# *The legacy of past human land use in current patterns of mammal distribution*

# The legacy of past human land use in current patterns of mammal distribution

Ester Polaina (e.polaina@gmail.com)
Conservation Biology Department, Doñana Biological Station-CSIC, Seville, Spain
Ecology Department, Swedish University of Agricultural Science, Uppsala, Sweden
ORCID ID: 0000-0002-5064-5881

Manuela González-Suárez (manuela.gonzalez@reading.ac.uk)
Conservation Biology Department, Doñana Biological Station-CSIC, Seville, Spain
School of Biological Sciences, University of Reading, Reading, UK
ORCID ID: 0000-0001-5069-8900

Eloy Revilla (revilla@ebd.csic.es)
Conservation Biology Department, Doñana Biological Station-CSIC, Seville, Spain
ORCID ID: 0000-0001-5534-5581

**Corresponding author:** Ester Polaina

**Acknowledgements**

We thank the IUCN Red List team for making and maintaining its database freely available online. We also acknowledge the Laboratory for Anthropogenic Landscape Ecology, directed by Dr. Erle C. Ellis, for publicly sharing anthropogenic land cover change datasets. This work was funded by the program '*Junta para la Ampliación de Estudios*' (JAEPre022. BOE-A-2011-10745, co-funded by the European Social Fund), the European Community's Seventh Framework Programme (FP7/ grant n° 235897 and

EU BON project nº 308454), and the Spanish Ministry of Science and Innovation co-funded by FEDER (CGL2009-07301/BOS, CGL2012-35931/BOS and JCI-2011-09158).

**Abstract**

Multiple environmental factors are known to shape species distributions at the global scale, including climate and topography, but understanding current extents of occurrence and biodiversity patterns requires considering anthropogenic factors as well. Numerous studies have explored the relationship between contemporary human activities and different biodiversity metrics, but the influence of past activities, such as land-use, remains poorly understood despite being one of the oldest human impacts. Here we evaluate the role of past land-use modifications in the current distribution and conservation status of mammals worldwide using spatial data characterizing human land use from c.B.C.6000 to c.A.D.2000. First, we applied a clustering method that revealed three generalized past human land-use trajectories that represent *low-*, *recently-* and *steadily-used* areas widely represented across the globe. Second, we fitted boosted regression trees to predict total and threatened mammalian richness, globally and within trajectory-clusters, testing the role of environmental factors and multiple human land-use metrics reflecting: total used area at different time spans, rates of land-use change, and the occurrence of remarkable land-use shifts. Environmental factors were identified as the main correlates of current mammalian richness, but several proposed metrics of past land-use were also relevant predictors. Overall, these results highlight the likely existence of a land-use legacy in some regions of the world that has influenced the distribution of extant mammals, particularly of those currently classified as threatened. Even if we cannot change that legacy, our results show that we need to account for past human impacts to understand present biodiversity patterns and, arguably, to guide future actions.

**Keywords:** land-use history, land-use trajectory, KK10 model, IUCN, threatened species, distribution, cluster analysis

## Introduction

Threatened species are unevenly distributed across the world, with remarkable differences among taxonomic groups (Grenyer et al. 2006). If extinctions occurred by chance, we would expect more threatened species in areas with higher species richness which is, in turn, largely determined by latitudinal gradients in climatic conditions affecting the availability of energy and water (Hawkins et al. 2003, Terribile et al. 2009). However, extinctions do not occur by chance but are instead largely shaped by human activities, such as direct persecution or habitat modifications (Russell et al. 2013); and extreme climatic events, such as past ice ages or current climate change (Thuiller et al. 2011, Varela et al. 2015). As a result, the relationship between environmental factors, overall species richness, and the number of threatened species is not straightforward and shows spatial heterogeneity (Orme et al. 2005, Ceballos and Ehrlich 2006, Brum et al. 2013). Environmental and evolutionary factors define the initial, pristine pool of species within a certain area, but human activities can lead to threats, and eventually, local extinctions that can result in lower overall species richness. Therefore, the number and/or intensity of threats within a given area can affect total richness and the number of threatened species.

From a global perspective, more humanized areas, such as those with increased accessibility or with more space allocated to agriculture, have been associated with relatively lower species richness at both fine (~10 x 10km grid size) and large resolutions (~100 x 100km grid size; Martins et al. 2014; Torres-Romero & Olalla-Tárraga 2014). Fewer studies have explored correlates of the global distribution of threatened species. These studies have often used broad study units such as ecoregions or countries, and have linked the occurrence of threatened species with human pressures, measured in terms of land use or human population density (Lenzen et al. 2009, Pekin and Pijanowski 2012, Brum et al. 2013) and with different socioeconomic profiles (Polaina et al. 2015). Polaina et al. (2018) analyzed threatened species richness at a finer resolution (1x1º grid cells) and showed that the relationship between this biodiversity metric and land-use-change derived impacts varied across regions, likely reflecting different stages of human development and appropriation of land. In some areas, threatened species were more abundant in more impacted areas, whereas in

others threatened species remained in less impacted zones. On the other hand, patterns of threatened species distribution have been studied within the framework of spatial conservation planning, considering high richness as an attribute to be prioritized, but not explicitly searching for correlates or drivers of those patterns (Bonn et al. 2002; Grenyer et al. 2006; spatial resolutions ranging from ~25 x 25km to ~100 x 100km grid sizes). It seems clear that whereas the most threatening activities are normally related to lower total species richness, the relationship between threats and threatened species likely varies across regions and spatial scales in more complex ways.

One additional but generally overlooked factor that may help explain current biodiversity patterns is past human pressure (Faurby and Svenning 2015). In general, human land uses develop from low impact extensive into high impact intensive industrial agriculture and urbanization (Foley et al. 2005). Since the beginning of sedentary human societies and the advent of agriculture, around B.C.8000-6000, the amount of land under human dominance has grown at an accelerating pace (Ellis 2011). Some events marked particularly notable transitions, such as the European invasions in the 15$^{th}$ century, the 19$^{th}$ century Industrial Revolution with its productive, technological and demographic shifts and, more recently, the Green Revolution that triggered the so-called 'great acceleration' (c.A.D.1950, Steffen et al. 2015). Throughout all this time, the evidence of human impacts on other mammals has become increasingly apparent. Some known examples include the extinctions of megafauna in the late Pleistocene-early Holocene (~B.C.8000; Sandom et al. 2014); the declining population trend of many European species coinciding with the decline of forested areas (c.A.D.1000; Kaplan et al. 2009; Crees et al. 2016); or the range contractions of many North American species after European settlement (Laliberte and Ripple 2004). The available evidence on the importance of these past human-induced events suggests past land-use modifications could help explain contemporary biodiversity patterns at broad scales (Loehle and Eschenbach 2012, Faurby and Svenning 2015).

How past human activities influence current species richness distribution remains largely unexplored at broad scales, partly because data have only recently become available (Goldewijk et al. 2010, Kaplan et al. 2011). Therefore, studies exploring the correlation between past human

activities and current distribution at the macroecological scale are very scarce, although some have demonstrated that these indicators may improve the understanding of present patterns of species distribution (Dullinger et al. 2013). Although the effects of past land-use change on current biodiversity distribution are indirect, mediated by unknown or uncertain past species' distributions; finding correspondences between both factors would provide evidence of an extinction debt. Thus, current biodiversity patterns would actually be an inflated picture of what it will come in the future, increasing current concerns regarding the sixth mass extinction (Ceballos et al. 2017, WWF 2018).

Our study aims to provide an understanding of how past human land use relates with current global mammalian biodiversity patterns. We used terrestrial mammals as a model study group because their present conservation status (IUCN 2014) and their past dynamics of decline (e.g. Turvey & Fritz 2011; Prescott et al. 2012) are generally well known. First, we proposed a new application of an approach to identify generalized trajectory-clusters areas on the common history of human land-use expansion or reduction according to estimated changes in the portion of land classified as used by humans at different time spans ranging from c. B.C.6000 to c. A.D.2000. Second, we proposed and tested how diverse metrics characterizing past land use and the defined trajectory-clusters can contribute to our understanding of the current distribution of total mammalian richness, number of threatened species and proportion of threatened species, globally and within each trajectory-cluster. The proposed metrics reflect three processes we hypothesize could affect current observed distribution of mammals richness: (1) proportion of human-used land at different time spans, if there is a time lag between human pressure and species response, past human land use extent may explain species richness distribution better than present land use; (2) rates of land-use change, if rapid changes limit the ability of species to respond via adaptation, areas that had been modified at a faster pace may have fewer and more threatened species today; and (3) remarkable land-use events, if abrupt past human impacts, even if later reversed, have lasting and irreversible effects, areas in which extensive land use occurred in the past may have fewer total and more threatened species.

## Material and Methods

*Data sources and selection*

Data of mean proportion of land use per unit area at different time spans were obtained from Ellis et al. (2013; available at http://ecotope.org/products/datasets/used_planet/). We chose the KK10 model (Kaplan et al. 2011), which assumes that humans tend to use land more intensively –against extensively– when population density is high and land scarce (Boserup 1965). This model is considered as more realistic than the alternative HYDE (Goldewijk et al. 2010), which provides predictions based only on nearly linear relationships between population and surface of used land (Ellis et al. 2013). We considered the 10 time breaks available within the selected dataset: B.C.6000, B.C.3000, B.C.1000, A.D.0, A.D.1000, A.D.1500, A.D.1750, A.D.1900, A.D.1950 and A.D. 2000. From the original 5-arc-minute resolution (~10 km), mean values were averaged for a 110 x 110 km grid (≈ 12,100 km²) in a Behrmann cylindrical equal area projection. Grid-cells with an emerged area smaller than 10,000 $km^2$ were excluded to avoid comparing grid-cells with very unequal areas. Working at finer resolutions without overestimating species richness would not be possible using available species geographic range data (Hurlbert and Jetz 2007). The portion of present human land use (per grid-cell) was used as an indicator of current land-use pressure, given its known importance as an anthropogenic driver of habitat loss and deterioration (Foley et al. 2005). For the purpose of the present work, it was represented as the value of human land use at c.A.D.2000, the most recent time break available on this data source. We considered several past land-use metrics based on the three hypotheses presented in the introduction: (1) proportions of human-used land at each of the 10 time spans, (2) rates of land-use change for different time periods, and (3) remarkable land-use events. The proportion of land intended for human use at each time break is directly provided by the KK10 model (Kaplan et al. 2011). Land-use rates of change were calculated as the difference in the portion of a grid-cell defined as used in a given time period standardized per 1000 elapsed years. Time periods were defined based on available time spans and aimed to capture major historical land-use periods at our spatial resolution and extension: prehistoric (c.B.C.6000-c.A.D.0), pre-industrialization (c.A.D.0-1750), industrialization

(c.A.D.1750-1950) and post-industrialization (c.A.D.1950-2000). Remarkable land-use changes were considered at three different thresholds: 20% grid-cell intended for human use ($LU_{+20}$), defined as the first significant use by Ellis et al. (2013); 50% grid-cell intended for human use ($LU_{+50}$), taken to reflect a high human-use value (> 6000 $km^2$); and maximum value of use per grid-cell for the whole time series ($LU_{max}$), a relative value to account for the expected differences among regions, i.e., the highest value in low-used areas may be the lowest observed in high-used regions. It is worth to mention that not all grid-cells exceeded the proposed absolute thresholds (20% and 50%), but all of them had a relative maximum value. For each threshold we calculated, *when* it occurred (time break as presented above), *how much* land was intended for human land (exact portion of land used at each time of remarkable land use), and *how long* this value was maintained (duration in years until A.D.2000; Table S2.1).

