



OPEN

Human disturbance is the major driver of vegetation changes in the Caatinga dry forest region

Helder F. P. Araujo^{1,2,3✉}, Nathália F. Canassa^{1,2}, Célia C. C. Machado⁴ & Marcelo Tabarelli⁵

Drastic changes in vegetation structure caused by exceeding ecological thresholds have fueled the interest in tropical forest responses to climate and land-use changes. Here, we examine the potential successional trajectories experienced by the largest dry tropical forest region in South America, driven by climate conditions and human disturbance. We built potential distribution models for vertebrate taxa associated with forest or shrub habitats to estimate natural vegetation cover. Distribution patterns were compared to current vegetation across the entire region to identify distinct forest degradation levels. Our results indicate the region has climatic and soil conditions suitable for more forest cover than is currently found, even in some areas with limited precipitation. However, 11.04% of natural cover persists across such an immense region, with only 4.34% consisting of forest cover. Forest degradation is characterized by the dramatic expansion of shrubland (390%), farming, and non-vegetation cover due to changes in land-use, rather than climatic conditions. Although different climate conditions have been the principal drivers for natural forest distribution in the region, the forest seems unable to resist the consequences of land-use changes, particularly in lower precipitation areas. Therefore, land-use change has exceeded the ecological thresholds for the persistence of forests, while climate change may exacerbate vegetation-type transitions.

Climate and land-use change are expected to drastically alter patterns of natural vegetation cover, particularly in the tropics^{1–3}. This emerging trend is not just related to the replacement of natural vegetation by other types of anthropogenic land use (e.g., agriculture, pasture, urban areas), but is also related to drastic changes in vegetation structure and functionality as ecological thresholds (i.e., tipping points) are exceeded^{4–6}. In general, projections indicate dry forests may be replaced by deserts (bare soils) in response to reduced precipitation and increased human disturbance, such as habitat loss and fragmentation, logging, harvesting of forest products, fire, and other chronic disturbances⁷. Such trajectories are expected to change the ability of natural vegetation to deliver ecosystem services, with impacts on both regional and global sustainability. Accordingly, there has been increasing interest in the drivers of changes in vegetation cover and their local- and global-scale effects, including the relative contribution of both climate and local human pressure on native vegetation.

Drylands have been expanding across the globe since the 1960s, and are projected to expand further throughout the twenty-first century under future climate changes^{8–11}. This transition towards treeless vegetation states is argued to emerge as a result of reductions in precipitation^{12,13}. However, such projections are based on the associations between present-day tree cover distribution and precipitation data, without the consideration of anthropogenic pressures, which may lead to a misinterpretation of the importance of climate for the distribution of natural semiarid vegetation, i.e. precipitation being assumed to be the main driver determining the distribution of natural vegetation in dry forests, when long-term anthropogenic land-use is a critical factor¹⁴. When ecohydrological characteristics and plant physiological regulations in natural drylands are considered, adverse effects of warming and drying can be mitigated by reducing water losses from soils^{15,16}. Therefore, the real contribution of land use and climate conditions to current and future changes in natural vegetation cover needs to be verified at the local- and regional-scale for drylands.

Land-use change has converted 50% of the world's natural land cover over the last 300 years¹⁷. Drylands in the Neotropics, for instance, are old colonization frontiers that are mostly covered by second-growth vegetation¹⁸. Empirical data suggests that long-term land use has reduced the resilience of the remaining natural vegetation

¹Present address: Department of Biosciences, Federal University of Paraíba, Areia, Paraíba 58397-000, Brazil. ²Postgraduate Program of Biological Sciences-Zoology, Federal University of Paraíba, João Pessoa, Paraíba, Brazil. ³Department of Biosciences, Federal University of Paraíba, Areia, PB CEP: 58397-000, Brazil. ⁴Center of Applied Biological and Social Sciences, State University of Paraíba, João Pessoa, Paraíba, Brazil. ⁵Department of Botany, Federal University of Pernambuco, Recife, Pernambuco, Brazil. ✉email: helder@cca.ufpb.br

in semiarid regions, with the rainy season offering little in the way of a boost to rates of recovery¹⁹. Landscape structure also plays a major role in controlling and defining successional vegetation trajectories in dry tropical regions, as faster recovery occurs in sites surrounded by higher forest cover²⁰. Therefore, in addition to desertification, second-growth vegetation can represent a degradation-driven alternative stable state in dry forest regions²¹. Forest degradation is defined as a state of anthropogenically-induced arrested succession, where ecological processes that underlie forest dynamics are disturbed or severely constrained²². Consequently, alternative stable states via forest-to-shrub conversions can be found²³, and these phenomena can mask the real interaction between climate and natural vegetation cover in semi-arid regions.

The Caatinga region is one of the most populated and biodiverse seasonally dry tropical forests (SDTF) globally, which has undergone intense transformations (including desertification) over the past five centuries, increasing its vulnerability to climate change²⁴. Historical records from the eighteenth and nineteenth centuries demonstrate that the Caatinga was characterized by extensive dry forests with high levels of biodiversity^{25,26}. However, slash-and-burn agriculture, free-ranging livestock, exploitation of forest products, and cycles of intensive agriculture (e.g., cotton) have converted large tracts of old-growth forest into vegetation mosaics dominated by shrubby vegetation and regenerating forest stands, areas of bare soil, and agricultural fields and pastures, across extensive areas during the twentieth and twenty-first centuries²⁷. Based on remote sensing data, some estimates suggest that the Caatinga has more than 60% natural vegetation cover^{28,29}, but these estimates consider all shrubland as natural vegetation and do not consider historical forest degradation caused by forest-to-shrub conversions (Table 1). Thus, the potential ecosystem services provided by undisturbed natural vegetation may be underestimated in the Caatinga, since soil and vegetation carbon stocks drop drastically following land-use change, even from dry forest to shrubland, for instance³⁰. Nevertheless, we recognize that distinguishing between degradation-induced forest-to-shrub conversion and undisturbed natural shrubland is hard since this is also a natural vegetation type throughout the region, and remains a challenge.

In this context, paleoecological patterns (including faunal composition) can provide insights into the relationships between climate and natural vegetation types³¹. Due to past climate change, larger-sized mammals were narrowly distributed across scattered patches of suitable habitats in South America throughout the late Quaternary³². In the Holocene, the loss of larger-sized mammal lineages was related to the decrease of open vegetation areas and an increase in dry forests in South American semiarid regions³³. On the other hand, the life-history traits of current endemic bird species associated with different vegetation types (forest or shrubland, for instance) support the occurrence of a natural gradient from forest to open vegetation composing the original cover in South American drylands³⁴. However, some current species associated with open vegetation can increase their distribution because they benefit from anthropogenic disturbance, and, therefore, these contemporary changes hamper efforts to identify current patterns of natural vegetation distribution.

