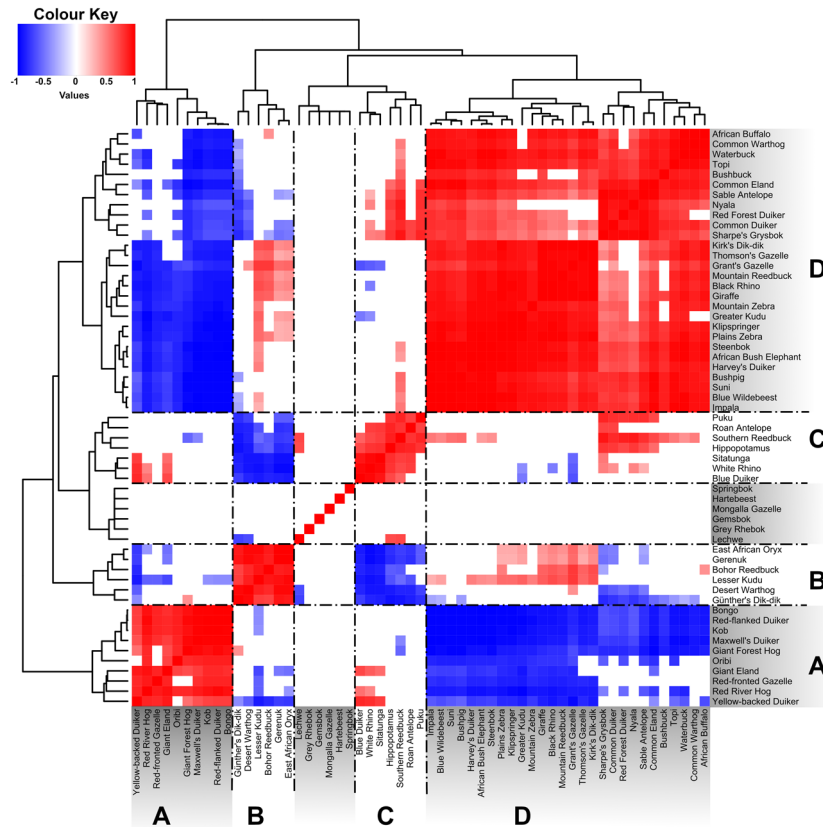


Figure 2. Correlation plots showing residual associations between species in the open habitats model. Red indicates that the residual association between two species is significant and positive, i.e. species are more likely to co-occur than would be expected from the environmental factors included as fixed effects in the model. Blue indicates the reverse, i.e., species are less likely to co-occur than would be expected from the environmental factors included as fixed effects in the model. Only significant values are shown, white indicates non-significant values. The dendrograms on the top and left axes show clustering of species with similar values, i.e. species that are broadly associated with each other. Clusters have been labelled from left to right/bottom to top and separated with dashed lines for visual clarity. Note that six species in the centre of the graph do not cluster with any other and so are not labelled as a cluster.

MCMC chains had not fully converged so these estimates should be interpreted cautiously.

We plotted the eta values of the latent variables at each point in the model (Fig. 4–5), with Kriging interpolation to more easily visualise gradients over the entire biome. Each model estimated a total of five latent variables to explain the residual variation in species occupancy; however, in the open habitats model, only the first two latent variables made a substantial contribution to the model, while in the tropical forests model, only the first one did. Therefore, only these variables are discussed here.

In the open habitats model, the first latent variable (Fig. 4A) shows that the main axis of difference was between the west/central African and the southern/eastern African savannahs. The second latent variable (Fig. 4B) shows the secondary axis of difference as between the Horn of Africa and the rest of the biome. Additionally, in both variables, the Namib Desert in southwest Africa appears to stand out as distinct from the rest of southern Africa. Taken together, this suggests that these four regions (West African savannahs, Horn of Africa, eastern/southern African savannahs and the Namib) all have differences in ungulate community composition not explained by climate effects alone.

In the tropical forests model, the first latent variable (Fig. 5) shows that the main axis of difference was between the Congolian rainforests areas and the western Guinean rainforest, while the East African coastal and montane rainforests had no clear pattern. Overall, this suggests that these two regions (Congolian rainforests and western Guinean forests) have differences in ungulate community composition not explained by climate effects alone, with differences also existing between parts of the East African coastal and montane rainforests and Congolian rainforests.

Trait dissimilarity analysis

There was a very weak but significant negative relationship between pairwise trait dissimilarity (measured using Gower's distance) and residual association (measured using the omega parameter estimated in HMSC) in the open habitats model ($R^2 = 0.018$, $F_{1,843} = 15.78$, $p < 0.001$), suggesting that, contrary to our hypothesis, more dissimilar species were less likely to co-occur with each other, albeit only marginally so. In the tropical forests model, there was no significant relationship ($R^2 < 0.001$, $F_{1,115} = 0.007$, $p = 0.93$).