Present environmental conditions were synthesized for each 110 km grid-cell in terms of annual mean actual evapotranspiration (AET), obtained from Zhang et al. (2010); annual mean temperature and precipitation, obtained from WorldClim2 (Fick and Hijmans 2017); and mean elevation, extracted from the global digital elevation model GTOPO30 (LP DAAC 2004; Table S2.2). Environmental indicators were included to account for the known latitudinal gradient in species diversity distribution, mainly associated with water and energy availability (Hillebrand 2004, Torres-Romero and Olalla-Tárraga 2014).

Distribution maps for 5237 terrestrial mammal species were gathered from the International Union for Conservation of Nature (IUCN 2014), selecting only polygons classified as native, extant and probably extant. Distribution data were intersected with the 110 x 110 km grid, and species were considered as present in a particular grid-cell when any overlap existed. Total richness was the sum of all terrestrial mammal species (with any conservation status) occurring in a grid-cell. The number of threatened mammals represented the sum of all species categorized as vulnerable (VU), endangered (EN) or critically endangered (CR) by the IUCN (IUCN 2014) occurring in a grid-cell. The proportion of threatened mammals was calculated by dividing the number of threatened mammals over the total mammal richness per grid-cell. Total mammalian richness, in models where

total threatened richness was the response, controlled for the fact that the more species within an area, the greater the chances to find more of them under threat.

*Statistical analyses*

To synthetize trajectory trends in longitudinal data of global land use we employed a clustering method that incorporates a *k*-means algorithm (Celeux and Govaert 1992) implemented in the *kml* package ('kml' function; Genolini et al. 2015) in R v.3.2.3 (R Core Team 2015). This method allowed us to group grid-cells with similar land use trajectories along the considered temporal range (c.B.C.6000 - c.A.D. 2000). The optimal number of clusters was defined using five non-parametric quality indices according to different criteria: three variants of the Calinski & Harabanz criterion, the Ray-Turi criterion and the Davies-Bouldin criterion (Calinski and Harabasz 1972, Davies and Bouldin 1979, Ray and Turi 1999, Kryszczuk and Hurley 2010). All of them try to minimize the within-cluster compactness while maximizing the between-cluster spacing, and are standardized within the *kml* package to allow comparisons (Genolini et al. 2015). Using this approach, we classified areas that had followed similar land-use transformations across time. The obtained trajectory-clusters were used as a categorical past land-use predictor in global models to test differences among clusters in terms of mammalian diversity and vulnerability. In addition, we fit separated models for each trajectory-cluster to identify specific past land-use indicators that explain current mammalian richness distribution within clusters.

As expected, some of the initially considered variables were highly correlated (Spearman's $\rho \geq |0.7|$), so we kept only one variable from each correlated couple, according to the following rules. From correlated pairs with one control variable (present land use, total richness -when applicable- and environmental indicators) and one past land-use indicator we selected the control variable. From correlated pairs with two control variables we selected the one with the highest correlation to the response. Finally, from correlated pairs of two past land-use indicators we selected the past land-use indicator representing the oldest temporal span to increase the contrast with present land use (a control variable). We checked for correlation separately for the global and each cluster's

subset of data (Figs. S2.1-S2.4), so the retained predictors varied among the global and by-trajectory-cluster models, but all models contained at least one indicator from each of the hypothesized effects (proportion of human-used land at different time spans, rates of land-use change for different time periods, and remarkable land-use events).

The selected indicators were included as predictors in global and by-trajectory-cluster boosted regression trees (BRT), testing three different response variables: mammalian richness, number or proportion of threatened mammals over total richness. BRT modeling uses a boosting technique to combine large numbers of relatively simple tree models to optimize predictive performance. BRT allow for detecting nonlinear relationships and including variables of very different nature and units (Elith et al. 2008). To fit the models we used the function 'gbm.step', which calculates the optimal number of boosting trees using 10-fold cross validation, and it is included in the *dismo* package (Hijmans et al. 2013) in R. A bagging fraction of 0.5, a tree complexity of 5 (up to 5-way interactions) and a learning rate of 0.005 were fixed for all BRT models to achieve a minimum of 1,000 trees, according to the guidelines in Elith et al. (2008). To account for spatial autocorrelation, all models included a residuals-based autocovariate (RAC) that specifies the relationship between the value of the residuals at each location and those at neighboring locations (the 8 immediate grid-cells surrounding each cell as neighbors, approximately within a 165 km distance in our case) from a model excluding spatial autocorrelation. Deriving the autocovariate from the residuals allows including only the remaining deviance unexplained by the explanatory variables, thus the actual influence of the predictors is better captured (Crase et al. 2012). A Poisson error structure for the response variable was assumed for total and threatened mammals' counts. When the proportion of threatened mammals was explored as a response, this was arcsine square root transformed to approximate normality and a Gaussian error structure was used instead. Grid-cells containing a proportion of threatened mammals greater than 0.999 were excluded (N=12) from the analyses for being considered outliers corresponding to grid-cells with very few species (proportion of threatened mammals per grid-cell, mean=0.049, third quartile= 0.667).

A predictor was considered to be relevant when its relative importance was greater than expected by chance (100% divided by the number of variables included in each model; e.g. Müller et al. 2013). The explanatory power of each model was calculated as the percentage of deviance explained respect to a null model, defined as one without any splits –equivalent to an intercept only model in linear regression (Ferrier and Watson 1997). The effect of each predictor was described in relation to the fitted model in which all other predictors were set to their average.

**Results**

All quality criteria supported the differentiation of three generalized trajectory-clusters describing global temporal patterns in past land use from c.B.C.6000 to c.A.D.2000 (Fig. S1.1). These three trajectory-clusters (Fig.1A) correspond to three broad global patterns of land-use progression that we named as: *low-used areas* (51.9% of grid-cells) where land-use values were low with only small increases in use over time; *recently-used areas* (32.3% of grid-cells), where the rate of land encroachment was moderate until relatively recent times (~A.D.1750) when a strong increase in land use was observed; and *steadily-used areas* (15.8% of grid-cells), where initial land use was higher than in the rest of trajectories and increased at a relatively constant rate with a soft steepening around B.C.1000 and a very recent decline. *Low-used areas* are mainly located in regions of low primary productivity, such as deserts and boreal forests; but also in productive biomes that may have remained largely unused due to difficulties for humans to access them, like the tropical forests of Borneo and the Amazon (Fig. 1B). *Recently-used areas* largely correspond to territories of relatively modern human colonization and expansion, such as North America, Australia or southern and East Africa. *Steadily-used areas* include the main cradles of ancient human settlements, including parts of the Middle East, Europe, India, eastern China, the Sahel and Central America (Fig. 1B). While quality criteria supported three clusters, there were additional configurations that had partial support. These suggested further division of *recently* and *steadily-used areas* into more clusters with *low-used areas* always remaining as a single group (Fig. S1.2).

All models showed overall good values of percentage of explained deviance (77-96%) and most of the initial spatial autocorrelation was accounted for by including the spatial autocovariate, although it did not control it completely, which may be explained by the high spatial aggregation of the biodiversity metrics employed in this study (Tables 1-3 vs S3.1-S3.3).

Global models showed no relevance of the trajectory-cluster to explain differences in total mammalian richness, number or proportion of threatened mammals' distributions, being the least relevant variable in all BRT models (Tables 1-3). Globally, total mammal richness was, as expected, predominantly influenced by environmental factors, namely actual evapotranspiration, with the highest mammalian richness coinciding with intermediate values of AET (Table 1; Fig. 3). The number of threatened species was primarily affected by total richness, being positively related (Table 2; Fig. 2), while the proportion of threatened species was predominantly related to mean annual temperature. Higher proportions of threatened mammals were found where mean annual temperatures were extremely low or high, and the same trend applies in relation to AET, and at high elevations (Table 3; Fig. 3).

Within-cluster trends were broadly similar, although in some cases past land-use indicators appeared as relevant. Both total mammalian richness and proportion of threatened species distribution were mainly explained by environmental indicators; with the exception of the relevance of past land use (c.A.D.1000) on total richness (Table 1) and pre-industrial rate of land-use change on proportion of threatened mammals within *recently-used* areas (Table 3). Slightly more indicators of past land use appeared as important to explain the distribution of total threatened mammals within all trajectory-clusters (Table 2).

Within *low-used* areas, effects directions were the same as described for the global model and, additionally, more threatened mammals were found where pre-industrial land-use change was relatively faster (Fig. 4). In *recently-used* areas, mammal richness was greater where AET and temperatures were higher, at intermediate elevations, and in areas where land use was relatively high c.A.D. 1000; more threatened mammals tended to concentrate where total richness and land use c.B.C.6000 were larger; as for the proportion of threatened mammals, this tended to be higher at

higher elevations, extremely low or high mean annual temperatures and AET values, and where changes in land use were more rapid during the pre-industrial period (Figs.3 and 4). *Steadily used* areas showed higher total richness at relatively high values of AET, temperature and altitude. More threatened mammals predominantly occurred where total mammalian richness, temperature and precipitation were higher, and where pre-industrial land-use changes were either slow (or even negative) or relatively fast. Finally, the highest proportions of threatened mammals tended to co-occur with relatively high temperatures, precipitations and elevations (Figs.3 and 4).

**Discussion**

Our results show that land-use history across the world can be broadly summarized into three trajectory-clusters: *low-*, *recently-* and *steadily-used* areas. Although there are not net differences in the total or threatened (absolute or proportional to the total) number of terrestrial mammals among clusters, there are disparities in the influence of different predictors in explaining mammal diversity metrics within each cluster, and in the shape of the relationship between predictors and responses.

As shown in previous research, we found environmental variables are generally the best predictors of global differences in the distribution of mammal richness and proportion of threatened mammals, and absolute numbers of threatened mammals are generally correlated with total richness. However, when disaggregating by trajectory-cluster, we detected descriptors of past land-use change as relevant. This signal of the past is most noticeable for *recently-used* areas and when assessing differences in total numbers of threatened mammals. Current land use or descriptors of remarkable past land-use changes were not identified as relevant in any model.

*Low-used areas*

According to our results, more than 50% of the global land area analyzed (excluding Antarctica and most of Greenland) have followed a *low-used* trajectory. This high share of the global surface may explain why general trends in the effects of environmental and land-use predictors within these areas resemble the global trends. These areas broadly coincide with last-of-the-wild regions,

traditionally seen as opportunities to preserve biodiversity given the relatively low human influence to which they are exposed (Sanderson et al. 2002). Lack of historical human pressure may be explained by two different reasons: low primary productivity associated to biomes in the northern hemisphere, such as the boreal forests, the Tundra, and deserts worldwide; and relative remoteness, limiting accessibility in some tropical forests, e.g. Amazon or Borneo (Fig. 1).

Total mammal richness and proportion of threatened mammals follow the classical latitudinal gradients reflecting variation in environmental conditions, with the proportion of threatened mammals being also high at high elevations, likely corresponding to the Andes Mountains and the Tibetan plateau, which present high levels of endemism but are not particularly diverse (Miehe et al. 2014; Figs. 3 & S4.1-4). Thus, *a priori*, within these areas, tropical regions such as the Congo Basin or areas of South East Asia –rich in mammals, with many of them under threat and relatively unmodified, as shown here– might be candidates to preserve some of the most threatened fauna of the world if no further changes occur. However, future predictions are not very optimistic for some of these regions, which are identified as at high risk of severe biodiversity loss under different scenarios of agricultural development (Kehoe et al. 2017).

More interesting for our aim is the high relative importance of pre-industrial rate of land-use change to explain the distribution of total numbers of threatened mammals. The fact that more threatened species tend to occur in areas more rapidly transformed during the period c.A.D. 0-1750 suggests the existence of a land-use legacy on these parts of the Earth, where mammals remain negatively influenced by past human impacts and local extinctions have not yet occurred given the low magnitude of land-use changes (Bürgi et al. 2017).