Here, we (1) estimate the potential distribution of natural Caatinga vegetation cover, considering climate and soil; (2) estimate land-cover change; and (3) identify the relative contribution of humans and climate as drivers of these land-cover changes. Three questions guided the study: (a) What would the expected distribution of dry forest and shrubland be in the absence of acute human disturbance? (b) How much of the natural distribution has been modified? (c) What are the relative contributions of human activities and climate conditions to vegetation change in distinct degradation transitions (forest to shrub, forest/shrub to farming, forest/shrub to non-vegetation) (Table 1)? To answer these questions, we used potential distribution models of strictly forest-dependent endemic birds as a proxy of forest habitats and potential distribution models of extinct larger-sized mammals for LGM, projected to current climate conditions, as a proxy of shrubland distribution across the region. Further, (1) these models were contrasted with the present-day vegetation across the entire region, and (2) the documented vegetation changes were contrasted with anthropogenic and climatic variables.

Results

Potential vegetation distribution

Our vegetation distribution models indicated a high probability of potential forest distribution stretching across the south, center, and west of the Caatinga region, and patches in the east. On the other hand, our models indicated a high probability of shrub habitats occurring in the Caatinga, particularly in the center and east of the region (Fig. 1, Supplementary Fig. S1). All models were influenced by environmental variables, with area under the curve (AUC) values greater than 0.8 for all models (i.e., a non-random distribution of vegetation types). Of the environmental variables we assessed, the mean temperature of wettest quarter (25.3%), temperature seasonality (16.9%), and annual precipitation (13.5%) were the most important, explaining 55.7% of forest distribution. While annual precipitation (57.3%) and annual mean temperature (10.7%) explained 68% of the distribution of shrub habitats (Supplementary Table S1).

Change in vegetation cover

Considering the spatial limits of the Caatinga biota (i.e., 863,752 km²), we predicted a total potential area of 731,211 km² for forest cover, and 132,195 km² for shrub cover, which corresponds to approximately 84.6% and 15.3% of the entire region, respectively. However, our analysis of present-day cover indicated that only 31,793 km² is covered by forest, 63,639 km² by natural shrubland, and 766,235 km² by degraded vegetation (452,128 km² by forest to shrubland, 300,822 km² by forest/shrubland to farming, and 13,284 km² by non-vegetation cover) (Fig. 2). These scores indicated an extensive loss of forest, i.e., only 4.34% of their potential distributions remaining. On the other hand, shrubland has expanded by 390% by grouping undisturbed natural shrubland and forest to shrubland change (Fig. 2). Furthermore, 35% were modified to farming and 1.6% to non-vegetation cover; i.e. desertification areas.











View		LULC type	
Ground	Aerial	Present paper	MapBiomias
		Forest (<i>no change</i>)	Forest formation
		Shrub (<i>no change</i>)	Savanna formation
		Forest to shrub (<i>degradation-level 1</i>)	Savanna formation
		Forest/shrub to farming (<i>degradation-level 2</i>)	Pasture or agriculture
		Forest/shrub to non vegetation (<i>degradation-level 3</i>)	Other non vegetated areas

Table 1. Ground and aerial views of the land use and land cover (LULC) found in the Caatinga region. The LULC types used in this paper were classified in degradation levels (*italics*) and contrasted with LULC types derived from current satellite data (MapBiomias²⁹).

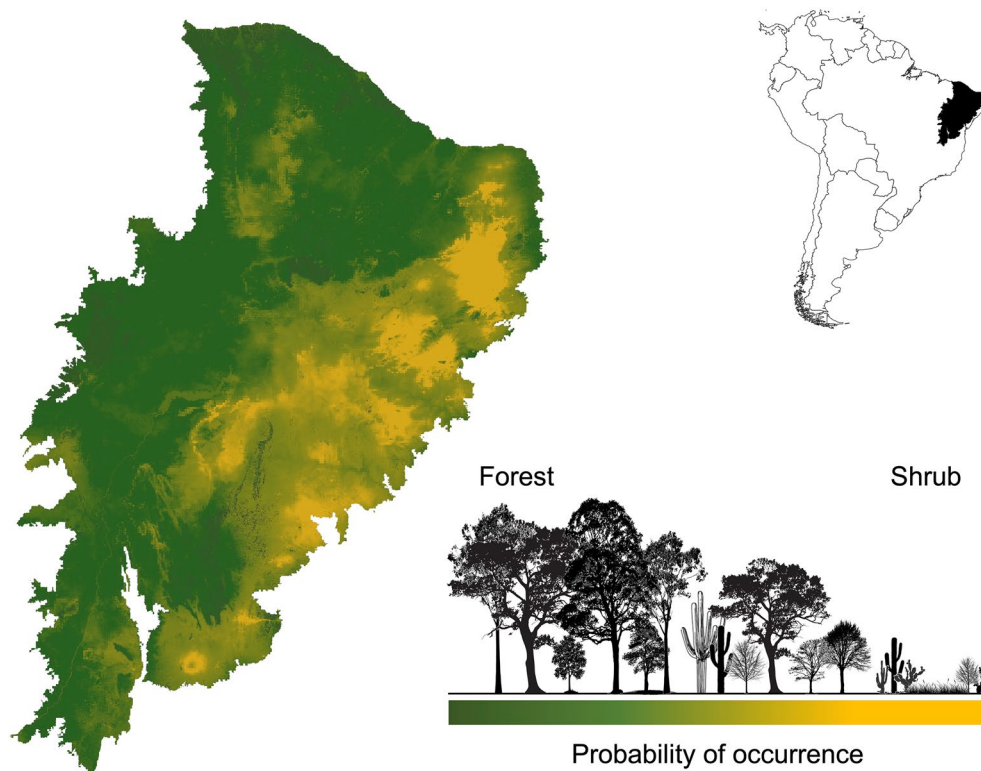


Figure 1. Potential distribution of forest and shrubland in the Caatinga for current climatic and soil conditions. The map was generated from the Maxent model results (see methods for details) using the software QGIS version 3.20.0-Odense (<https://qgis.org/en/site/>).

Drivers of vegetation cover change

Our random forest model predicted c. 73% ($R^2=0.73$, $MAE=0.14$, $RMSE=0.20$) vegetation cover change. The human footprint was the variable with the highest relative importance (100%), consistently appearing in all models and the best-performing model. On the other hand, climatic variables (precipitation seasonality, annual precipitation, and temperature seasonality) contributed less than 15% each (Fig. 3). However, the highest degradation levels were weakly associated with areas where current temperature seasonality is higher, and annual precipitation and precipitation seasonality are lower (Supplementary Fig. S2).