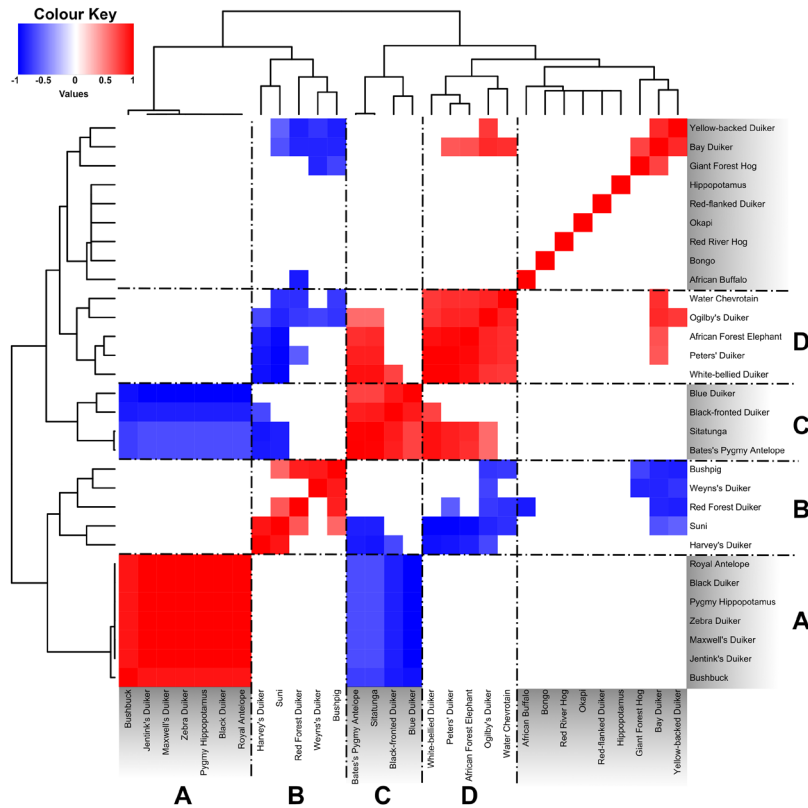


Figure 3. Correlation plots showing residual associations between species in the tropical forests model. Red indicates that the residual association between two species is significant and positive, i.e. species are more likely to co-occur than would be expected from the environmental factors included as fixed effects in the model. Blue indicates the reverse, i.e. species are less likely to co-occur than would be expected from the environmental factors included as fixed effects in the model. Only significant values are shown, white indicates non-significant values. The dendrograms on the top and left axes show clustering of species with similar values, i.e., species that are broadly associated with each other. Clusters have been labelled from left to right/bottom to top and separated with dashed lines for visual clarity.

Discussion

Our results indicate that while climate is the primary determinant of ungulate distribution in the Afrotropics (after controlling for the influence of human activity), it is not the only important factor with the latent variables revealing significant differences in ungulate community composition even after controlling for climatic conditions, with around 30% of all explained variance attributable to these latent effects. Contrary to our expectations, however, trait dissimilarity did not correlate positively with the residual variation. In fact, in the open habitat model a very weak negative association was found, with more dissimilar species being marginally less likely to be significantly associated after controlling for the environment; however, this association was so weak that it is unlikely to be biologically meaningful and may be statistically significant simply due to our high sample size. The tropical forests model showed no significant relationship.

This study uses HMSC to estimate the proportion by which species distributions are determined by climate (modelled by the fixed effects) versus other factors such as biotic interactions (modelled by the latent variables). We find that, consistent with other recent studies (Cosentino et al. 2023), there is evidence for a substantial role for non-climatic factors

in determining species distributions at broad as well as fine-scale spatial extents. However, we note two important caveats in our estimate. Firstly, the exact values of these estimated proportions will be influenced by the extent over which one considers the species ranges. Here, we have considered the tropical forest and open habitat biomes in two separate models, and this has likely influenced the relative influence of these factors in our results, as the influence of climate on species distributions is likely to be different than if we had combined these areas in a single model. Secondly, as mentioned previously, the interpretation of these latent variables as representing biotic factors alone is not straightforward, as they may represent unexplained environmental variation. In the remainder of this paper, we will discuss the results of the latent variables in more detail and propose how they might best be interpreted.

HMSC successfully detects biogeographic patterns

Our results show geographical differences between ungulate communities across the Afrotropics beyond that predicted from climatic differences. In the tropical forests model, more significant differences exist between ungulate communities in the Congolian rainforests and the western Guinean and parts of the eastern African montane and coastal rainforests than would be expected from the climatic fixed effects alone. This