*Recently-used areas*

Around 32% of global land is classified as a *recently-used* trajectory, which coincide with areas humanized after the great colonization events of the 15$^{th}$ century onwards. Many of these regions are located within highly developed countries, such as the United States or Australia (Fig. 1). Today, these areas do not present particularly high species richness or accumulations of threatened

mammals (Table S2.1), thus they are not generally considered a global priority (Brooks et al. 2006). However, some of the regions included in this category, such as southern South America, have been previously labelled as crisis ecoregions, where habitat conversion is occurring much faster than protection (Hoekstra et al. 2005).

Differently from *low-used* areas, environmental conditions correlated with mammal diversity metrics do not present such a clear geographic pattern and may reflect local conditions (Figs. S4.7-S4.9). As for the land-use-history variables of interest, total mammalian richness is lower where the proportion of land use c.A.D.1000 had been relatively low (Fig. 4), opposite to what we expected under a time-lag effect (lower species richness predicted for areas heavily modified in the past). Therefore, there is not a lagged effect of past land-use changes on current total richness, which remains relatively high within these areas (Table S2.1). On the other hand, the direction of the relationships between land use c.B.C.6000 and numbers of threatened mammals (more threatened species where areas were more converted in the past), and pre-industrial rate of change and proportion of threatened mammals (Fig. 4) seems to agree with the initial predictions. Overall, there is evidence of the existence of an extinction debt in threatened species but not for total richness. Considering that these areas were largely unmodified until the second half of the 18$^{th}$ century, it is possible that we would see changes in total richness as biodiversity inventories are updated in the future at the global scale.

*Steadily-used areas*

Less than 16% of global land belongs to this trajectory-cluster, which is characterized by a relatively high and long-lasting human land use covering different tropical and temperate regions. In these areas, steep changes in use were rare, but their average level of human appropriation of land by A.D.0 was already higher than levels observed today in *low-used* areas (Table S2.1).

In this trajectory-cluster trends in distribution of mammal diversity according to the three tested metrics show a latitudinal gradient, related to climatic factors (Figs. 3 & S4.14-16). In addition, total richness and proportion of threatened mammals tend to be higher at higher elevations

(Fig. 3), likely consequence of the higher human impact that eroded certain species from flatter areas –more favorable for human use– in the past (Laliberte and Ripple 2004). Furthermore, areas with more threatened mammals are also characterized by rapid, very low, or even negative pre-industrial changes (i.e. where human land use decreased; Fig.4). Here, we are likely observing two different processes; on the one hand, areas like Central and South America are naturally rich in mammal species, which make them more likely to harbor more threatened species even if they did not have a human-activity explosion in the pre-industrial times. On the other hand, in areas like Eastern Europe and African regions from Guinea to Chad, changes were pronounced and our prediction of species not having had time to recover from past human pressures may hold (Fig. S4.18). Nevertheless, it is important to note that, within these areas, environmental factors have the greatest relevance in explaining diversity, therefore, time-lagged effects must be cautiously interpreted (Table 2).

*Conclusions*

To our knowledge, the use of temporal trajectory-clusters' delineation has never been applied in biogeographical or ecological studies. We show the power of this method to synthesize complex information, in this case reflecting human land-use history at a global scale. Its application at finer scales might allow disentangling diversity patterns at local or regional scales, and exploring which specific prehistoric and historic events may be associated with those patterns.

From all the tested predictors, environmental factors were identified as the main drivers of total and threatened mammal richness distribution across the world. The importance of environmental predictors offers cautious good news in the current context of global change. Natural processes still appear to best explain current mammalian biodiversity distribution patterns. However, we should not overlook the fact that these natural processes also drive specific land uses, such as agriculture, which may not be fully captured with our general ‘present-land-use’ metric and that likely are also shaping contemporary mammals’ distribution (Licker et al. 2010). Interestingly, the second group of predictors in terms of importance was past human land use. In particular, the

rate of change during the period c.A.D.0-1750 repeatedly appears as relevant in tested models, highlighting the importance of this time-span. Understanding land-use drivers before industrialization may give us clues to explain current global biodiversity patterns (Redman 1999). The fact that past land-use metrics are more relevant in explaining total and especially threatened mammalian richness distributions than present land use also has implications for future studies given the widespread practice of including present land use, but not past use, as a predictor of numbers of threatened species (Lenzen et al. 2009, Koh and Ghazoul 2010, Brum et al. 2013). On the other hand, what we call current mammalian biodiversity patterns may be an optimistic picture of what actually remains, since rapid changes may not be captured by global biodiversity databases and, although our resolution of 1° is recommended at the global scale, IUCN range maps are known to overestimate species geographic ranges (Hurlbert and Jetz 2007). Furthermore, we represented land use with data from c.A.D.2000 to use the same source, but land-use changes have likely occurred in the last two decades. The difficulty of obtaining global updated and matching land-use and biodiversity data may add some noise to our analyses, and could partially explain why current land use has not relevance in explaining current patterns of mammal distribution.

Extinction debts have been described at local or landscape scales, but less frequently at wider geographical scales (Cowlishaw 1999, Helm et al. 2006, Bommarco et al. 2014). Here, we present evidence of global extinction debt, consistent with previous macroecological studies showing the greater relevance of past vs present land-use metrics, using countries as units of study and focusing on temperate regions, or using much simpler statistical methods for all vertebrate groups (Dullinger et al. 2013, Chen and Peng 2017). Our results point to the existence of a geographically widespread land-use legacy –more evident in *recently*-used areas– and warn us about possible extinction debts (Brooks et al. 1999). Extinction debts at regional scales have been found to reflect more recent habitat modifications and particularly affect sensitive species, agreeing with our global results (Uezu and Metzger 2016). Worryingly, if past changes explain today's patterns, it is possible that biodiversity responses to today's changes are not detectable yet, even if there are impacts (Jackson and Sax 2009).

We complement previous findings by including different biodiversity metrics, which improves the interpretation of land-use legacies; and providing a trajectory-cluster zonification, which not only summarizes global trends but also reveals three distinct land-use areas with specific conservation needs. *Low-used* areas present a good opportunity to proactively preserve mammal species before future impacts and before signs of emerging extinction-debt are confirmed. *Recently-used* areas, which have been rarely prioritized for conservation up to now, are likely most affected by extinction debt which may be best addressed via restoration initiatives. In *steadily-used* areas conservation actions should focus in tropical higher-diversity areas and those that were rapidly-modified-in-the-past areas where some susceptible species may remain. On the other hand, areas modified at a more constant pace throughout history (e.g. Europe or India) may have more resilient species, as these areas are more likely to have suffered more severe extinction filters in the past (Newbold et al. 2018, Polaina et al. 2018).

Overall, our study shows current biodiversity patterns reflect a signal of environmental drivers but also of past land use. Separating the land surface based on land-use history also warns us about overlooked processes in more anthropized regions, which encompass a relatively small area (*recently*- and *steadily-used* areas) affected by different events that may not be detected at the global scale. Human impacts have not completely altered natural biogeographical patterns, but we note with concern that in most areas of the world the peak of human land-use has occurred only recently while the transformation of primary areas continues. Deeper impacts on species biodiversity may have already occurred, but they may only be apparent to the conservation biologists of the future.

**References**

Bommarco, R. et al. 2014. Extinction debt for plants and flower-visiting insects in landscapes with contrasting land use history. - Divers. Distrib. 20: 591–599.

Bonn, A. et al. 2002. Threatened and endemic species: are they good indicators of patterns of biodiversity on a national scale? - Ecol. Lett. 5: 733–741.

Boserup, E. 1965. Population and technological change: a study of long term trends. - Chicago University Press.

Brooks, T. M. et al. 1999. Time Lag between Deforestation and Bird Extinction in Tropical Forest Fragments. - Conserv. Biol. 13: 1140–1150.

Brooks, T. M. et al. 2006. Global biodiversity conservation priorities. - Science (80-. ). 313: 58–61.

Brum, F. T. et al. 2013. Land Use Explains the Distribution of Threatened New World Amphibians Better than Climate. - PLoS One 8: 4–11.

Bürgi, M. et al. 2017. Legacy effects of human land use: ecosystems as time-lagged systems. - Ecosystems 20: 94–103.

Calinski, T. and Harabasz, J. 1972. A dendrite method for cluster analysis. - Commun. Stat. - Theory methods 3: 1–27.

Ceballos, G. and Ehrlich, P. R. 2006. Global mammal distributions, biodiversity hotspots, and conservation. - Proc. Natl. Acad. Sci. U. S. A. 103: 19374–19379.

Ceballos, G. et al. 2017. Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. - Proc. Natl. Acad. Sci. 114: E6089–E6096.

Celeux, G. and Govaert, G. 1992. A classification EM algorithm for clustering and two stochastic versions. - Comput. Stat. Data Anal. 14: 315–332.

Chen, Y. and Peng, S. 2017. Evidence and mapping of extinction debts for global forest-dwelling reptiles , amphibians and mammals. - Sci. Rep. 7: 44305.

Cowlishaw, G. 1999. Predicting the Pattern of Decline of African Primate Diversity: an Extinction Debt from Historical Deforestation. - Conserv. Biol. 13: 1183–1193.

Crase, B. et al. 2012. A new method for dealing with residual spatial autocorrelation in species distribution models. - Ecography (Cop.). 35: 879–888.

Crees, J. J. et al. 2016. Millennial-scale faunal record reveals differential resilience of European large mammals to human impacts across the Holocene. - Proceeding R. Scoiety London, Ser. B. 283: 20152152.

Davies, D. L. and Bouldin, D. W. 1979. A cluster separation measure. - IEEE Trans. Pattern Anal.

Mach. Intell. 1: 224–227.

Dullinger, S. et al. 2013. Europe ’ s other debt crisis caused by the long legacy of future extinctions. 110: 7342–7347.

Elith, J. et al. 2008. A working guide to boosted regression trees. - J. Anim. Ecol. 77: 802–813.

Ellis, E. C. 2011. Anthropogenic transformation of the terrestrial biosphere. - Phil. Trans. R. Soc. A 369: 1010–1035.

Ellis, E. C. et al. 2013. Used planet: a global history. - Proc. Natl. Acad. Sci. U. S. A. 110: 7978–85.

Faurby, S. and Svenning, J. C. 2015. Historic and prehistoric human-driven extinctions have reshaped global mammal diversity patterns. - Divers. Distrib. 21: 1155–1166.

Ferrier, S. and Watson, G. 1997. Evaluation of modelling techniques: I. Development of evaluation methodology. - In: An evaluation of the effectiveness of environmental surrogates and modelling techniques in predicting the distribution of biological diversity. Environment Australia, pp. 10.

Fick, S. and Hijmans, R. J. 2017. WorldClim2: new 1-km spatial resolution climate surfaces for global land areas. - Int. J. Climatol. 37: 4302–4315.

Foley, J. a et al. 2005. Global consequences of land use. - Science (80-. ). 309: 570–4.

Genolini, C. et al. 2015. kml and kml3d: R packages to cluster longitudinal data. - J. Stat. Softw. 65: 1–34.

Goldewijk, K. K. et al. 2010. The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12,000 years. - Glob. Ecol. Biogeogr. 20: 73–86.

Grenyer, R. et al. 2006. Global distribution and conservation of rare and threatened vertebrates. - Nature 444: 93–96.

Hawkins, B. A. et al. 2003. Energy, water and broad-scale geographic patterns of species richness. - Ecology 84: 3105–3117.

Helm, A. et al. 2006. Slow response of plant species richness to habitat loss and fragmentation. - Ecol. Lett. 9: 72–77.

Hijmans, R. J. et al. 2013. dismo: Species distribution modeling. R package version 0.9-3. in press.

Hillebrand, H. 2004. On the generality of the latitudinal diversity gradient. - Am. Nat. 163: 192–211.