Discussion

Our results indicate that the Caatinga has climatic and soil conditions able to support more forest cover than is currently found throughout the region, regardless of the current occurrence of low-precipitation areas (Supplementary Fig. S3). However, our estimates indicate that only 11.04% of native vegetation cover persists across the whole region, with only 4.34% being forest cover. On the other hand, forest degradation is characterized by the dramatic expansion of shrubland (390%), farming, and non-vegetation cover due to changes in land-use, rather than climatic conditions. Climate conditions have been the major driver for forest and natural shrubland distribution in the Caatinga region, with areas climatically suitable for both vegetation types. However, the forest seems unable to resist the consequences of land-use changes, particularly in those areas where the temperature seasonality is high and annual precipitation is low. The patterns we have uncovered highlight the resistance the Caatinga has to different climatic conditions while reinforcing the prominence of land-use change as the main driver of the rapid decline in forest cover and the proliferation of reduced tree-cover vegetation types such as shrublands.

Although our approach was able to identify the potential forest distribution across areas with past forest occurrence, confirmed by historical records from the eighteenth and nineteenth centuries^{25,26}, it is worth mentioning the potential limitations of the model sensitivity. Even using climate extremes and soil variables to predict potential vegetation distribution cover at a regional scale (Supplementary Table S2), local conditions may support particular vegetation types, which are not detected by regional-level assessments. However, our findings indicated that the forest and shrubland could potentially occur across a wide range of climate conditions, which naturally reduces the prominence of local conditions (particularly soil conditions) as the main driver of vegetation type. This perspective is supported by the present-day occurrence of Caatinga dry forest across a wide range of soil conditions, such as deep sandy soils to shallow and clay-based relatively rich soils³⁵.

Globally, precipitation and water availability are the major drivers of vegetation distribution³⁶ by controlling biological processes from plant physiology to ecosystem dynamics and structure³⁷. In the Caatinga region, the

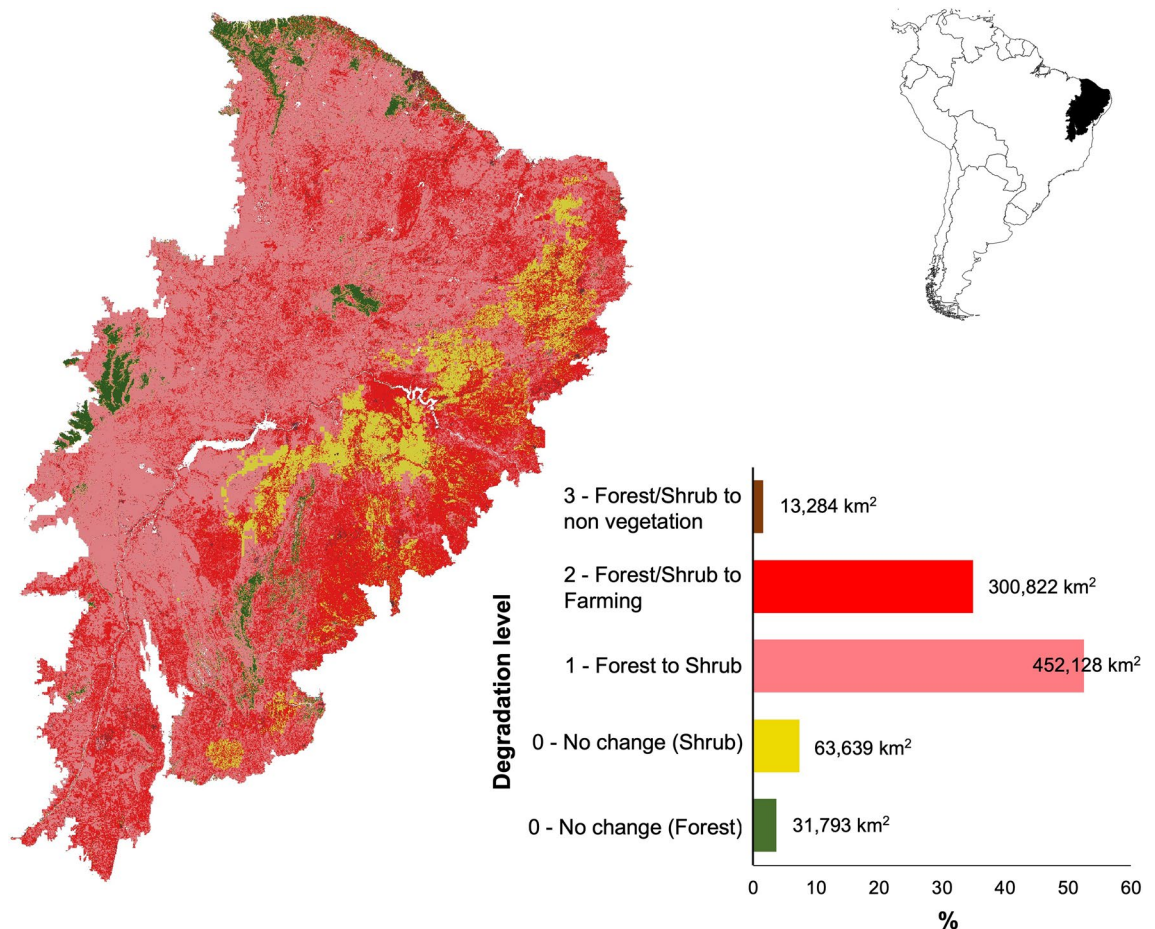


Figure 2. Distribution and amount of each forest degradation level in the Caatinga region. Forest degradation is defined as a state of anthropogenically induced arrested succession, where ecological processes that underlie forest dynamics are diminished or severely constrained²². The map was generated by comparing potential distribution and current land cover (see methods for details) using the software QGIS version 3.20.0-Odense (<https://qgis.org/en/site/>).

precipitation regime is mainly associated with warm and humid air masses coming from the west (Continental Equatorial), southeast (Atlantic Tropical), and north-northeast (Atlantic Equatorial), as well as with the relief which contributes to the occurrence orographic rains³⁸. The Caatinga biota dates back to the Miocene and since then it has experienced climate changes^{39,40}. Moreover, the Caatinga dry forest stretches over a wide spectrum of soils and precipitation conditions, which results in gradients of water availability occurring from local to regional spatial scales. In addition to successive climate changes, the Caatinga biota has long been exposed to frequent droughts, commonly associated with El Niño Southern Oscillation events⁴¹. This evolutionary and ecological context has probably favored plant lineages able to cope with increased ecological filtering such as those with conservative resource-use strategies, high plasticity, and the ability to resprout as previously documented⁴². In fact, the resistance of the Caatinga to different climatic conditions and its ability to cope with climate oscillations through time has been considered a key asset in a global change scenario⁴³. However, this apparent resistance to climate oscillations has not been sufficient to resist local human disturbances, resulting in the expansion of shrublands and deserts⁴⁴. The Caatinga has been subjected to intense human disturbances since the arrival of Europeans in the sixteenth century⁴⁵. We refer to an immense package of disturbances including slash-and-burn agriculture, extensive cattle raising, and the exploitation of forest products. Such disturbance is varied and intense including the collection of firewood, wood for charcoal production, timber, and fodder, plus the periodic cycles of commercial agriculture devoted to commodities such as cotton^{46–48}.