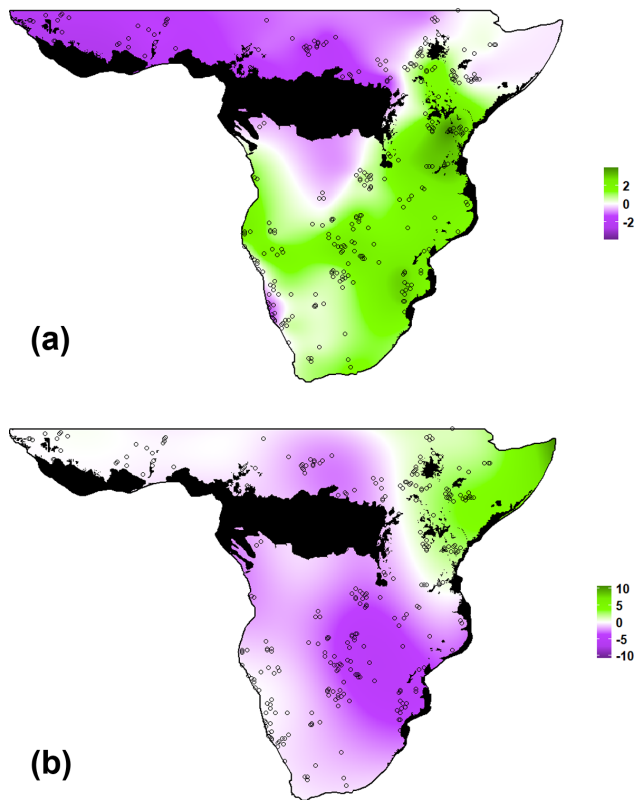


Figure 4. The site loadings values, or eta parameter, of the first (A), second (B) latent variable from the open habitats model for all points in the model interpolated across the Afrotropics (excluding tropical forests) using Kriging interpolation. These site loadings show differences in ungulate communities after controlling for climate differences included in the fixed effects of the model. Areas with positive values have more different communities compared to areas with negative values than expected for climate alone. Sample sites in the model are shown by the hollow circles. Areas in black indicate areas belonging to the moist tropical forest biome which were not included in this model and hence not interpolated over. Site loadings values plotted without interpolation, are in the supplementary materials (Supporting information).

result was not unexpected as the African tropical rainforest biome is non-contiguous, with the areas of tropical grassland potentially acting as barriers between areas of rainforest, such as the Dahomey Gap between the western Guinean region and the Congolian region (Salzmann and Hoelzmann 2005). It is likely that the latent variables in our model are capturing these barriers to dispersal that result in differences in community composition even after controlling for climate. These barriers have not been stable over evolutionary history; pollen records indicate that the Dahomey Gap, for example, has fluctuated in size over the last 150 000 years, including disappearing completely during periods of maximum forest expansion (Dupont and Weinel 1996). Furthermore, multiple lines of evidence suggest expansion of open savannah grasslands, and corresponding contraction of wooded habitats, in East Africa over the last 10 million years (Cerling et al. 1997, 2011, Bobe et al. 2002, Uno et al. 2016). These fluctuations in forest cover, and the expansion and contraction of barriers of

unsuitable habitat, have likely led to some community differences between regions while still allowing some admixture, with species such as the bay duiker *Cephalophus dorsalis* distributed across both the Congolian and Guinean forests. Genetic differentiation in other taxa, such as coffee *Coffea canephora* (Gomez et al. 2009, Labouisse et al. 2020), supports the idea that the Dahomey Gap acts as a biogeographic barrier to dispersal and gene flow. Similar differentiation can be seen in primates between populations in Congolian forests and in Kenyan and Tanzanian coastal rainforests (McDonald et al. 2022).

Historical biogeographic barriers may offer a similar explanation for the results seen in our open habitats model. Correlation plots and the site loadings from the latent variables suggest differences in community composition between West African savannahs and East Africa, and differences between the Horn of Africa and the rest of the continent. The Great Rift Valley lies perpendicular to the main gradient in the first latent variable from the open habitats model (Fig. 4). The Great Rift Valley acts as both a topographical and climatic barrier for some highland specialist species in Ethiopia (Evans et al. 2011, Freilich et al. 2016) and a mosquito species in Kenya (Lehmann et al. 1999). However, the extent to which it has posed a barrier to ungulate populations over evolutionary history requires more research.

In addition to the Great Rift Valley, historical vegetation barriers may explain the results seen in this model. As described above, extent of tropical forests in sub-Saharan Africa has fluctuated significantly, with alternating periods of relative dryness and relatively high precipitation (Dupont 2011), with climate modelling suggesting that in wet periods, tropical forest may have expanded to cover most of central Africa in a continuous belt from the west to east coasts (Cowling et al. 2008). This would have prevented species dispersal between remaining areas of savannah habitat in West Africa and the Horn of Africa with savannah habitats in southern Africa. Additionally, there is fossil evidence from sites in East Africa indicating dry phases with expanded grassy cover relative to modern conditions, indicating considerable flux in this region over recent geological time (Faith et al. 2020).

The boundary between these regions of differing community structure in the open habitats model appears to match closely the proposed boundaries between Sudanian, Somalian and Zambesian biogeographic regions proposed by Linder et al. (2012; see their Fig. 2). As well as representing boundaries between areas with distinct ungulate community composition, these boundaries reflect genetic differences within species that range across these zones which have been shown to have genetic differentiation between populations in West Africa and South/East Africa, e.g. the common warthog *Phacochoerus africanus* (Lorenzen et al. 2012). Similar genetic differentiation is seen in savannah species other than ungulates, such as ostriches (Miller et al. 2011) and lions (Barnett et al. 2006).

Do biotic effects influence ungulate communities?