Hoekstra, J. M. et al. 2005. Confronting a biome crisis: global disparities of habitat loss and protection. - Ecol. Lett. 8: 23–29.

Hurlbert, A. H. and Jetz, W. 2007. Species richness, hotspots, and the scale dependence of range maps in ecology and conservation. - Proc. Natl. Acad. Sci. U. S. A. 104: 13384–13389.

IUCN 2014. The IUCN Red List of Threatened Species. Version 2014.3. - Int. Union Conserv. Nat. in press.

Jackson, S. T. and Sax, D. F. 2009. Balancing biodiversity in a changing environment : extinction debt , immigration credit and species turnover. - Trends Ecol. Evol. 25: 153–160.

Kaplan, J. O. et al. 2009. The prehistoric and preindustrial deforestation of Europe. - Quat. Sci. Rev. 28: 3016–3034.

Kaplan, J. O. et al. 2011. Holocene carbon emissions as a result of anthropogenic land cover change. - The Holocene 21: 775–791.

Kehoe, L. et al. 2017. Biodiversity at risk under future cropland expansion and intensification. - Nat. Ecol. Evol. 1: 1129–1135.

Koh, L. P. and Ghazoul, J. 2010. A matrix-calibrated species-area model for predicting biodiversity losses due to land-use change. - Conserv. Biol. 24: 994–1001.

Kryszczuk, K. and Hurley, P. 2010. Estimation of the number of clusters using multiple clustering validity indices. - In: International Workshop on Multiple Classifier Systems. Springer, pp. 114–123.

Laliberte, A. S. and Ripple, W. J. 2004. Range Contractions of North American Carnivores and Ungulates. - Bioscience 54: 123–138.

Lenzen, M. et al. 2009. Effects of land use on threatened species. - Conserv. Biol. 23: 294–306.

Licker, R. et al. 2010. Mind the gap: how do climate and agricultural management explain the ‘yield gap’ of croplands around the world? - Glob. Ecol. Biogeogr. 19: 769–782.

Loehle, C. and Eschenbach, W. 2012. Historical bird and terrestrial mammal extinction rates and

causes. - Divers. Distrib. 18: 84–91.

LP DAAC 2004. Global 30 arc-second elevation data set GTOPO30. - L. Process Distrib. Act. Arch. Cent.

Martins, I. S. et al. 2014. The unusual suspect: Land use is a key predictor of biodiversity patterns in the Iberian Peninsula. - Acta Oecologica 61: 41–50.

Miehe, G. et al. 2014. How old is the human footprint in the world's largest alpine ecosystem? A review of multiproxy records from the Tibetan Plateau from the ecologists' viewpoint. - Quat. Sci. Rev. 86: 190–209.

Müller, D. et al. 2013. Comparing the determinants of cropland abandonment in Albania and Romania using boosted regression trees. - Agric. Syst. 117: 66–77.

Newbold, T. et al. 2018. Widespread winners and narrow-ranged losers: land use homogenizes biodiversity in local assemblages worldwide. - PLoS Biol.: 1–24.

Orme, C. D. L. et al. 2005. Global hotspots of species richness are not congruent with endemism or threat. - Nature 436: 1016–9.

Pekin, B. K. and Pijanowski, B. C. 2012. Global land use intensity and the endangerment status of mammal species (R Mac Nally, Ed.). - Divers. Distrib. 18: 909–918.

Polaina, E. et al. 2015. Socioeconomic correlates of global mammalian conservation status. - Ecosphere 6: 146.

Polaina, E. et al. 2018. From tropical shelters to temperate defaunation: The relationship between agricultural transition stage and the distribution of threatened mammals. - Glob. Ecol. Biogeogr. 27: 647–657.

Prescott, G. W. et al. 2012. Quantitative global analysis of the role of climate and people in explaining late Quaternary megafaunal extinctions. - Proc. Natl. Acad. Sci. U. S. A. 109: 4527–4531.

R Core Team 2015. R a language and environment for statistical computing. in press.

Ray, S. and Turi, R. H. 1999. Determination of number of clusters in k-means clustering and application in colour image segmentation. - Proc. 4th Int. Conf. Adv. pattern Recognit. Digit.

Tech.: 137–143.

Russell, G. J. et al. 2013. Present and Future Taxonomic Selectivity in Bird and Mammal Extinctions Present Mammal and Future Extinctions Taxonomic Selectivity in. 12: 1365–1376.

Sanderson, E. W. et al. 2002. The Human Footprint and the Last of the Wild. - Bioscience 52: 891–904.

Sandom, C. et al. 2014. Global late Quaternary megafauna extinctions linked to humans, not climate change. - Proc. R. Soc. B 281: 20133254.

Steffen, W. et al. 2015. Planetary Boundaries: Guiding human development on a changing planet. - Science (80-. ). 347: 1259855.

Terribile, L. C. et al. 2009. Global richness patterns of venomous snakes reveal contrasting infuences of ecology and history in two different clades. - Oecologia 159: 617–626.

Thuiller, W. et al. 2011. Consequences of climate change on the tree of life in Europe. - Nature 470: 531–534.

Torres-Romero, E. J. and Olalla-Tárraga, M. Á. 2014. Untangling human and environmental effects on geographical gradients of mammal species richness: a global and regional evaluation. - J. Anim. Ecol. 84: 851–860.

Turvey, S. T. and Fritz, S. a 2011. The ghosts of mammals past: biological and geographical patterns of global mammalian extinction across the Holocene. - Philos. Trans. R. Soc. Lond. B. Biol. Sci. 366: 2564–76.

Uezu, A. and Metzger, J. P. 2016. Time-lag in responses of birds to Atlantic forest fragmentation: Restoration opportunity and urgency. - PLoS One 11: 1–16.

Varela, S. et al. 2015. Differential effects of temperature change and human impact on European Late Quaternary mammalian extinctions. - Glob. Chang. Biol. 21: 1475–1481.

WWF 2018. Living Planet Report - 2018: Aiming higher (M Grooten and REA Almond, Eds.).

Zhang, K. et al. 2010. A continuous satellite - derived global record of land surface evapotranspiration from 1983 to 2006. - Water Resour. Res. 46: 1–21.

## Figures

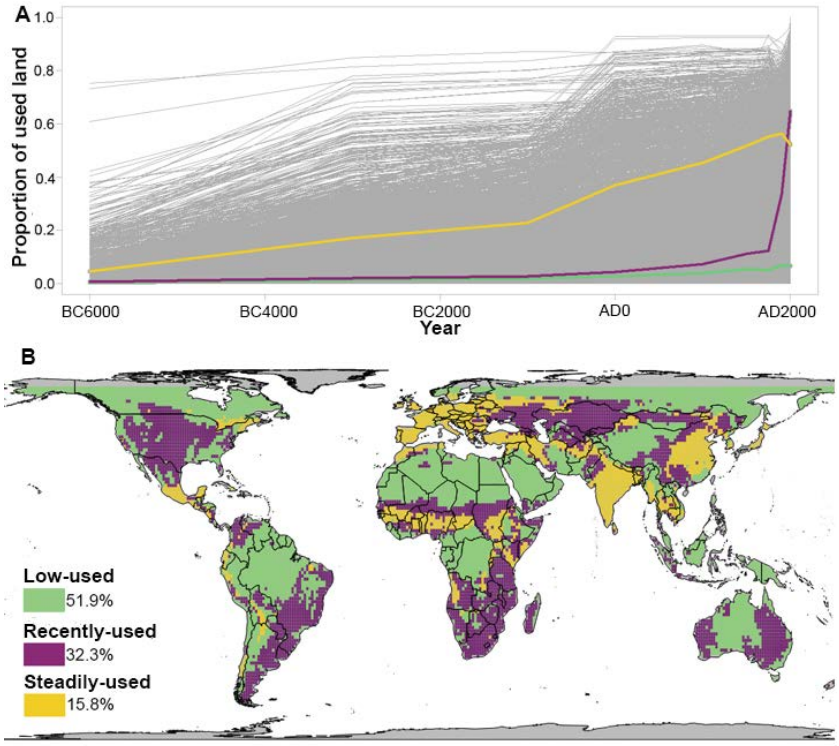



**Figure 1.** Overall trajectories (A) and spatial location (B) of trajectory-clusters from B.C.6000 to A.D.2000. Three was the optimal number of clusters according to the different quality criteria implemented in *kml* package (Genolini et al. 2015). In panel A, the X-axis represents the approximate year for which land use estimations are available in the KK10-model database, and Y-axis the proportion of land use (grid-cell proportion). Each grey line in panel A corresponds to a grid-cell on panel B. Legend (applicable for both panels) shows the proportion of grid cells assigned to each trajectory-cluster. Green color represents *low-used* areas*;* purple *recently-used* areas, and yellow *steadily-used* areas. Grey areas on the map show areas without historical land-use data. Map projection: Berhmann cylindrical equal area.

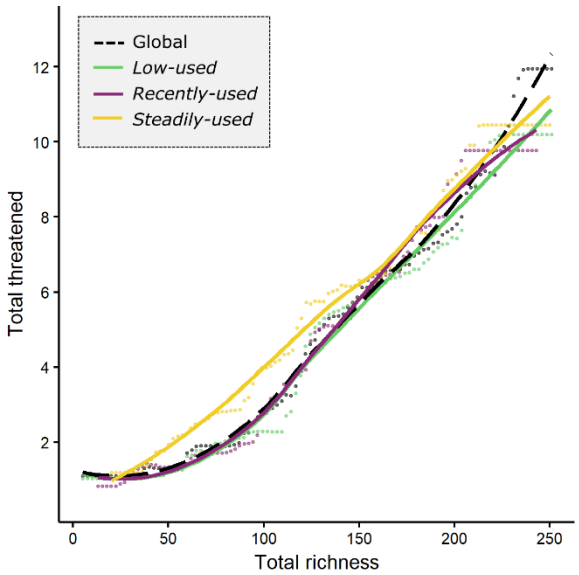


**Figure 2.** Partial dependency plot showing the relationship between total and threatened richness according to the global (dashed black line), *low-used* (green), *recently-used* (purple) and *steadily-used* (yellow) models. Point clouds show the actual response prediction when the rest of the explanatory variables are fixed to the mean; lines are LOESS smoothers to facilitate visualization.

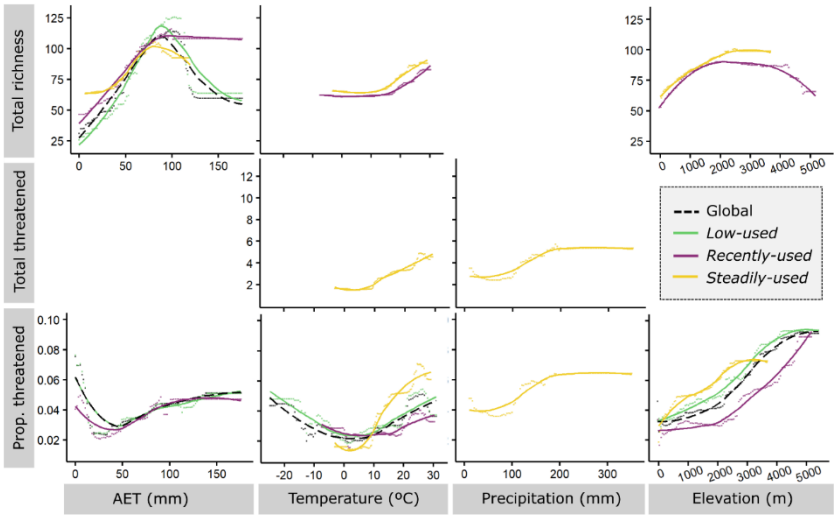



**Figure 3.** Partial dependency plots of the relationships between relevant environmental predictors of total mammalian richness (first row), total threatened richness (second row) and proportion of threatened mammals (third row). Dashed black lines symbolize results from the global model, continuous lines represent *trajectory-cluster* models' outcomes: *low-used* (green), *recently-used* (purple) and *steadily-used* (yellow). Point clouds show the actual response prediction when the rest of the explanatory variables are fixed to the mean; lines are LOESS smoothers to facilitate visualization.