Our findings confirm land-use change as the main driving force of forest loss and distinctive successional trajectories towards shrublands and anthropogenic-induced deserts. This trajectory affects ecosystem services provided by the natural vegetation cover, such as soil/vegetation carbon stocks, which drastically drop in the case of biomass, when forest, for example, is converted to shrublands^{30,43,44}. Present-day satellite data indicates that aboveground forest biomass can reach 80–130 Mg/ha in the Caatinga, but such a high-biomass forest covers only about 7% of the region today, while 73% of its area support aboveground biomass < 40 Mg/ha⁴⁹. Landscape-level data obtained through direct measurements have also found reduced aboveground forest biomass across successional mosaics resulting from slash-and-burn agriculture²⁷, with several sites exhibiting low rates of biomass recovery (i.e. reduced resilience). We refer to secondary/regenerating forest stands (up to 70-yr

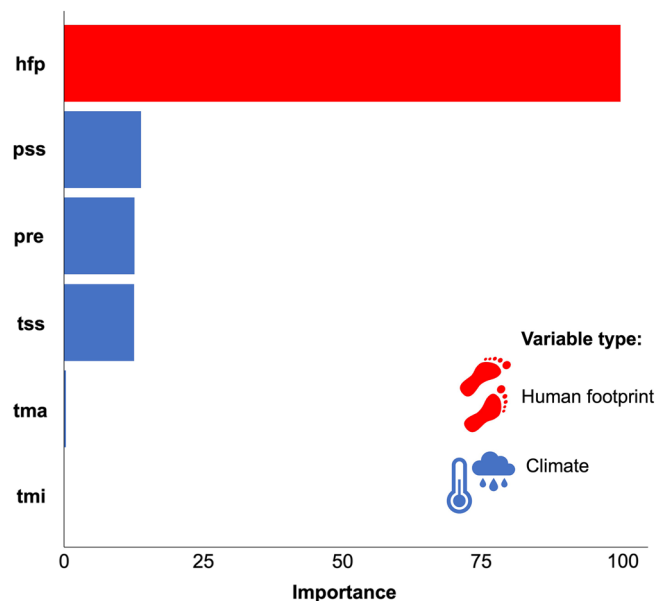


Figure 3. Importance values of human footprint (hfp) and climate variables on the forest degradation in the Caatinga region. Importance values were calculated based on a Random Forest model that explained 73% of the variation in the data. Climate variables were: precipitation seasonality (pss), annual precipitation (pre), temperature seasonality (tss), max temperature (tma), and min temperature (tmi).

old) limited to 50 Mg/ha on average^{49,50}. In fact, dry forest regeneration and dynamics are influenced by local land use and landscape context, where long-term land use affects the potential for recovery, particularly when soils are degraded⁵¹, regeneration mechanisms (e.g. seeds and sprouts) are compromised, and sites are located far from old-growth forest patches^{19,20}. In other words, forest persistence has already been compromised across the majority of the Caatinga region, because more than 90% of the potential forest distribution has already been modified, as a result of disruptions in forest regeneration mechanisms at the regional scale. Therefore, the land-use change has exceeded ecological thresholds (i.e., tipping points) for the persistence of forests. This scenario demands active forest restoration initiatives across the entire region if society is to have any chance of protecting the irreplaceable biodiversity and ecosystem services of the Caatinga.

Considering the socioecological context of the Caatinga, with the prevalence of subsistence farming, we argue that this transition of dry forests naturally dominated by single-stem tree species towards deserts (i.e. bare soils) or vegetation type dominated by multi-stemmed shrubs (here defined as shrubland) is a response by the forest ecosystems to a continuous exposition to conversion into pastures^{17,52}, slash-and-burn agriculture⁵³, shifting cultivation^{27,54}, coppicing²³, and browsing by livestock^{48,55}, as well as the exploitation of forest products. Such an alternative successional trajectory is not exclusively related to tree species replacement by shrub species, but also by the predominance of tree species occurring as multi-stem shrub-sized individuals, as is already occurring with almost all dominant tree species in human-modified Caatinga landscapes^{56–58}. Such forest transition towards shrublands as an alternative successional trajectory has been previously proposed for other dry forests exposed to intensive land use⁵⁹. However, to what extent regeneration favoring shrubs rather than trees (one of the stages towards human-induced desertification in the Caatinga) meets the concept of alternative stable states, remains to be investigated.

In synthesis, natural cover in the Caatinga has been replaced by deserts⁴⁴ and shrublands due to historical and current anthropogenic land use. Land-use intensification is already causing tree cover and aboveground biomass decline, not only due to deforestation but also due to changes in the structural status of the remaining vegetation, such as the spread of shrublands, at the expense of forests. This transition toward treeless land cover reduces ecosystem services, as is already occurring in tropical dry forests around the world (i.e. forest degradation), affecting billions of people who directly depend on forest products and their regulating services^{7,12,59}. Although much attention has been given to the occurrence of potential precipitation thresholds which, if exceeded, may permanently reduce tree cover and biomass, comparatively little attention has been paid to land-use intensification, which is already causing fast declines in tree cover and biomass. However, climate change may be a synergistic force that will intensify (where precipitation is reduced and drought events are more frequent) or retard (where precipitation increases) regeneration dynamics representing alternative trajectories in human-modified landscapes¹⁵. In this context, it might be misleading to adopt current patterns of vegetation distribution across human-modified landscapes as a reliable indicator to infer vegetation sensitivity to climate change. Finally, we propose that landscapes composed of a combination of forests and crop and/or cattle farming may allow high food productivity, biodiversity, and ecosystem services, even during extreme drought events^{60,61}. However, dry forest restoration and the adoption of better practices to prevent further degradation are urgently needed to help the recovery of ecosystem productivity and resilience for the sake of global sustainability.

Methods

Study area

The Caatinga region in northeast Brazil is one of the world's largest SDTFs⁶², which originally covered almost one million km²⁶³. This biota is restricted to the Brazilian territory, and it covers crystalline basement surfaces and sand sedimentary basins supporting flattened surfaces cut by narrow valleys, residual hills, and high-altitude plateaus (Supplementary Fig. S4). Soil types are diverse, ranging from shallow, rocky, and quite fertile to deep, sandy, and unfertile. The climate is classified as BSh (hot semi-arid) and As (with dry summer) across most of the region⁶⁴. Most of the rainfall is concentrated in three consecutive months, but the region experiences spatial and annual variations. Due to the great inter-annual variability of precipitation, droughts can endure for years³⁸.