Taking the above context into account, we conclude the latent variables estimated by HMSC are likely capturing a

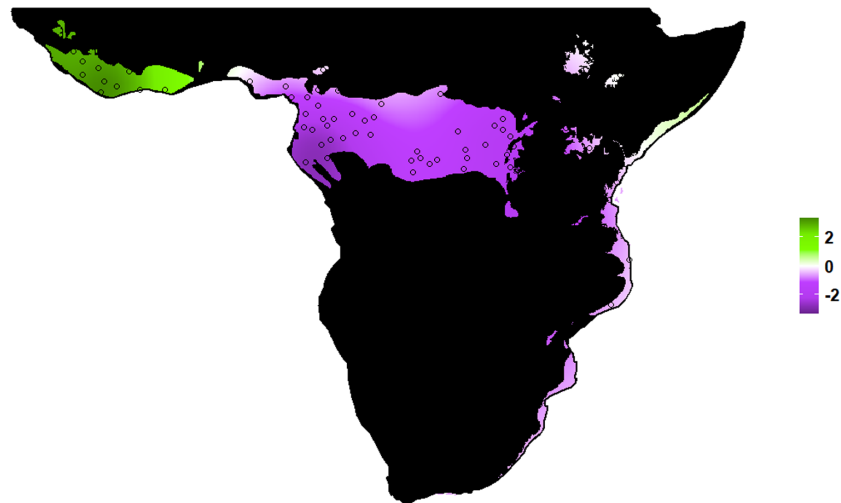


Figure 5. The site loadings values, or eta parameter, of the first latent variable from the tropical forests model for all points in the model interpolated across the entire moist tropical forest biome in the Afrotropics, using Kriging interpolation. These site loadings show differences in ungulate communities after controlling for climate differences included in the fixed effects of the model. Areas with positive values have more different communities compared to areas with negative values than expected for climate alone. Sample sites from the model used to produce this interpolation are shown by the hollow circles; note that to successfully interpolate, these points required thinning as some were very close together and therefore only 74 of the 300 used in the JSDM were used for Kriging interpolation. Areas in black indicate areas that do not belong to the moist tropical forest biome which were not included in this model and hence were not interpolated over. Site loadings values plotted without interpolation are in the supplementary materials (Supporting information).

combination of dispersal limitations due to present-day biogeographic boundaries and biotic effects. Under this interpretation, historically these barriers will have prevented population movement and gene flow between areas of contiguous suitable habitat. This would enable genetic differentiation and allopatric speciation. In the present day, these biogeographic barriers have partially disappeared (in the case of the savannah habitats covered by the open habitats model) but our model suggests that differences in community are maintained by differences in climate and by a combination of biotic effects and dispersal limitations. It is important to reiterate that the latent variables cannot be attributed automatically to either biotic effects or dispersal limitations as they estimate all variation not explained by the fixed effects in the model. Therefore, they could equally represent unexplained environmental variation not captured by the fixed effects. Indeed, if the latent variables partly capture unexplained environmental variation, this could explain some of our findings.

Firstly, it may explain the distinctness of western Namibia from the rest of southern Africa in Fig. 4A– B. This region has a very different environment to the rest of southern Africa and accordingly has a different ungulate community. Ideally, the climatic variables we have chosen should control for the abiotic factors that influence communities, but they are unlikely to do so perfectly. Specifically, climate differences alone may not fully account for the environmental differences between the Namib Desert in western Namibia, dominated by sand dunes, with the arid environments of the Kalahari that covers much of the rest of south-west Africa. We speculate that this is why this region appears distinct and may provide a good example of why latent variables must not

be treated as straightforwardly representing biotic factors (or any other one single variable).

Second, if the latent variables represent environmental variables not included as fixed effects, it may also explain the weak association between residual association and ecological trait similarity (rather than dissimilarity) in the open habitats model. In this case, we would expect that latent variables align with environmental filtering, leading to a positive association with trait similarity.

All in all, we conclude that – apart from capturing any environmental effects not explicitly modelled – the latent variables in this study most likely reflect dispersal limitations rather than biotic interactions because we did not find any evidence of a negative association between trait similarity and residual association (i.e. after controlling for climatic effects). A major role for dispersal limitations in the tropical forests model in particular would make sense, as the areas of moist tropical forest are highly fragmented in the Afrotropics whereas open habitats are broadly contiguous, with limited evidence existing for the effect of potential barriers such as the Great Rift Valley on ungulate dispersal.