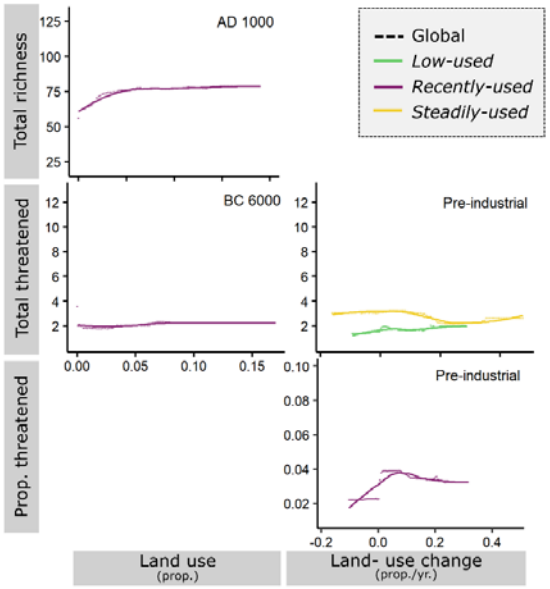



**Figure 4.** Partial dependency plots of the relationships between relevant land-use predictors of total mammalian richness (first row), total threatened richness (second row) and proportion of threatened mammals (third row). Dashed black lines symbolize results from the global model, continuous lines represent *trajectory-cluster* models' outcomes: *low-used* (green), *recently-used* (purple) and *steadily-used* (yellow). Point clouds show the actual response prediction when the rest of the explanatory variables are fixed to the mean; lines are LOESS smoothers to facilitate visualization.

## Tables

**Table 1.** Results of the boosted regression trees when total mammals richness was the response. Bold numbers denote relevant variables (% relevance greater than expected by chance, i.e. threshold = 100% divided by the number of variables included in each model). Significantly correlated variables ($\rho \geq |0.7|$) that did not enter any model are not included in this table. Dashes represent variables excluded only from some models due to high correlation or because they were not pertinent (e.g. including trajectory-cluster as a predictor within specific cluster models).

| | **Global** | **Low-used** | **Recently-used** | **Steadily-used** |
|---|---|---|---|---|
| **No. trees** | 5150 | 4700 | 5000 | 4700 |
| **Residuals' Moran's I** | 0.16*** | 0.09*** | 0.11*** | 0.03** |
| **% Deviance explained** | 96 | 96 | 91 | 89 |
| **Variables (relevance, %)** | | | | |
| RAC | **22.1** | **11.51** | **21.34** | **23.17** |
| Trajectory-cluster | 0.01 | - | - | - |
| *Environmental* | | | | |
| AET | **57.66** | **71.38** | **43.56** | **32.39** |
| Temperature | 5.92 | 4.13 | **10.86** | **8.69** |
| Precipitation | - | - | - | 7.61 |
| Elevation | 3.2 | 1.55 | **8.67** | **14.99** |
| *Land use* | | | | |
| $LU_{BC6000}$ | 1.04 | 0.67 | 1.26 | - |
| $LU_{AD1000}$ | - | - | **9.02** | 2.01 |
| $LU_{AD1750}$ | - | - | - | 2.91 |
| $LU_{AD1900}$ | - | - | 0.4 | - |
| $LU_{AD2000}$ | 0.82 | 0.43 | 0.53 | 0.63 |
| *Land-use change* | | | | |
| Prehistoric[1] | - | 1.18 | - | - |
| Pre-industrialization[2] | 6.25 | 5.68 | 3.55 | 4.96 |
| Post-industrialization[3] | 0.13 | 0.27 | - | - |
| *Remarkable land-use changes* | | | | |
| Time break | | | | |
| $LU_{+20}$ | 2.74 | 0.74 | 0.45 | - |
| $LU_{+50}$ | - | 2.02 | - | 1.52 |
| Land use | | | | |
| $LU_{+20}$ | - | - | 0.34 | 1.00 |
| Duration | | | | |
| $LU_{max}$ | 0.14 | 0.43 | 0.03 | 0.12 |

[1]BC6000-AD0; [2]AD0-1750; [3]AD1750-AD1950

**Table 2.** Results of the boosted regression trees when total threatened mammals richness was the response. Bold numbers denote relevant variables (% relevance greater than expected by chance, i.e. threshold = 100% divided by the number of variables included in each model). Significantly correlated variables ($\rho \geq |0.7|$) that did not enter any model are not included in this table. Dashes represent variables excluded only from some models due to high correlation or because they were not pertinent (e.g. including trajectory-cluster as a predictor within specific cluster models).

| | **Global** | **Low-used** | **Recently-used** | **Steadily-used** |
|---|---|---|---|---|
| **No. trees** | 10550 | 8600 | 10400 | 6700 |
| **Residuals' Moran's I** | 0.28*** | 0.23*** | 0.23*** | 0.19*** |
| **% Deviance explained** | 89 | 90 | 81 | 85 |
| **Variables (relevance, %)** | | | | |
| RAC | **27.94** | **14.99** | **20.34** | **8.3** |
| Trajectory-cluster | 0.06 | - | - | - |
| Total richness | **47.15** | **51.31** | **38.39** | **24.45** |
| *Environmental* | | | | |
| AET | - | - | 6.02 | 5.61 |
| Temperature | 3.28 | 1.66 | 3.68 | **9.02** |
| Precipitation | - | - | - | **22.12** |
| Elevation | 2.31 | 1.75 | 5.94 | 4.11 |
| *Land use* | | | | |
| $LU_{BC6000}$ | 3.6 | 1.4 | **9.38** | - |
| $LU_{AD1000}$ | - | - | 2.27 | 1.63 |
| $LU_{AD1750}$ | - | - | - | 0.94 |
| $LU_{AD1900}$ | - | - | 2.17 | - |
| $LU_{AD2000}$ | 1.82 | 2.16 | 2.91 | 3.35 |
| *Land-use change* | | | | |
| Prehistoric[1] | - | 5.3 | - | - |
| Pre-industrialization[2] | 5.61 | **10.39** | 5.47 | **12.01** |
| Post-industrialization[3] | 3.13 | 1.51 | - | - |
| *Remarkable land-use changes* | | | | |
| Time break | | | | |
| $LU_{+20}$ | 4.93 | 3.24 | 2.01 | - |
| $LU_{+50}$ | - | 6.26 | - | 7.29 |
| Land use | | | | |
| $LU_{+20}$ | - | - | 0.82 | 0.92 |
| Duration | | | | |
| $LU_{max}$ | 0.17 | 0.04 | 0.61 | 0.24 |

[1]BC6000-AD0; [2]AD0-1750; [3]AD1750-AD1950

**Table 3.** Results of the boosted regression trees when proportion of threatened mammals over total mammal richness was the response. Bold numbers denote relevant variables (% relevance greater than expected by chance, i.e. threshold = 100% divided by the number of variables included in each model). Significantly correlated variables ($\rho \geq |0.7|$) that did not enter any model are not included in this table. Dashes represent variables excluded only from some models due to high correlation or because they were not pertinent (e.g. including trajectory-cluster as a predictor within specific cluster models).

| | **Global** | **Low-used** | **Recently-used** | **Steadily-used** |
|---|---|---|---|---|
| **No. trees** | 9400 | 5000 | 6800 | 3800 |
| **Residuals' Moran's I** | 0.03*** | 0.11*** | -0.02 | -0.03 |
| **% Deviance explained** | 84 | 84 | 77 | 82 |
| **Variables (relevance, %)** | | | | |
| RAC | **45.44** | **29.19** | **33.79** | **18.14** |
| Trajectory-cluster | 0.07 | - | - | - |
| *Environmental* | | | | |
| AET | **14.54** | **23.01** | **10.93** | 6.04 |
| Temperature | **16.3** | **17.52** | **13.46** | **30.53** |
| Precipitation | - | - | - | **11.15** |
| Elevation | **9.84** | **9.08** | **11.91** | **17.11** |
| *Land use* | | | | |
| $LU_{BC6000}$ | 2.93 | 3.85 | 5.73 | - |
| $LU_{AD1000}$ | - | - | 6.19 | 2.13 |
| $LU_{AD1750}$ | - | - | - | 0.82 |
| $LU_{AD1900}$ | - | - | 1.84 | - |
| $LU_{AD2000}$ | 1.04 | 1.02 | 2.87 | 1.1 |
| *Land-use change* | | | | |
| Prehistoric[1] | - | 4.6 | - | - |
| Pre-industrialization[2] | 3.8 | 4.38 | **10.64** | 7.71 |
| Post-industrialization[3] | 3.03 | 1.7 | - | - |
| *Remarkable land-use changes* | | | | |
| Time break | | | | |
| $LU_{+20}$ | 2.64 | 1.26 | 1.15 | - |
| $LU_{+50}$ | - | 3.8 | - | 4.14 |
| Land use | | | | |
| $LU_{+20}$ | - | - | 1.28 | 0.94 |
| Duration | | | | |
| $LU_{max}$ | 0.36 | 0.57 | 0.2 | 0.18 |