At the regional level, the Caatinga dry forest was originally a vegetation mosaic dominated by dry forest stands of varying structure (high- to low-statured forests)⁶⁵, which also supported enclaves of both semideciduous rainforest and Cerrado (i.e. Brazilian savanna), particularly due to the occurrence of patchy orographic rain. The Caatinga dry forest is highly diverse, with levels of species endemism ranging from 5 to 25%⁶⁵. This unique biota has been intensively modified since the arrival of Europeans in the sixteenth century by a combination of (1) slash-and-burn agriculture, (2) cycles of intensive agriculture devoted to commodities across particular sites, (3) free-ranging livestock (principally goats and cattle), and (4) exploitation of forest products, such as firewood for domestic use and for charcoal production, fodder, and poles for farming facilities⁴⁶. Rural smallholdings are prevalent throughout the region, with residents' livelihoods depending on natural resources and ecosystem services such as nutrient provision for subsistence agriculture.

To address our questions, three complementary and integrative steps were performed as described below.

Potential distribution of forest and shrub vegetation

The potential forest and shrub distribution at the biota scale under current climatic conditions was estimated using (1) taxa associated with forest habitats under present climate conditions, and (2) vertebrate taxa associated with open vegetation during the Last Glacial Maximum (LGM; 21 ka BP). More precisely, we used records of strictly forest-dependent endemic bird species as a proxy, since these species do not benefit from anthropogenic changes in the landscape. In the Caatinga, 31% of bird species are forest-dependent and such habitat filtering reduces the influences of spatial dynamics associated with the conservation status of the region³⁴. Our database consisted of occurrence records of six endemic bird species: *Conopophaga cearae*, *Hylopezus ochroleucus*, *Xiphocolaptes falcirostris*, *Lepidocolaptes wagleri*, *Phylloscartes roquettei*, *Arremon franciscanus*. Species' occurrences were retrieved from the avifauna literature and the following online databases: Global Biodiversity Information Facility (Gbif—<https://www.gbif.org>), Species link (<http://www.splink.org.br>), and Wikiaves (<https://www.wikiaves.com.br>).

We did not use present-day endemic species to model open vegetation cover because these species can benefit from anthropogenic processes that reduce forest cover³⁴. Therefore, we used extinct megafauna occurrences in the Caatinga to model potential shrub-dominated environments. Occurrences of megafauna taxa that lived at the end of the Pleistocene were collected from papers on animal paleoecology. The interpretations of paleoecology were based on dental morphology and studies involving carbon isotopes. Taxa were selected based on literature focusing on grazing behavior, C4 plant-based diets (grasses), and occurrence in open areas (Supplementary Table S3). Our database consisted of occurrence records of eight species: *Equus neogaeus*, *Glossotherium sp.*, *Glyptodon clavipes*, *Glyptotherium cylindricum*, *Pampatherium humboldtii*, *Panochthus greslebini*, *Panochthus jaguaribensis*, and *Xenorhinotherium bahiense*. The occurrences of the species were collected from bibliographic surveys of fossil findings in the Caatinga, paleontology collections (Department of Geology of the Federal University of Pernambuco, Vertebrate and Paleontology Laboratory of the Federal University of Paraíba), Federal University of Bahia (from Prof. Dr. Mario Dantas), and Paleobiology networks (<https://www.paleobiodb.org>).

Distribution modeling was carried out using 393 geographic points in the Caatinga region, 196 from forest cover indicators under current climatic and soil conditions, and 197 from shrub environment indicators under LGM (21 ka BP) climatic conditions. We used 19 climatic variables obtained from Worldclim (<http://www.worldclim.org>) and PaleoClim (<http://www.paleoclim.org>). The climatic variables used to model forest cover, obtained through present-day forest bird occurrence, were selected after an autocorrelation analysis of the current climatic conditions in the Caatinga region. Variables used for shrub environments were selected after an autocorrelation analysis of the LGM climatic conditions (Supplementary Table S2). We included in the models the variables with a correlation < 0.8 to avoid redundancy of climatic variables. We used Pearson's correlation, calculated using the *Vegan* package for R (version 3.5.3). Physical soil properties (Clay and Sand content in g/kg) also were used as variables to model current forest cover. These variables were obtained from digital soil mapping (SoilGrids) based on a global compilation of soil profile data (WoSIS) and environmental layers (<http://www.soilgrids.org>). Further information about the SoilGrids and WoSIS projects is available on <http://isric.org>.

The distribution models were generated with the Maximum Entropy (Maxent) algorithm⁶⁶, which models a probability distribution where each grid cell has predicted the suitability of conditions, from a set of environmental variables and georeferenced occurrence records. The highest value of area under the curve (AUC) was used to select the best models. An AUC ≥ 0.75 indicated that the modeled distribution was not random but influenced by environmental variables. We use Receiver Operator Characteristic statistics (ROC) to assess model accuracy with 10 repetitions of 10,000 interactions, with 10% of replicas being randomly selected as test data and the remaining 90% used for model training in each replica.

To map the potential distribution of natural vegetation cover, we summed the raster file of the current potential distribution of forest with the reprojected raster file of the current potential distribution of shrubs, derived from the LGM model. For that, we reproject the LGM model to the current climate scenario using Maxent and used the raster calculator to multiply the likelihood of shrub occurrence by - 1. So, we summed the positive

probability (0 to 1) for the potential distribution of forest with the negative probability (0 to -1) that represents the potential distribution of shrub environments.

Estimating vegetation cover change

To assess regional changes in vegetation cover, the potential distribution of forest and shrub vegetation for current conditions was compared with present-day vegetation cover, assessed by MapBiomass²⁹ for the year 2021. We use current forest-bird occurrence data to calibrate the threshold expected of the probability distribution of forest, transforming the raster file of the potential distribution of natural vegetation covers into a binary map (forest vs shrub). This binary map was compared with MapBiomass land-use and land-cover type, and the difference was categorized in forest degradation levels (0—no change, when forest or shrub remained in Forest and Savanna Formation by MapBiomass, respectively; 1—forest to shrub, when Savanna Formation occur where the forest was predicted; 2—forest/shrub to farming, when Pasture or Agriculture occur where forest or shrub was predicted; 3—forest/shrub to non-vegetation, when Non-Vegetated Areas occur where forest or shrub was predicted (Table 1).