An alternative explanation for the unexpected relationship between residual association and similarity may be a function of how HMSC estimates values for residual association. HMSC does not estimate the residual association between each individual species pair, as in large species communities, this would lead to a prohibitively large number of parameters requiring estimation. Instead, HMSC uses latent variables which model residual associations for all species pairs simultaneously, effectively summarising the main axes of residual variation. This greatly reduces the number of parameters that need to be estimated but it introduces the limitation

of necessarily simplifying the covariance matrix between all species in the model (Warton et al. 2015). This reduction in the number of parameters may introduce bias into the estimates of covariance between species by unintentionally forcing spurious associations between species whilst trying to approximate the overall structure of associations with fewer parameters than there are species pairs (Pichler and Hartig 2021). In certain cases, this leads to clustering of species that may show limited co-occurrence in reality, which in turn will introduce noise that may explain why the relationships between association and similarity are so weak and not in the direction we predicted. For example, in both models, we see species with little or no overlap in their ranges, such as the mountain zebra *Equus zebra* and lesser kudu *Tragelaphus imberbis* in the open habitats model (found in the south-west of Africa and in the Horn of Africa respectively), have a high omega value, indicating higher than expected residual association, despite no actual association in the data. This is because the residual association in these models are dominated by a small number of latent variables; the first shows a gradient between west African savannahs and the majority of the rest of the biome (Fig. 5), effectively inducing residual associations among species that do not fall within either of these two regions, even if they have very different ranges.

Conclusions

We found that climatic conditions were the primary driver of community assembly in Afrotropical ungulate communities, while dispersal and biotic effects (as interpreted from the latent variables estimated by HMSC) played a secondary role. These latter effects led to differences in community structure that closely match the conclusions of previous studies around the phylogeographic structure of African ungulates (Lorenzen et al. 2012). HMSC can identify ecoregions characterised by more similar species communities than would be expected from the environment alone. In addition to offering an insight into the factors determining ungulate community composition, joint species distribution models such as the one presented here may be used to predict how those communities are likely to respond to anthropogenic climate change. HMSC compares very favourably both to SDMs and other joint species distribution models in terms of predictive power (Norberg et al. 2019). Moreover, our results show that the performance of both the tropical forests model and the open habitats model is relatively high in terms of explanatory and predictive power. The latent variable approach implemented in HMSC therefore offers potential insights into factors affecting community assembly and exciting possibilities for making more powerful predictions about future distributions than we can from climatic effects alone. This is crucial at a time when anthropogenic effects are increasingly threatening ungulate populations across the African continent.

Acknowledgements – The authors gratefully acknowledge Otso Ovaskainen, Gleb Tikhonov and Jari Oksanen for providing access

to the HMSC pipeline that formed the basis of the core analyses for this paper.

Funding – AC was funded by a NERC PhD studentship under the Adapting to the Challenges of a Changing Environment (ACCE) DTP.

Author contributions

Alex Cranston: Conceptualization (equal); Formal analysis (lead); Visualization (lead); Writing – original draft (lead); Writing – review and editing (equal). **Natalie Cooper:** Conceptualization (equal); Project administration (supporting); Supervision (supporting); Writing – review and editing (equal). **Jakob Bro-Jørgensen:** Conceptualization (equal); Project administration (lead); Supervision (lead); Writing – review and editing (equal).

Transparent peer review

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/ecog.07209>.

Data availability statement

Data are available from the Dryad Digital Repository: <https://doi.org/10.5061/dryad.8pk0p2ntj> (Cranston et al. 2024).

Supporting information

The Supporting information associated with this article is available with the online version.

References

- Abraham, J. O., Hempson, G. P., Faith, J. T. and Staver, A. C. 2022. Seasonal strategies differ between tropical and extratropical herbivores. – *J. Anim. Ecol.* 91: 681–692
- Abrams, P. 1983. The theory of limiting similarity. – *Annu. Rev. Ecol. Syst.* 14: 359–376.
- Barnett, R., Yamaguchi, N., Barnes, I. and Cooper, A. 2006. The origin, current diversity and future conservation of the modern lion (*Panthera leo*). – *Proc. R. Soc. B* 273: 2119–2125.
- Belmaker, J., Zarnetske, P., Tuanmu, M.-N., Zonneveld, S., Record, S., Strecker, A. and Beaudrot, L. 2015. Empirical evidence for the scale dependence of biotic interactions. – *Global Ecol. Biogeogr.* 24: 750–761.
- Bobe, R., Behrensmeier, A. K. and Chapman, R. E. 2002. Faunal change, environmental variability and Late Pliocene hominin evolution. – *J. Hum. Evol.* 42: 475–497.
- Booth, D. J. and Murray, B. R. 2008. Coexistence. – In: Jørgensen, S. E. and Fath, B. D. (eds), *Encyclopedia of ecology*. Academic Press, pp. 664–668.
- Bullock, J. M., Edwards, R. J., Carey, P. D. and Rose, R. J. 2000. Geographical separation of two *Ulex* species at three spatial scales: does competition limit species' ranges? – *Ecography* 23: 257–271.
- Caro, T. 2005. *Antipredator defenses in birds and mammals*. – Univ. of Chicago Press.
- Cerling, T. E., Harris, J. M., MacFadden, B. J., Leakey, M. G., Quade, J., Eisenmann, V. and Ehleringer, J. R. 1997. Global vegetation change through the Miocene/Pliocene boundary. – *Nature* 389: 153–158.