[1]BC6000-AD0; [2]AD0-1750; [3]AD1750-AD1950

## SUPPLEMENTARY MATERIAL

## APPENDIX S1. Additional results of the trajectory-cluster analyses

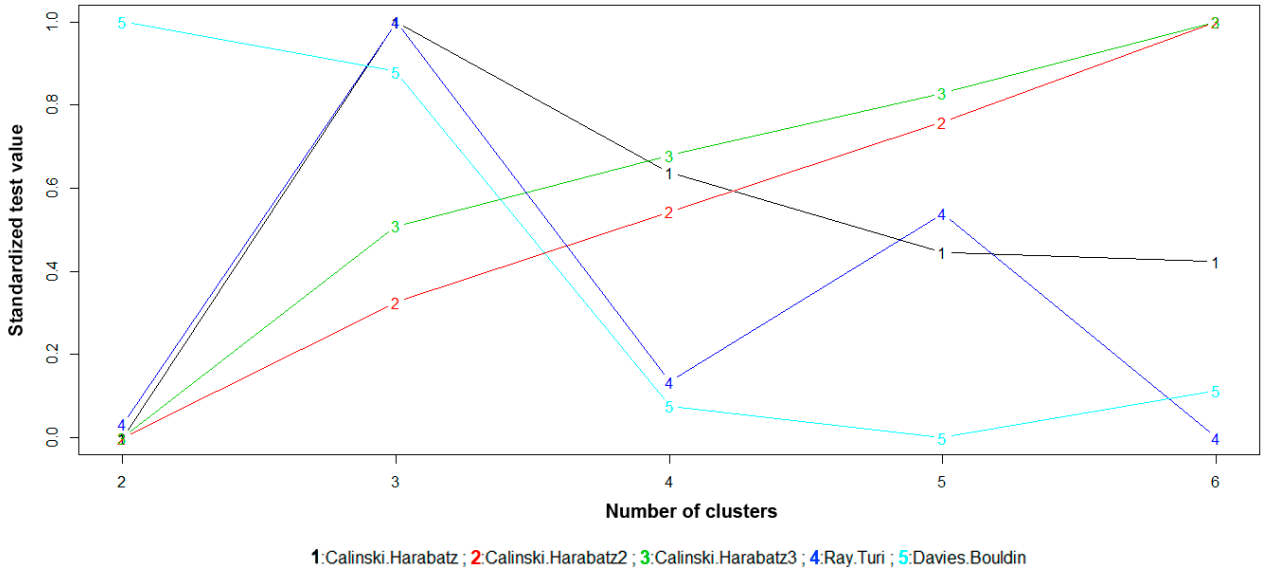
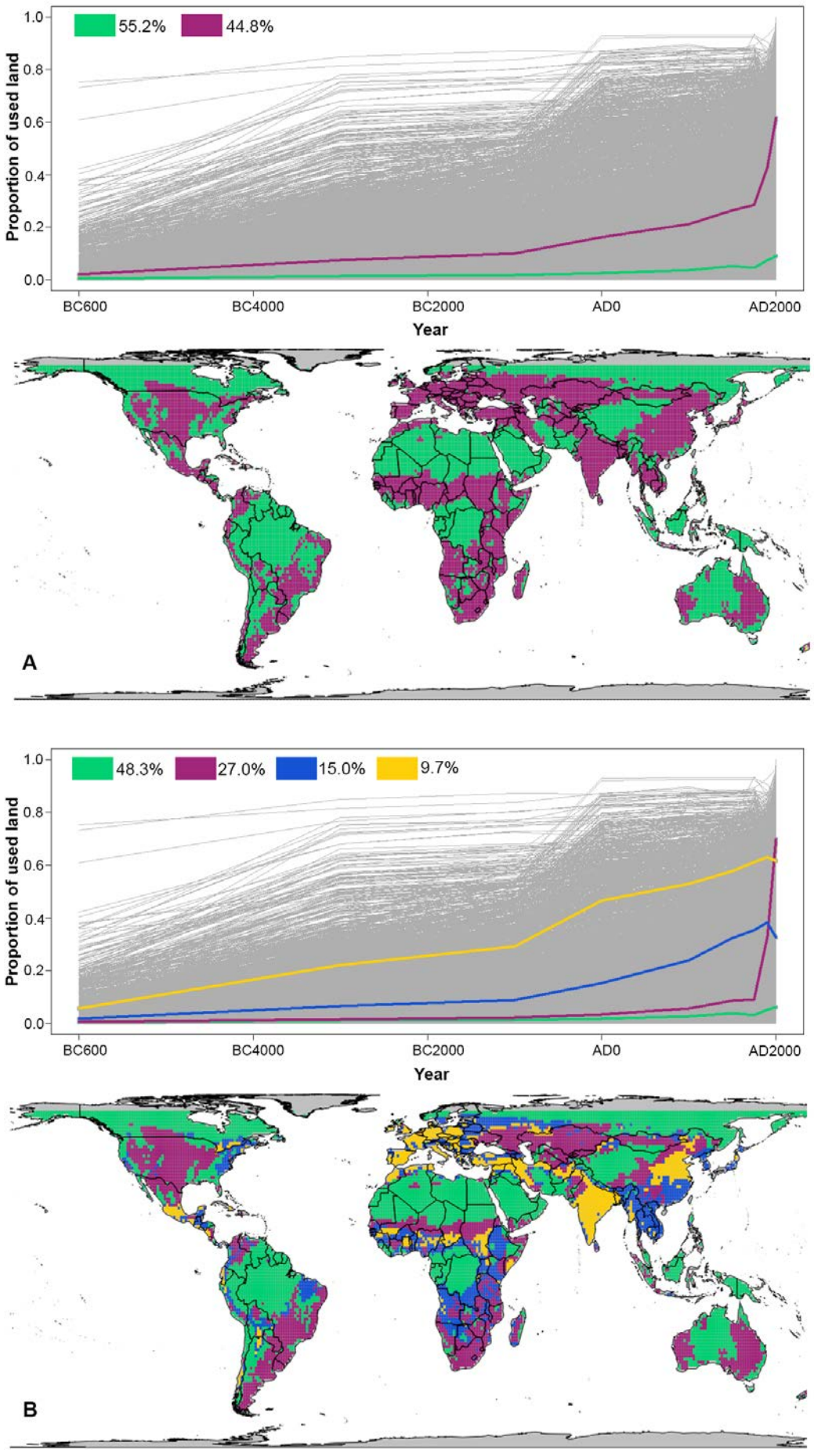



**Figure S1.1.** Optimal number of trajectory clusters according to different criteria. X-axis represent the number of proposed clusters, Y-axis shows the standardized [0-1] test values (the higher, the better). Different numbers and colors represent different quality tests, 1(black): Calinski-Harabatz (Calinski & Harabasz 1972); 2 (red): Calinski-Harabatz2, Kryszczuk variant (Kryszczuk & Hurley 2010); 3 (green): Calinski-Harabatz3, Genolini variant; 4 (blue): Ray-Turi (Ray & Turi 1999); 5 (cyan): Davies-Bouldin (Davies & Bouldin 1979).


55.2%
44.8%
Proportion of used land
1.0
0.8
0.6
0.4
0.2
0.0
BC600
BC4000
BC2000
AD0
AD2000
Year
A
48.3%
27.0%
15.0%
9.7%
Proportion of used land
1.0
0.8
0.6
0.4
0.2
0.0
BC600
BC4000
BC2000
AD0
AD2000
Year
B

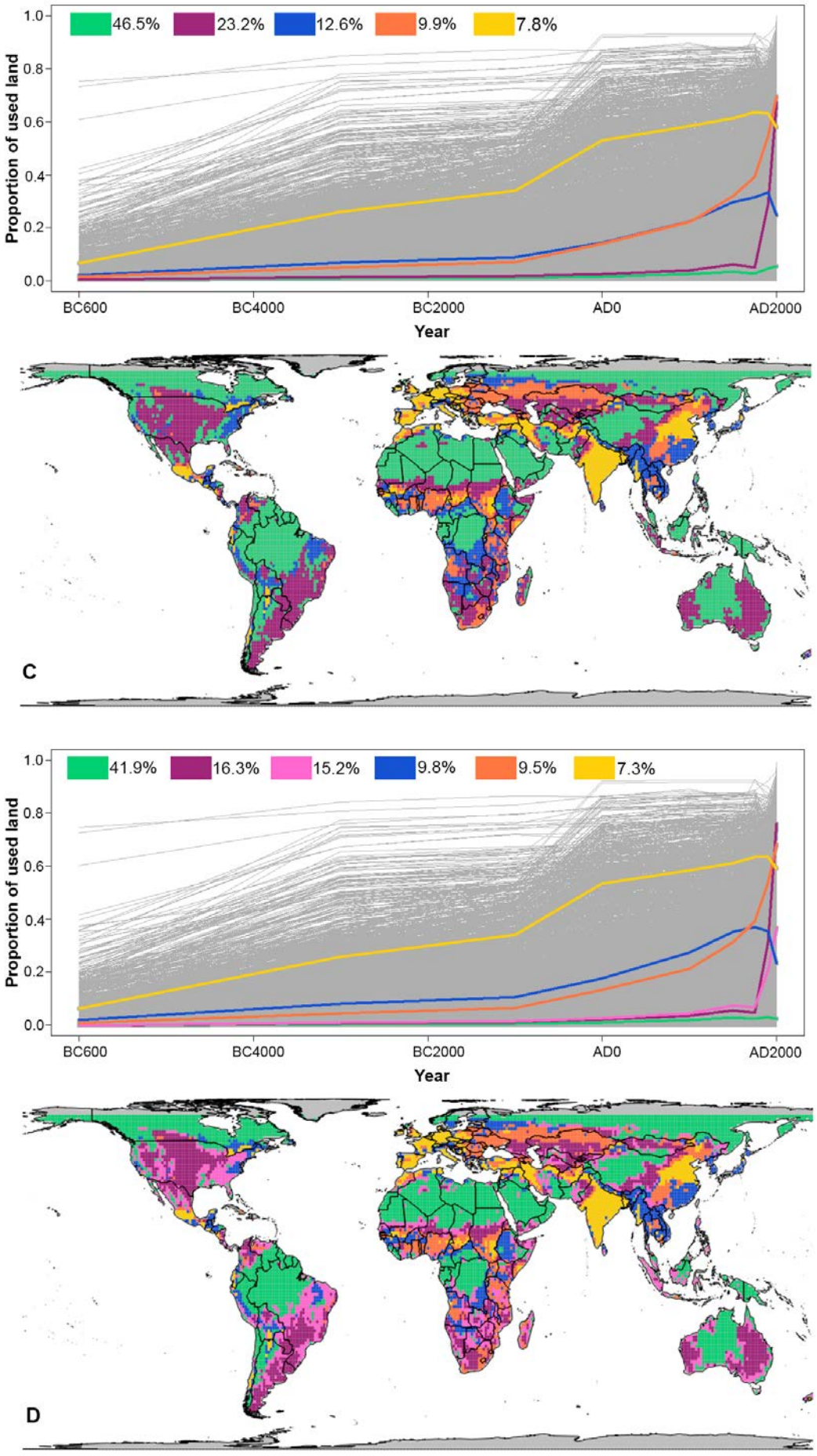



**Figure S1.2.** Overall trajectories and spatial location of land-use trajectories from B.C.6000 to A.D.2000, considering two (A), four (B), five (C) and six (D) cluster divisions. In all plots, X axis represent the

approximate year for which land use estimations are available in the KK10-model database, and Y axis the proportion of land use (grid-cell proportion). Legend shows the percentage of grid cells assigned to each trajectory-cluster. Map projection: Berhmann cylindrical equal area.

## APPENDIX S2. Data description: global and by trajectory-cluster

**Table S2.1.** Global and per-trajectory summary of land-use, environmental and mammalian diversity indicators calculated for a 110 x 100km grid (~1º at the Equator). Past land-use indicators are based on the KK10 model (Kaplan et al. 2011; spatial data available at <http://ecotope.org/anthromes/data/>). *Land use* refers to the mean proportion of grid cell classified as intended for human use at each time break (original temporal resolution); *land -use change* was calculated as the difference of proportion of used grid cell between temporal milestones, standardized per 1000-years time; *remarkable land uses* considered are: 20% grid cell intended for human use ($LU_{+20}$), defined as the first significant use by Ellis et al. (2013); 50% grid cell intended for human use ($LU_{+50}$), taken as an arbitrarily high value; and maximum value of use per grid cell for the whole time series ($LU_{max}$). For each threshold, we show: *time break* (when the threshold was exceeded), *land use* (how much land was used by humans when the threshold was exceeded) and *duration* (how long this value was maintained). Environmental indicators are recalculated from the original data (see Table S2.2) as the mean per grid-cell. Total and threatened mammalian richness are based on IUCN Red List Data (IUCN 2014).