Drivers of vegetation cover change

To identify the relative contribution of humans and climate as drivers of the land-cover changes, we used a hexagonal grid covering the entire Caatinga. We use 12,976 hexagons with a side of 5km (area 64.95 km²), and statistical zonal to obtain information from degradation level as response variable and human footprint and climate as predictor variables. The degradation level raster was obtained from the previous step, where each pixel (30 × 30 m) has a value between 0 and 3 (0—no change, 1—forest to shrub, 2—forest/shrub to farming, 3—forest/shrub to non-vegetation). The Human Footprint (hfp) was used as a human pressure variable and was obtained from a global map terrestrial Human Footprint⁶⁷ (1 km² resolution). The Human Footprint map indicates human pressure weighted within that range according to estimates of their relative levels of human pressure for the extent of built environments, population density, electricity infrastructure, crop lands, pasture lands, roads, railways, and navigable waterways⁶⁷. Climate variables used as predictors were precipitation seasonality (pss), annual precipitation (pre), temperature seasonality (tss), max temperature (tma), and min temperature (tmi) (1 km² resolution). These climate data were obtained from Worldclim⁶⁸. We extracted the mean value of each variable from each hexagon (Fig. 4).

To evaluate the relationship between the predictor variables and the response variable, we used the Random Forest (RF) predictive analysis (Predictive Model with Machine Learning)⁶⁹. In general, the RF algorithm is

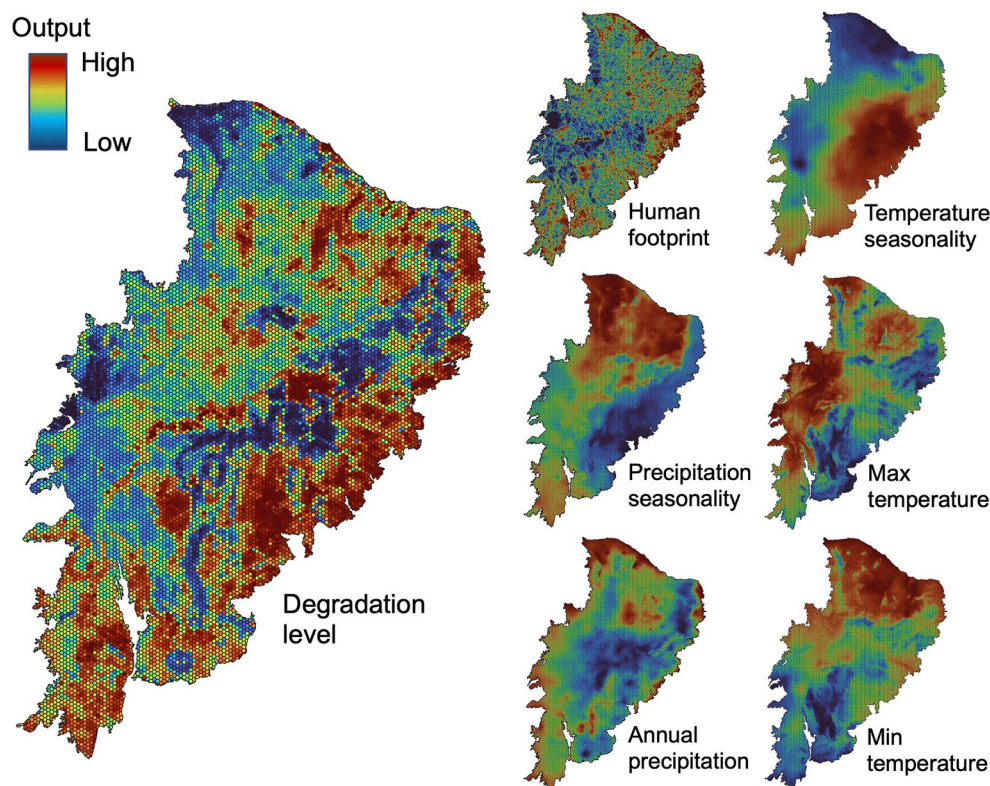


Figure 4. Distribution of the degradation levels (as a response variable) and human pressure and climate variables (as predictor variables) in the Caatinga region. The relationship between the predictor variables and the response variable was evaluated using a Random Forest predictive model (see results and Fig. 3). The maps were generated using the software QGIS version 3.20.0-Odense (<https://qgis.org/en/site/>).

robust and one of the most used in predictive model analysis. It has few assumptions and data pre-processing and generation of predictive models with high performance, designed to avoid overfitting⁶⁹. Through decision tree algorithms, RF produces classifications or regression models, as in the case evaluated here. Tuning of the model parameters was carried out to increase the performance of the model. In our RF analysis, hfp, pss, pre, tss, tma, and tmi are used as predictor variables, and degradation level (0 to 3) as the response variable. Analyses were performed using the caret package in R⁷⁰.

To evaluate the performance of the models, RMSE (Root Mean Square Error), R², and MAE (mean absolute error) were used. We selected the best model using the bestTune and finalModel functions from the caret package. In the final model, the selected variables are organized in order of importance. The importance of the variable is a value that measures the number of times the performance of the model increased when a specific predictor was included in the model. Therefore, the importance of the variable is closely related to the performance of the model and the relative values are scaled between 0 and 100⁷⁰.

Data availability

The datasets used in this study are available from the corresponding author upon reasonable request.