- Cerling, T. E., Wynn, J. G., Andanje, S. A., Bird, M. I., Korir, D. K., Levin, N. E., Mace, W., Macharia, A. N., Quade, J. and Remien, C. H. 2011. Woody cover and hominin environments in the past 6 million years. – *Nature* 476: 51–56.
- Cosentino, F., Seamark, E. C. J., Van Cakenberghe, V. and Maiorano, L. 2023. Not only climate: the importance of biotic interactions in shaping species distributions at macro scales. – *Ecol. Evol.* 13: e9855.
- Cowling, S. A., Cox, P. M., Jones, C. D., Maslin, M. A., Peros, M. and Spall, S. A. 2008. Simulated glacial and interglacial vegetation across Africa: implications for species phylogenies and trans-African migration of plants and animals. – *Global Change Biol.* 14: 827–840.
- Cranston, A., Cooper, N. and Bro-Jorgensen, J. 2024. Data from: Using joint species distribution modelling to identify climatic and non-climatic drivers of Afrotropical ungulate distributions. – Dryad Digital Repository, <https://doi.org/10.5061/dryad.8pk0p2ntj>.
- Dinerstein, E. et al. 2017. An ecoregion-based approach to protecting half the terrestrial realm. – *BioScience* 67: 534–545.
- Du Toit, J. T. and Cumming, D. H. M. 1999. Functional significance of ungulate diversity in African savannas and the ecological implications of the spread of pastoralism. – *Biodivers. Conserv.* 8: 1643–1661.
- Dupont, L. 2011. Orbital scale vegetation change in Africa. – *Quat. Sci. Rev.* 30: 3589–3602.
- Dupont, L. M. and Weinelt, M. 1996. Vegetation history of the savanna corridor between the Guinean and the Congolian rainforest during the last 150,000 years. – *Veg. Hist. Archaeobot.* 5: 273–292.
- Evans, B., Bliss, S., Mendel, S. and Tinsley, R. 2011. The Rift Valley is a major barrier to dispersal of African clawed frogs (*Xenopus*) in Ethiopia. – *Mol. Ecol.* 20: 4216–4230.
- Faith, J. T., Rowan, J., O'Brien, K., Blegen, N. and Peppe, D. J. 2020. Late Pleistocene mammals from Kibogo, Kenya: systematic paleontology, paleoenvironments, and non-analog associations. – *J. Vertebr. Paleontol.* 40: e1841781.
- Ferretti, F. and Mori, E. 2020. Displacement interference between wild ungulate species: does it occur? – *Ethol. Ecol. Evol.* 32: 2–15.
- Ferry, N., Dray, S., Fritz, H. and Valeix, M. 2016. Interspecific interference competition at the resource patch scale: do large herbivores spatially avoid elephants while accessing water? – *J. Anim. Ecol.* 85: 1574–1585.
- Fick, S. E. and Hijmans, R. J. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. – *Int. J. Climatol.* 37: 4302–4315.
- Freeman, B. G., Strimas-Mackey, M. and Miller, E. T. 2022. Interspecific competition limits bird species' ranges in tropical mountains. – *Science* 377: 416–420.
- Freilich, X. et al. 2016. Comparative phylogeography of Ethiopian anurans: impact of the Great Rift Valley and Pleistocene climate change. – *BMC Evol. Biol.* 16: 206.
- Gagnon, M. and Chew, A. E. 2000. Dietary preferences in extant African Bovidae. – *J. Mammal.* 81: 490–511.
- Gomez, C., Dussert, S., Hamon, P., Hamon, S., Kochko, A. d. and Poncet, V. 2009. Current genetic differentiation of *Coffea canephora* Pierre ex A. Froehn in the Guineo-Congolian African zone: cumulative impact of ancient climatic changes and recent human activities. – *BMC Evol. Biol.* 9: 167.
- Gower, J. C. 1971. A general coefficient of similarity and some of its properties. – *Biometrics* 27: 857–871.
- Hardin, G. 1960. The competitive exclusion principle. – *Science* 131: 1292–1297.