| | **Global** (N=9867) | **Low-used** (N=5119) | **Recently-used** (N=3189) | **Steadily-used** (N=1559) |
|---|---|---|---|---|
| **Environmental [median, min-max]** | | | | |
| Temperature | 18.14 (-24.90-30.92) | 18.78 (-24.90-30.92) | 18.71 (-7.54-30.32) | 16.12 (-3.15-29.14) |
| AET | 30.49 (0-175.71) | 22.32 (0-175.38) | 36.02 (0-175.71) | 39.69 (6.41-119.63) |
| Precipitation | 44.61 (0-447.26) | 37.81 (0-447.26) | 42.30 (1.35-303.66) | 66.22 (10.43-351.06) |
| Elevation | 430.63 (-39.91-5414.87) | 410.86 (-39.91-5414.87) | 510.42 (-27.29-5173.28) | 397.43 (2.09-3677.44) |
| **Land use [median, min-max (grid-cell proportion)]** | | | | |
| BC6000 | 0.004 (0-0.754) | 0.001 (0-0.181) | 0.004 (0-0.169) | 0.024 (0.002-0.754) |
| BC3000 | 0.008 (0-0.848) | 0.004 (0-0.332) | 0.009 (0-0.230) | 0.138 (0.003-0.848) |
| BC1000 | 0.012 (0-0.871) | 0.005 (0-0337) | 0.012 (0-0.278) | 0.204 (0.003-0.871) |
| AD0 | 0.019 (0-0.928) | 0.009 (0-0.380) | 0.021 (0-0.323) | 0.347 (0.016-0.928) |
| AD1000 | 0.034 (0-0.930) | 0.013 (0-0.392) | 0.046 (0-0.380) | 0.435 (0.102-0.930) |
| AD1500 | 0.055 (0-0.931) | 0.018 (0-0.442) | 0.082 (0-0.609) | 0.494 (0.166-0.931) |
| AD1750 | 0.034 (0-0.936) | 0.011 (0-0.520) | 0.057 (0-0.559) | 0.547 (0.108-0.936) |
| AD1900 | 0.197 (0-0.928) | 0.030 (0-0.491) | 0.314 (0.141-0.783) | 0.560 (0.203-0.928) |
| AD1950 | 0.247 (0-0.946) | 0.023 (0-0.332) | 0.490 (0.271-0.853) | 0.556 (0.144-0.946) |
| AD2000 | 0.250 (0-0.999) | 0.005 (0-0.404) | 0.652 (0.166-0.999) | 0.538 (0.003-0.979) |
| **Land-use change [median, min-max (grid-cell proportion/1000 yrs)]** | | | | |
| Prehistoric[1] | 0.002 (0-0.143) | 0.001 (0-0.062) | 0.003 (0-0.052) | 0.050 (0.002-0.143) |
| Pre-industrialization[2] | 0.003 (-0.160-0.476) | 0.001 (-0.0944-0.287) | 0.008 (-0.103-0.316) | 0.081 (-0.160-0.476) |
| Industrialization[3] | 0.192 (-2.781-4.095) | -0.003 (-1.442-1.479) | 1.908 (-0.748-4.095) | -0.056 (-2.781-2.985) |
| Post-industrializtion[4] | 0.073 (-5.564-6.544) | -0.033 (-4.222-2.669) | 3.277 (-3.454-6.544) | -0.389 (-5.564-5.851) |
| **Remarkable land-use changes** | | | | |
| Time break [highest number of grid-cells transformed (yr)] | | | | |
| $LU_{+20}$ | AD1900 | AD1950 | AD1900 | BC3000 |
| $LU_{+50}$ | AD1950 | AD1950 | AD1950 | AD0 |
| $LU_{max}$ | AD2000 | AD2000 | AD2000 | AD1750 |
| Land use [median, min-max (grid-cell proportion)] | | | | |
| $LU_{+20}$ | 0.220 (0-0.871) | 0 (0-0.477) | 0.283 (0.2-0.783) | 0.286 (0.200-0.871) |
| $LU_{+50}$ | 0.253 (0-0.798) | 0.048 (0-0.477) | 0.436 (0.021-0.783) | 0.357 (0.018-0.798) |
| $LU_{max}$ | 0.346 (0-0.999) | 0.054 (0-0.520) | 0.652 (0.278-0.999) | 0.642 (0.294-0.979) |

| | **Global** (N=9867) | **Low-used** (N=5119) | **Recently-used** (N=3189) | **Steadily-used** (N=1559) |
|---|---|---|---|---|
| Duration [median, min-max (yrs)] | | | | |
| $LU_{+20}$ | 250 (0-8000) | 50 (0-5000) | 100 (0-5000) | 2750 (400-8000) |
| $LU_{+50}$ | 250 (0-8000) | 900 (0-8000) | 50 (0-5000) | 2000 (50-8000) |
| $LU_{max}$ | 0 (0-8000) | 0 (0-8000) | 0 (0-250) | 100 (0-2000) |
| **Mammalian diversity [median, min-max]** | | | | |
| Species richness | 55 (5-251) | 41 (5-251) | 63 (13-243) | 68 (20-250) |
| Threatened spp richness | 2 (0-40) | 1 (0-40) | 2 (0-33) | 3 (0-32) |

[1]BC6000-AD0; [2]AD0-1750; [3]AD1750-AD1950; [4]AD1950-2000.

**Table S2.2.** Data sources of environmental indicators included in the statistical models. All variables were resampled to our grid size (~110 x 110km ; 1º at the equator).

| **Environmental indicators** | | **Units, description** | **Year** | **Original resolution** | | **Reference** |
|---|---|---|---|---|---|---|
| *Long name* | *Short name* | | | *Spatial* | *Temporal* | |
| Mean annual actual evapotranspiration | AET | mm, accumulated | 2000 | 1 degree | month | Zhang et al. (2010, 2015) |
| Mean annual temperature | Temperature | ºC, average | 1970-2000 | 10 arc minutes | month | Fick & Hijmans (2017) |
| Mean annual precipitation | Precipitation | mm, average | 1970-2000 | 10 arc minutes | month | Fick & Hijmans (2017) |
| Global digital elevation model | Elevation | m | 1996 | 30 arc seconds | - | LP DAAC (2004) |

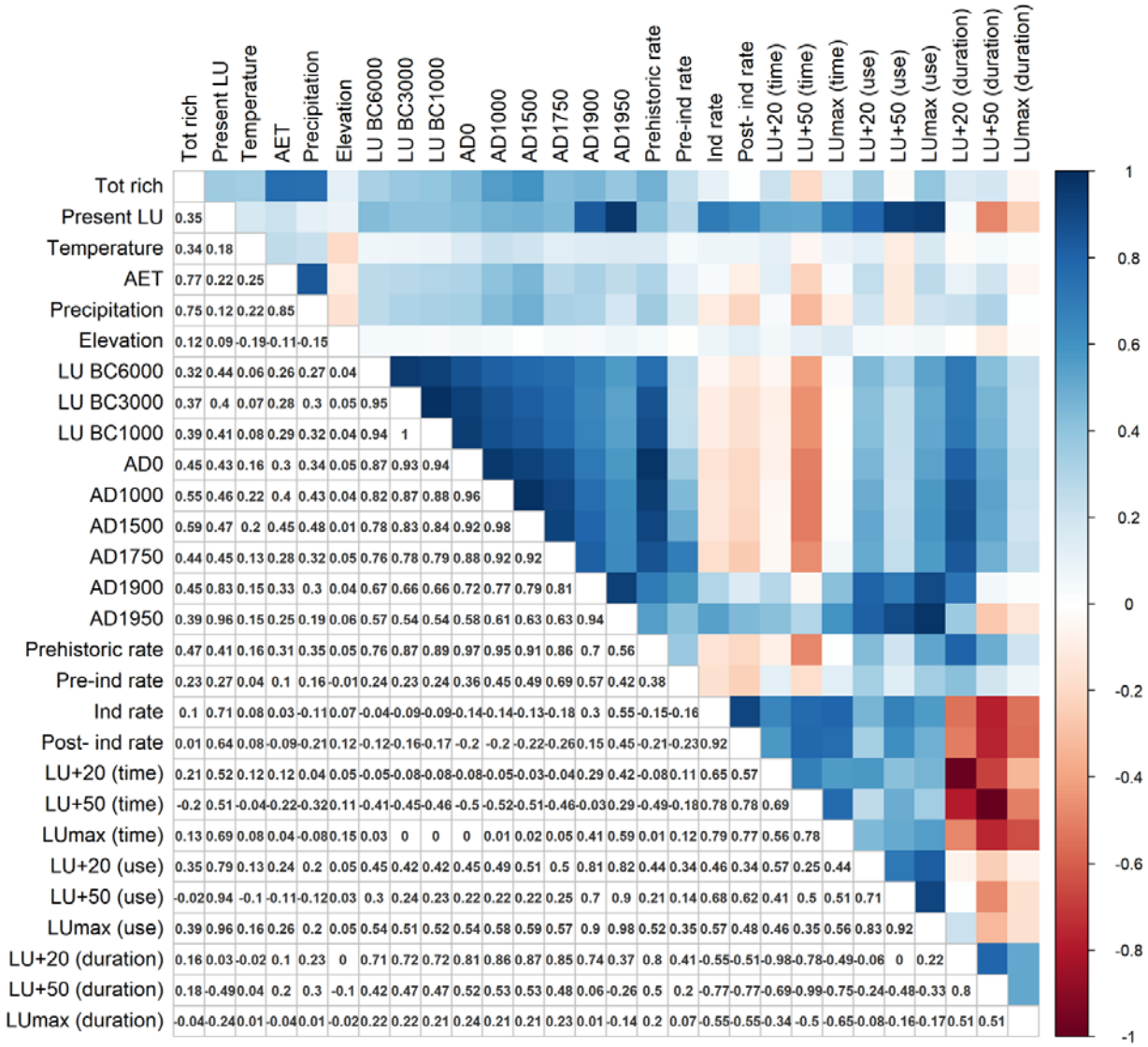



**Figure S2.1.** Spearman's rank correlation coefficient (ρ) calculated for all variables *a priori* considered for the global model. Blue colours show high positive correlation values; red colours show low positive correlation values.

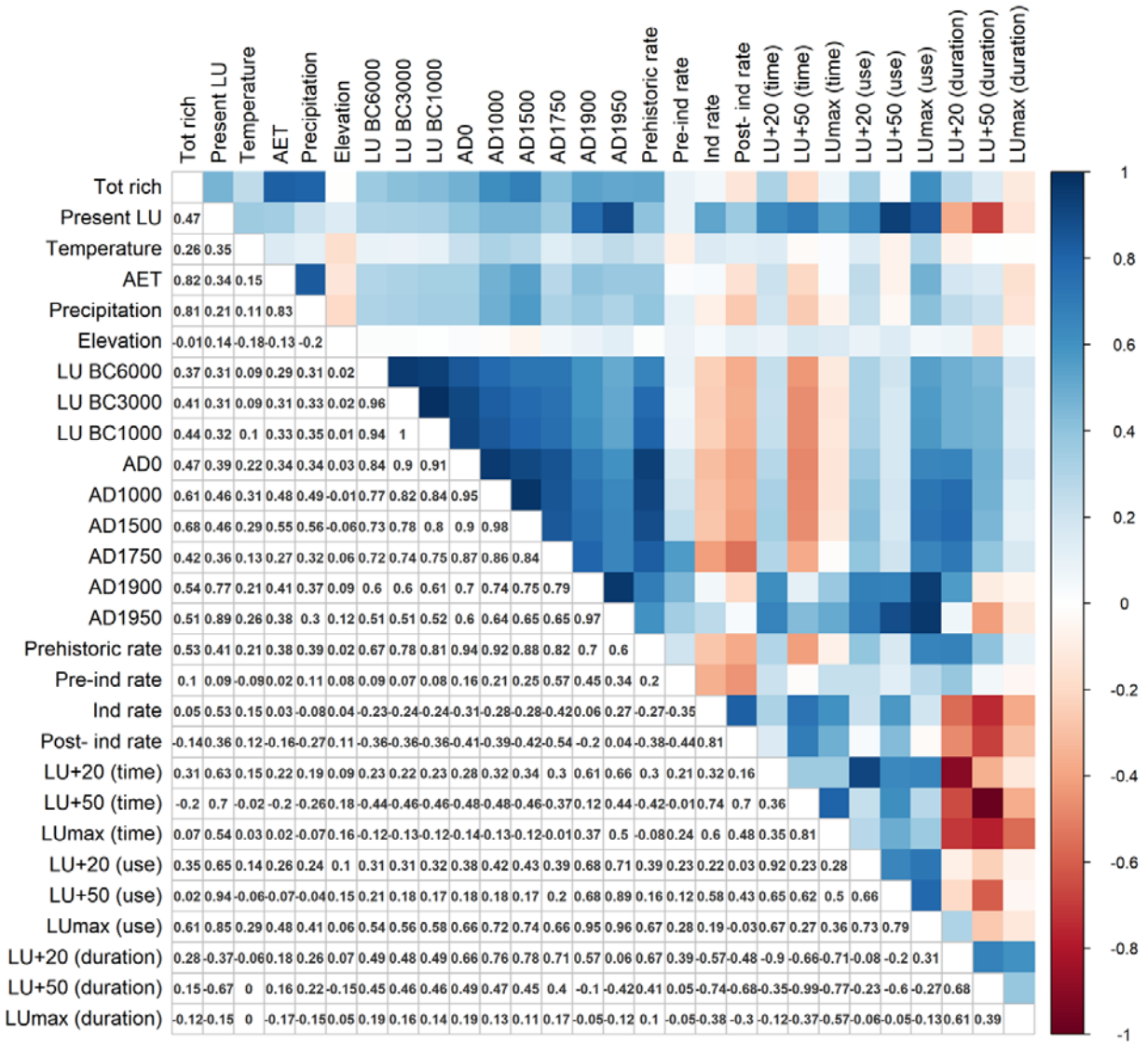



**Figure S2.2.** Spearman's rank correlation coefficient (ρ) calculated for all variables *a priori* considered for the *low-used*-areas model. Blue colours show high positive correlation values; red colours show low positive correlation values.