Received: 17 January 2023; Accepted: 21 October 2023

Published online: 27 October 2023

References

- Findell, K. L. *et al.* The impact of anthropogenic land use and land cover change on regional climate extremes. *Nat. Commun.* **8**, 989 (2017).
- Pecl, G. T. *et al.* Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. *Science* **355**, eaai9214 (2017).
- Sales, L. P., Galetti, M. & Pires, M. M. Climate and land-use change will lead to a faunal “savannization” on tropical rainforests. *Glob. Change Biol.* **26**, 7036–7044 (2020).
- Salazar, L. F. & Nobre, C. A. Climate change and thresholds of biome shifts in Amazonia. *Geophys. Res. Lett.* **37**, n/a-n/a (2010).
- Lenton, T. M. *et al.* Tipping elements in the Earth’s climate system. *Proc. Natl. Acad. Sci.* **105**, 1786–1793 (2008).
- Nobre, C. A. *et al.* Land-use and climate change risks in the Amazon and the need of a novel sustainable development paradigm. *Proc. Natl. Acad. Sci.* **113**, 10759–10768 (2016).
- Mirzabaev, A. *et al.* Desertification. In *Climate Change and Land: An IPCC Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse Gas Fluxes in Terrestrial Ecosystems* 249–338 (2019).
- Feng, S. & Fu, Q. Expansion of global drylands under a warming climate. *Atmos. Chem. Phys.* **13**, 10081–10094 (2013).
- Joubert, D. F., Rothauge, A. & Smit, G. N. A conceptual model of vegetation dynamics in the semiarid highland savanna of Namibia, with particular reference to bush thickening by *Acacia mellifera*. *J. Arid Environ.* **72**, 2201–2210 (2008).
- Loarie, S. R. *et al.* The velocity of climate change. *Nature* **462**, 1052–1055 (2009).
- Huang, J. *et al.* Dryland climate change: Recent progress and challenges. *Rev. Geophys.* **55**, 719–778 (2017).
- Hirota, M., Holmgren, M., Van Nes, E. H. & Scheffer, M. Global resilience of tropical forest and Savanna to critical transitions. *Science* **334**, 232–235 (2011).
- Anadón, J. D., Sala, O. E. & Maestre, F. T. Climate change will increase savannas at the expense of forests and treeless vegetation in tropical and subtropical Americas. *J. Ecol.* **102**, 1363–1373 (2014).
- Rito, K. F., Arroyo-Rodríguez, V., Queiroz, R. T., Leal, I. R. & Tabarelli, M. Precipitation mediates the effect of human disturbance on the Brazilian Caatinga vegetation. *J. Ecol.* **105**, 828–838 (2017).
- Lian, X. *et al.* Multifaceted characteristics of dryland aridity changes in a warming world. *Nat. Rev. Earth Environ.* **2**, 232–250 (2021).
- Berg, A. & McColl, K. A. No projected global drylands expansion under greenhouse warming. *Nat. Clim. Change* **11**, 331–337 (2021).
- Ellis, E. C., Klein Goldewijk, K., Siebert, S., Lightman, D. & Ramankutty, N. Anthropogenic transformation of the biomes, 1700 to 2000. *Glob. Ecol. Biogeogr.* **19**, 589–606 (2010).
- Chazdon, R. L. *et al.* Carbon sequestration potential of second-growth forest regeneration in the Latin American tropics. *Sci. Adv.* **2**, e1501639 (2022).
- Lipoma, M. L. *et al.* Low resilience at the early stages of recovery of the semi-arid Chaco forest—Evidence from a field experiment. *J. Ecol.* <https://doi.org/10.1111/1365-2745.13622> (2021).
- Pérez-Cárdenas, N. *et al.* Effects of landscape composition and site land-use intensity on secondary succession in a tropical dry forest. *For. Ecol. Manag.* **482**, 118818 (2021).
- Tabarelli, M., Filgueiras, B. K. C., Ribeiro, E. M. S., Lopes, A. V. & Leal, I. R. Tropical dry forests. In *Encyclopedia of Biodiversity* (Elsevier, 2023) (in Press).
- Ghazoul, J., Burivalova, Z., Garcia-Ulloa, J. & King, L. A. Conceptualizing forest degradation. *Trends Ecol. Evol.* **30**, 622–632 (2015).
- Bart, R. R., Tague, C. L. & Moritz, M. A. Effect of tree-to-shrub type conversion in lower Montane forests of the Sierra Nevada (USA) on streamflow. *PLoS ONE* **11**, e0161805 (2016).
- Silva, J. L. S. E., Cruz-Neto, O., Peres, C. A., Tabarelli, M. & Lopes, A. V. Climate change will reduce suitable Caatinga dry forest habitat for endemic plants with disproportionate impacts on specialized reproductive strategies. *PLoS ONE* **14**, e0217028 (2019).
- Kury, L., Neves, E. F., Gesteira, H. M. & Jakobsson, A. *Sertões Adentro-Viagens Nas Caatingas Séculos XVII A XIX* (Andrea Jakobsson Estúdio, 2013).
- de Souza, B. I. & Souza, R. S. The occupation process of the Cariri Velhos—PB and the effects in the vegetal cover: Contribution to the cultural biogeography of the semiarid. *Caderno de Geografia* **26**, 229–258 (2016).
- Souza, D. G. *et al.* Multiple drivers of aboveground biomass in a human-modified landscape of the Caatinga dry forest. *For. Ecol. Manag.* **435**, 57–65 (2019).
- Beuchle, R. *et al.* Land cover changes in the Brazilian Cerrado and Caatinga biomes from 1990 to 2010 based on a systematic remote sensing sampling approach. *Appl. Geogr.* **58**, 116–127 (2015).
- MapBiomias. Projeto MapBiomias—Coleção [versão] da Série Anual de Mapas de Uso e Cobertura da Terra do Brasil. <https://plataforma.brasil.mapbiomas.org> (2022).
- Menezes, R. S. C. *et al.* Soil and vegetation carbon stocks after land-use changes in a seasonally dry tropical forest. *Geoderma* **390**, 114943 (2021).
- Bakker, E. S. *et al.* Combining paleo-data and modern enclosure experiments to assess the impact of megafauna extinctions on woody vegetation. *Proc. Natl. Acad. Sci. USA* **113**, 847–855 (2016).