- Hubbell, S. P. 2001. The unified neutral theory of biodiversity and biogeography. – Princeton Univ. Press.
- IUCN 2021. The IUCN Red List of Threatened Species (spatial data), ver. 2018-2. – <https://www.iucnredlist.org>.
- IUCN 2023. The IUCN Red List of threatened species, ver. 2023-2. – <https://www.iucnredlist.org>.
- IUCN SSC Antelope Specialist Group 2022. Species account. – <https://www.iucn.org/ssc-groups/mammals/antelope-specialist-group/species-account>.
- IUCN SSC Red List Technical Working Group, A 2021. Mapping standards and data quality for the IUCN Red List spatial data. – <https://www.iucnredlist.org/resources/mappingstandards>.
- Kingdon, J., Butynski, T. M., Happold, D. C. D., Happold, M. and Hoffmann, M. 2013. Mammals of Africa. – Bloomsbury.
- Krebs, C. J. 2009. Ecology: the experimental analysis of distribution and abundance, 6th edn. – Benjamin Cummings.
- Labouisse, J., Cubry, P., Austerlitz, F., Rivallan, R. and Nguyen, H. A. 2020. New insights on spatial genetic structure and diversity of *Coffea canephora* (*Rubiaceae*) in upper Guinea based on old herbaria. – *Plant Ecol. Evol.* 153: 82–100.
- Land Cover (GLC-SHARE) Beta-Release 1.0 Database. FAO Land and Water Division. – <https://data.apps.fao.org/map/catalog/sru/api/records/7f4f5ce4-9a3e-406f-bae6-efe63867329a>.
- Latham, J., Cumani, R., Rosati, I. and Bloise, M. 2014. FAO Global, Land and Water Division.
- Lawton, J. H. 1999. Are there general laws in ecology? – *Oikos* 84: 177–192.
- Lehmann, T., Hawley, W. A., Grebert, H., Danga, M., Atieli, F. and Collins, F. H. 1999. The Rift Valley complex as a barrier to gene flow for *Anopheles gambiae* in Kenya. – *J. Hered.* 90: 613–621.
- Linder, H. P., de Klerk, H. M., Born, J., Burgess, N. D., Fjeldså, J. and Rahbek, C. 2012. The partitioning of Africa: statistically defined biogeographical regions in sub-Saharan Africa. – *J. Biogeogr.* 39: 1189–1205.
- Lorenzen, E. D., Heller, R. and Siegismund, H. R. 2012. Comparative phylogeography of African savannah ungulates. – *Mol. Ecol.* 21: 3656–3670.
- Lortie, C. J., Brooker, R. W., Choler, P., Kikvidze, Z., Michalet, R., Pugnaire, F. I. and Callaway, R. M. 2004. Rethinking plant community theory. – *Oikos* 107: 433–438.
- Louthan, A. M., Doak, D. F. and Angert, A. L. 2015. Where and when do species interactions set range limits? – *Trends Ecol. Evol.* 30: 780–792.
- McDonald, M. M., Cunneyworth, P. M. K., Anderson, A. G. and Wroblewski, E. 2022. Mitochondrial genetic diversity and divergence dating of Angolan colobus monkeys (*Colobus angolensis*) in the eastern forests of Kenya and Tanzania: implications for speciation and reconstructing historical biogeography. – *Am. J. Primatol.* 84: e23384.
- Meier, E. S., Kienast, F., Pearman, P. B., Svenning, J.-C., Thuiller, W., Araújo, M. B., Guisan, A. and Zimmermann, N. E. 2010. Biotic and abiotic variables show little redundancy in explaining tree species distributions. – *Ecography* 33: 1038–1048.
- Miller, J. M., Hallager, S., Monfort, S. L., Newby, J. E., Bishop, K., Tidmus, S. A., Black, P. M., Houston, B., Matthee, C. A. and Fleischer, R. C. 2011. Phylogeographic analysis of nuclear and mtDNA supports subspecies designations in the ostrich (*Struthio camelus*). – *Conserv. Genet.* 12: 423–431.
- Mod, H. K., Chevalier, M., Luoto, M. and Guisan, A. 2020. Scale dependence of ecological assembly rules: insights from empiri-

- cal datasets and joint species distribution modelling. – *J. Ecol.* 108: 1967–1977.
- Norberg, A. et al. 2019. A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. – *Ecol. Monogr.* 89: e01370.
- Ovaskainen, O. and Abrego, N. 2020. Joint species distribution modelling: with applications in R. Chapter 10, *Linking HMSC back to community assembly processes.* – Cambridge Univ. Press, p. 266.
- Ovaskainen, O., Tikhonov, G., Norberg, A., Guillaume Blanchet, F., Duan, L., Dunson, D., Roslin, T. and Abrego, N. 2017. How to make more out of community data? A conceptual framework and its implementation as models and software. – *Ecol. Lett.* 20: 561–576.
- Payne, B. L. 2015. The impact of climate change on the distribution and conservation status of African antelopes. – PhD thesis, Univ. of Liverpool, UK.
- Payne, B. L. and Bro-Jørgensen, J. 2016. Disproportionate climate-induced range loss forecast for the most threatened African antelopes. – *Curr. Biol.* 26: 1200–1205.
- Pearson, R. G. and Dawson, T. P. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? – *Global Ecol. Biogeogr.* 12: 361–371.
- Pichler, M. and Hartig, F. 2021. A new joint species distribution model for faster and more accurate inference of species associations from big community data. – *Methods Ecol. Evol.* 12: 2159–2173.
- Pigot, A. L. and Tobias, J. A. 2013. Species interactions constrain geographic range expansion over evolutionary time. – *Ecol. Lett.* 16: 330–338.
- Ramesh, V., Gopalakrishna, T., Barve, S. and Melnick, D. J. 2017. IUCN greatly underestimates threat levels of endemic birds in the Western Ghats. – *Biol. Conserv.* 210: 205–221.
- Rookmaaker, K. and Antoine, P.-O. 2012. New maps representing the historical and recent distribution of the African species of rhinoceros: *Diceros bicornis*, *Ceratotherium simum* and *Ceratotherium cottoni*. – *Pachyderm* 52: 91–96.
- Salzmann, U. and Hoelzmann, P. 2005. The Dahomey Gap: an abrupt climatically induced rain forest fragmentation in West Africa during the Late Holocene. – *Holocene* 15: 190–199.
- Sexton, J. P., McIntyre, P. J., Angert, A. L. and Rice, K. J. 2009. Evolution and ecology of species range limits. – *Annu. Rev. Ecol. Evol. Syst.* 40: 415–436.
- Sinclair, A. R. E. 1985. Does interspecific competition or predation shape the African ungulate community? – *J. Anim. Ecol.* 54: 899–918.
- Sirén, A. P. K., Sutherland, C. S., Bernier, C. A., Royar, K. J., Kilborn, J. R., Callahan, C. B., Cliché, R. M., Prout, L. S. and Morelli, T. L. 2021. Abiotic stress and biotic factors mediate range dynamics on opposing edges. – *J. Biogeogr.* 48: 1758–1772.
- Soberón, J. and Nakamura, M. 2009. Niches and distributional areas: concepts, methods, and assumptions. – *Proc. Natl Acad. Sci. USA* 106: 19644–19650.
- Soria, C. D., Pacifici, M., Di Marco, M., Stephen, S. M. and Rondinini, C. 2021. Combine: a coalesced mammal database of intrinsic and extrinsic traits. – *Ecology* 102: e03344.
- Staniczenko, P. P. A., Sivasubramaniam, P., Suttle, K. B. and Pearson, R. G. 2017. Linking macroecology and community ecology: refining predictions of species distributions using biotic interaction networks. – *Ecol. Lett.* 20: 693–707.
- Staniczenko, P. P. A., Suttle, K. B. and Pearson, R. G. 2018. Negative biotic interactions drive predictions of distributions for species from a grassland community. – *Biol. Lett.* 14: 20180426.
- Staver, A. C., Archibald, S. and Levin, S. 2011. Tree cover in sub-Saharan Africa: rainfall and fire constrain forest and savanna as alternative stable states. – *Ecology* 92: 1063–1072.
- Stegen, J. C., Lin, X., Fredrickson, J. K., Chen, X., Kennedy, D. W., Murray, C. J., Rockhold, M. L. and Konopka, A. 2013. Quantifying community assembly processes and identifying features that impose them. – *ISME J.* 7: 2069–2079.
- Sutton, L., Mueter, F. J., Bluhm, B. A. and Iken, K. 2021. Environmental filtering influences functional community assembly of epibenthic communities. – *Front. Mar. Sci.* 8: 736917.
- Tikhonov, G., Opedal, Ø. H., Abrego, N., Lehtikoinen, A., de Jonge, M. M. J., Oksanen, J. and Ovaskainen, O. 2020. Joint species distribution modelling with the r-package Hmsc. – *Methods Ecol. Evol.* 11: 442–447.
- UNEP-WCMC, and IUCN 2022. Protected planet: the world database on protected areas (WDPA) and world database on other effective area-based conservation measures (WD-OECM). – www.protectedplanet.net.
- Uno, K. T., Polissar, P. J., Jackson, K. E. and deMenocal, P. B. 2016. Neogene biomarker record of vegetation change in eastern Africa. – *Proc. Natl Acad. Sci. USA* 113: 6355–6363.
- van der Loo, M. 2022. gower: gower's distance. – R package ver. 1.0.1, <https://CRAN.R-project.org/package=gower>.
- Veldhuis, M. P., Kihwele, E. S., Cromsigt, J., Ogotu, J. O., Hopcraft, J. G. C., Owen-Smith, N. and Olf, H. 2019. Large herbivore assemblages in a changing climate: incorporating water dependence and thermoregulation. – *Ecol. Lett.* 22: 1536–1546.
- Vellend, M. 2010. Conceptual synthesis in community ecology. – *Q. Rev. Biol.* 85: 183–206.
- Vilela, B. and Villalobos, F. 2015. letsR: a new R package for data handling and analysis in macroecology. – *Methods Ecol. Evol.* 6: 1229–1234.
- Warnes, G., Bolker, B., Bonebakker, L., Gentleman, R., Huber, W., Liaw, A., Lumley, T., Mächler, M., Magnusson, A. and Möller, S. 2005. gplots: various R programming tools for plotting data. – R package ver. 2, <https://cran.r-project.org/package=gplots>.
- Warton, D. I., Blanchet, F. G., O'Hara, R. B., Ovaskainen, O., Taskinen, S., Walker, S. C. and Hui, F. K. C. 2015. So many variables: joint modeling in community ecology. – *Trends Ecol. Evol.* 30: 766–779.
- Wiens, J. J. 2011. The niche, biogeography and species interactions. – *Phil. Trans. R. Soc. B* 366: 2336–2350.
- Wisz, M. S. et al. 2013. The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. – *Biol. Rev.* 88: 15–30.
- Yancovitch Shalom, H., Granot, I., Blowes, S. A., Friedlander, A., Mellin, C., Ferreira, C. E. L., Arias-González, J. E., Kulbicki, M., Floeter, S. R., Chabanet, P., Parravicini, V. and Belmaker, J. 2020. A closer examination of the 'abundant centre' hypothesis for reef fishes. – *J. Biogeogr.* 47: 2194–2209.
- Zobel, M. 1997. The relative role of species pools in determining plant species richness. An alternative explanation of species coexistence? – *Trends Ecol. Evol.* 12: 266–269.