32. Lima-Ribeiro, M. S., Nogués-Bravo, D., Terribile, L. C., Batra, P. & Diniz-Filho, J. A. F. Climate and humans set the place and time of Proboscidean extinction in late Quaternary of South America. *Paleogeogr. Paleoclimatol. Paleoecol.* **392**, 546–556 (2013).
33. de Vivo, M. & Carmignotto, A. P. Holocene vegetation change and the mammal faunas of South America and Africa. *J. Biogeogr.* **31**, 943–957 (2004).
34. Araujo, H. F. P. & da Silva, J. M. C. The avifauna of the Caatinga: biogeography, ecology, and conservation. In *Caatinga* (eds Silva, J. M. C. *et al.*) 181–210 (Springer International Publishing, 2017). https://doi.org/10.1007/978-3-319-68339-3_7.
35. Castro Oliveira, G., Francelino, M. R., Arruda, D. M., Fernandes-Filho, E. I. & Schaefer, C. E. G. R. Climate and soils at the Brazilian semiarid and the forest-Caatinga problem: New insights and implications for conservation. *Environ. Res. Lett.* **14**, 104007 (2019).
36. Staver, A. C., Archibald, S. & Levin, S. A. The global extent and determinants of Savanna and forest as alternative biome states. *Science* **334**, 230–232 (2011).
37. Nemani, R. R. Climate-driven increases in global terrestrial net primary production from 1982 to 1999. *Science* **300**, 1560–1563 (2003).
38. Prado, D. As caatingas da América do Sul. In *Ecologia e Conservação da Caatinga* (eds Leal, I. R. *et al.*) 3–73 (Editora Universitária da UFPE, 2003).
39. Pennington, R. T., Lewis, G. P. & Ratter, J. A. *Neotropical Savannas and Seasonally Dry Forests: Plant Diversity, Biogeography, and Conservation* (CRC Press–Taylor & Francis Group, 2006).
40. de Araujo, H. F. P. *et al.* The Caatinga region is a system and not an aggregate. *J. Arid Environ.* **203**, 104778 (2022).
41. Pereira, M. P. S. *et al.* Brazilian dry forest (Caatinga) response to multiple ENSO: The role of Atlantic and Pacific Ocean. *Sci. Total Environ.* **705**, 135717 (2020).
42. Barros, M. F. *et al.* Resprouting drives successional pathways and the resilience of Caatinga dry forest in human-modified landscapes. *For. Ecol. Manag.* **482**, 118881 (2021).
43. Sfair, J. C., de Bello, F., de França, T. Q., Baldauf, C. & Tabarelli, M. Chronic human disturbance affects plant trait distribution in a seasonally dry tropical forest. *Environ. Res. Lett.* **13**, 025005 (2018).
44. Vieira, R. *et al.* Identifying areas susceptible to desertification in the Brazilian northeast. *Solid Earth* **6**, 347–360 (2015).
45. Siqueira-Filho, J. A. A extinção inexorável do rio São Francisco. In *Flora das caatingas do rio São Francisco: história natural e conservação* 24–65 (Andrea Jajobsson Estúdio, 2012).
46. Mamede, M. D. A. & de Araújo, F. S. Effects of slash and burn practices on a soil seed bank of caatinga vegetation in Northeastern Brazil. *J. Arid Environ.* **72**, 458–470 (2008).
47. Costa, S. R. & Bueno, M. G. *A saga do algodão: das primeiras lavouras à ação na OMC* (Insight Engenharia, 2004).
48. Schulz, K. *et al.* Grazing reduces plant species diversity of Caatinga dry forests in northeastern Brazil. *Appl. Veg. Sci.* **22**, 348–359 (2019).
49. Castanho, A. D. A. *et al.* A close look at above ground biomass of a large and heterogeneous Seasonally Dry Tropical Forest-Caatinga in North East of Brazil. *Anais da Academia Brasileira de Ciências* <https://doi.org/10.1590/0001-37652020190282> (2020).
50. Cabral, G. A. L., Sampaio, E. V. S. B. & Almeida-Cortez, J. S. Spatial Structure and aboveground biomass in different Caatinga succession stages, in Santa Terezinha, Paraíba. *Revista Brasileira de Geografia Física* **6**, 566–574 (2013).
51. Lawrence, D. *et al.* Ecological feedbacks following deforestation create the potential for a catastrophic ecosystem shift in tropical dry forest. *Proc. Natl. Acad. Sci.* **104**, 20696–20701 (2007).
52. Garcia, C. A. *et al.* The global forest transition as a human affair. *One Earth* **2**, 417–428 (2020).
53. Bezerra, J. S. *et al.* Drastic impoverishment of the soil seed bank in a tropical dry forest exposed to slash-and-burn agriculture. *For. Ecol. Manag.* **513**, 120185 (2022).
54. Sobrinho, M. S. *et al.* Land use, fallow period and the recovery of a Caatinga forest. *Biotropica* **48**, 586–597 (2016).
55. Lins, L. *et al.* Exotic goats do not affect the seed bank but reduce seedling survival in a human-modified landscape of Caatinga dry forest. *For. Ecol. Manag.* **522**, 120491 (2022).
56. Melo, F. P. L. The socio-ecology of the Caatinga: Understanding how natural resource use shapes an ecosystem. In *Caatinga* 369–382 (Springer International Publishing, 2017). https://doi.org/10.1007/978-3-319-68339-3_14.
57. Riegelhaupt, E. M. & Pareyn, F. G. C. A questão energética. In *Uso sustentável e conservação dos recursos florestais da Caatinga* (eds Gariglio, M. A. *et al.*) 65–75 (Serviço Florestal Brasileiro/MMA, 2010).
58. da Silva, J. M. C. & Barbosa, L. C. F. Impact of human activities on the Caatinga. In *Caatinga* 359–368 (Springer International Publishing, 2017). https://doi.org/10.1007/978-3-319-68339-3_13.
59. Murphy, P. G. & Lugo, A. E. Ecology of tropical dry forest. *Annu. Rev. Ecol. Syst.* **17**, 67–88 (1986).
60. de Araujo, H. F. P. *et al.* A sustainable agricultural landscape model for tropical drylands. *Land Use Policy* **100**, 104913 (2021).
61. Araujo, H. F. P., Machado, C. C. C., Alves, A. C. F., Costa Lima, M. & Silva, J. M. C. Vegetation productivity under climate change depends on landscape complexity in tropical drylands. *Mitig. Adapt. Strateg. Glob. Change* **27**, 54 (2022).
62. da Silva, J. M. C., Barbosa, L. C. F., Leal, I. R. & Tabarelli, M. The Caatinga: Understanding the challenges. In *Caatinga* 3–19 (Springer International Publishing, 2017). https://doi.org/10.1007/978-3-319-68339-3_1.
63. IBGE. *Biomass e sistema costeiro-marinho do Brasil: compatível com a escala 1:250 000*. (2019).
64. Alvares, C. A., Stape, J. L., Sentelhas, P. C., de Moraes Gonçalves, J. L. & Sparovek, G. Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift* **22**, 711–728 (2013).
65. de Queiroz, L. P., Cardoso, D., Fernandes, M. F. & Moro, M. F. Diversity and evolution of flowering plants of the Caatinga domain. In *Caatinga* (eds Silva, J. M. C. *et al.*) 23–63 (Springer International Publishing, 2017). https://doi.org/10.1007/978-3-319-68339-3_2.
66. Phillips, S. J., Anderson, R. P., Dudík, M., Schapire, R. E. & Blair, M. E. Opening the black box: An open-source release of Maxent. *Ecography* **40**, 887–893 (2017).
67. Venter, O. *et al.* Global terrestrial human footprint maps for 1993 and 2009. *Sci. Data* **3**, 160067 (2016).
68. Fick, S. E. & Hijmans, R. J. WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315 (2017).
69. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
70. Kuhn, M. caret: Classification and Regression Training. Preprint at (2021)

Acknowledgements

This work was supported by the Brazilian National Council for Scientific and Technological Development (CNPq) [grant numbers 441436/2017-0 and 405298/2021-8], Public Call n. 09/2021 Fundação de Amparo à Pesquisa do Estado da Paraíba (FAPESQ-PB) grant n. 3101/2021, and n. 03 Produtividade em Pesquisa PROPESQ/PRPG/UFPB grant n. PVN13311-2020. N. F. Canassa has a doctoral scholarship from the CNPq.

Author contributions

H.F.P.A. and C.C.C.M. conceived the ideas; H.F.P.A., N.F.C., and C.C.C.M. analyzed the data; all authors drafted and included relevant intellectual contributions to improve and conclude the manuscript.

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/s41598-023-45571-9>.

Correspondence and requests for materials should be addressed to H.F.P.A.

Reprints and permissions information is available at 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) 2023