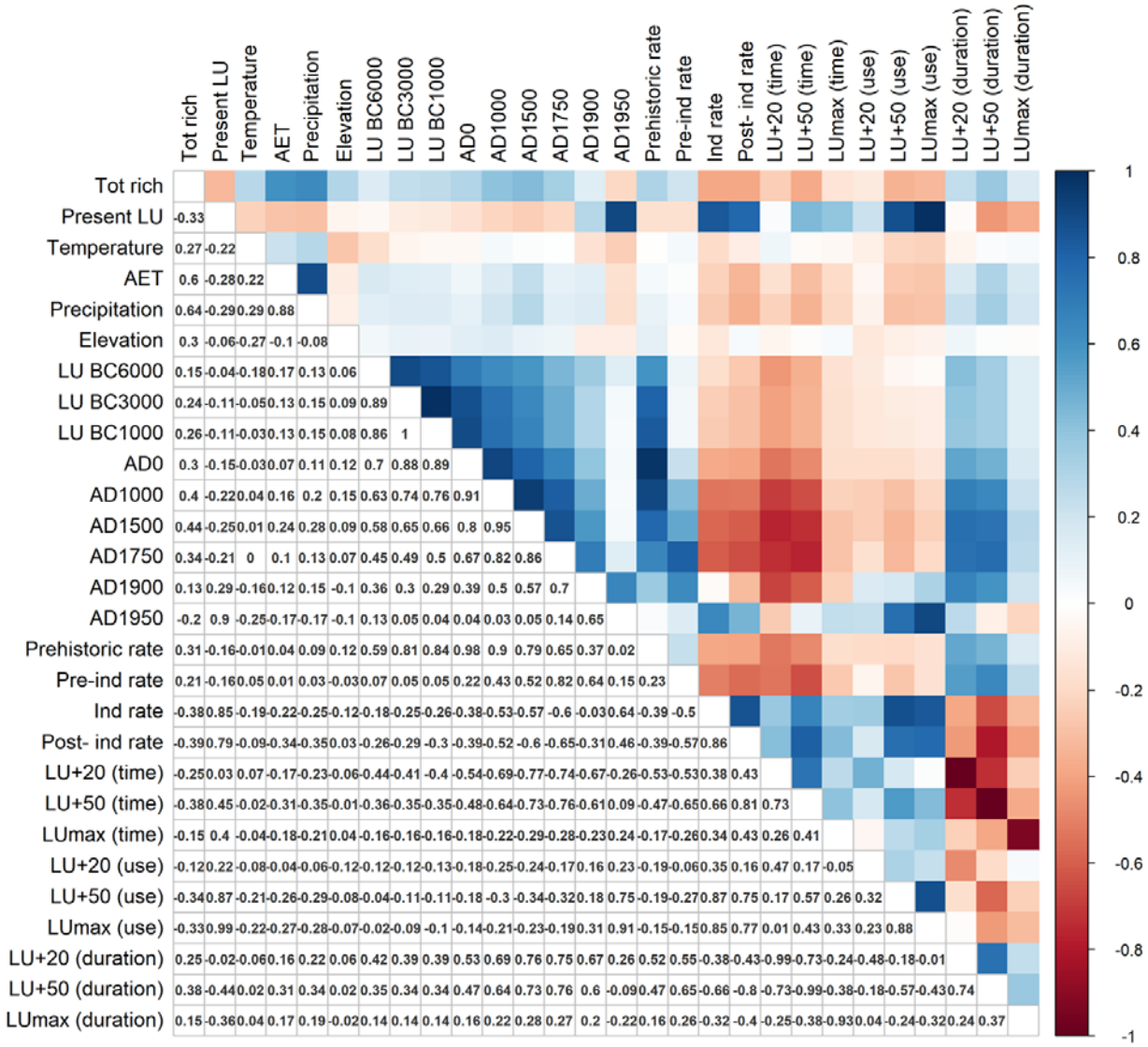



**Figure S2.3.** Spearman's rank correlation coefficient (ρ) calculated for all variables *a priori* considered for the *recently-used*-areas model. Blue colours show high positive correlation values; red colours show low positive correlation values.

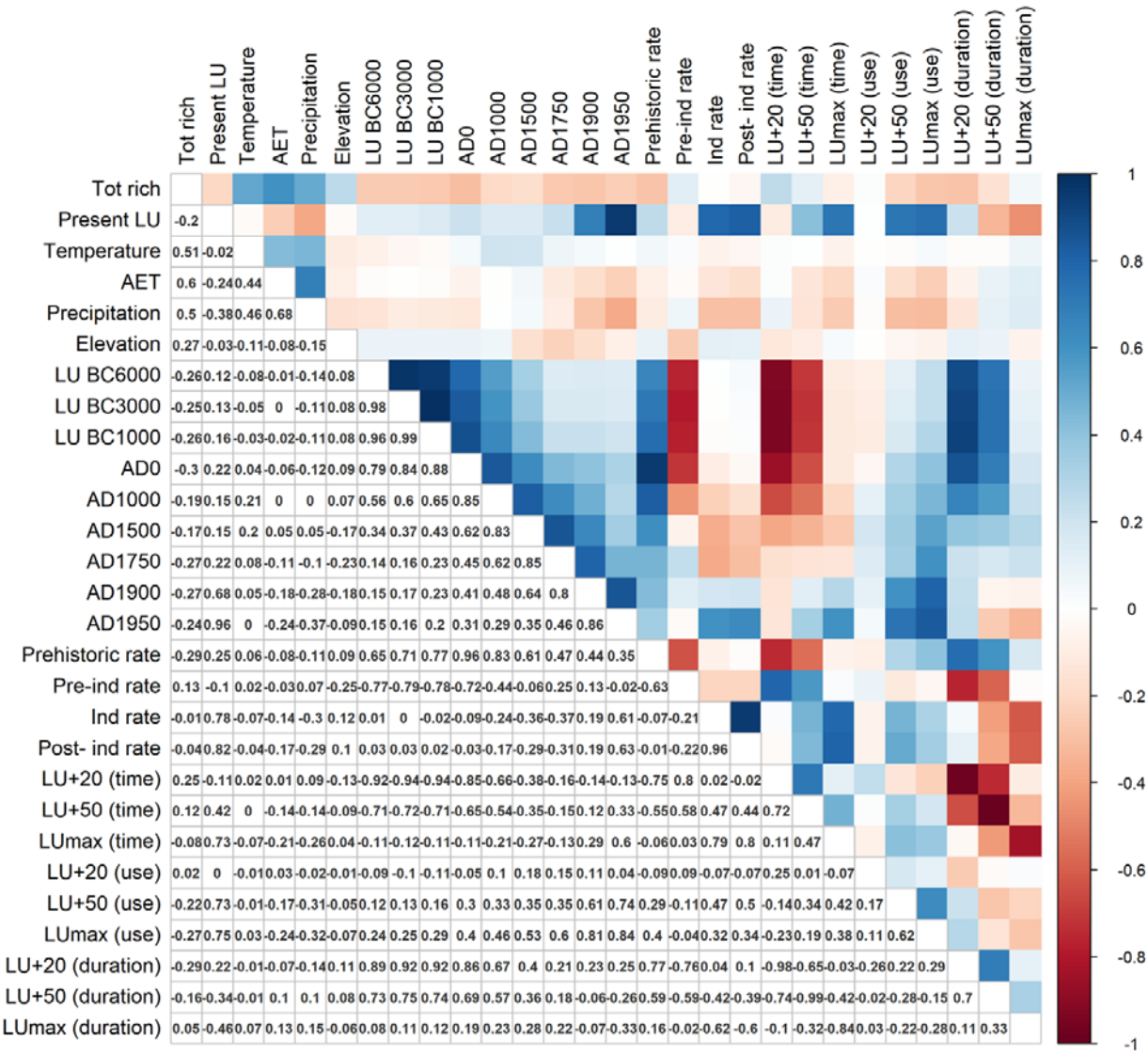



**Figure S2.4.** Spearman's rank correlation coefficient (ρ) calculated for all variables *a priori* considered for the *steadily-used*-areas model. Blue colours show high positive correlation values; red colours show low positive correlation values.

## APPENDIX S3. Additional results of the BRT Models

**Table S3.1.** Results of the boosted regression trees when total mammals richness was the response, excluding residuals autocovariate (RAC). Bold numbers denote relevant variables (% relevance greater than expected by chance, i.e. threshold = 100% divided by the number of variables included in each model). Significantly correlated variables (ρ≥ |0.7|) that did not enter any model are not included in this table. Dashes represent variables excluded only from some models due to high correlation or because they were not pertinent (e.g. including trajectory-cluster as a predictor within specific cluster models).

| | **Global** | **Low-used** | **Recently-used** | **Steadily-used** |
|---|---|---|---|---|
| **No. trees** | 7300 | 5650 | 9400 | 5700 |
| **Residuals' Moran's I** | 0.62*** | 0.51*** | 0.53*** | 0.42*** |
| **% Deviance explained** | 89 | 93 | 83 | 81 |
| **Variables (relevance, %)** | | | | |
| Trajectory-cluster | 0.04 | - | - | - |
| *Environmental* | | | | |
| AET | **62.63** | **73.75** | **45.36** | **36.25** |
| Temperature | **10.54** | 5.63 | **14.47** | **10.61** |
| Precipitation | - | - | | **10.91** |
| Elevation | 6.75 | 3.02 | **11.50** | **18.06** |
| *Land use* | | | | |
| $LU_{BC6000}$ | 2.69 | 1.77 | 3.79 | - |
| $LU_{AD1000}$ | - | - | **10.36** | 3.71 |
| $LU_{AD1750}$ | - | - | - | 4.08 |
| $LU_{AD1900}$ | - | - | 1.55 | - |
| $LU_{AD2000}$ | 1.87 | 0.85 | 1.86 | 1.16 |
| *Land-use change* | | | | |
| Prehistoric[1] | - | 1.91 | - | - |
| Pre-industrialization[2] | 8.93 | **7.87** | **8.24** | 9.98 |
| Post-industrialization[3] | 0.90 | 0.99 | - | - |
| *Remarkable land-use changes* | | | | |
| Time break | | | | |
| $LU_{+20}$ | 5.19 | 1.12 | 1.36 | - |
| $LU_{+50}$ | - | 2.54 | - | 2.56 |
| Land use | | | | |
| $LU_{+20}$ | - | - | 1.42 | 2.42 |
| Duration | | | | |
| $LU_{max}$ | 0.46 | 0.55 | 0.09 | 0.28 |

[1]BC6000-AD0; [2]AD0-1750; [3]AD1750-AD1950

**Table S3.2.** Results of the boosted regression trees when total threatened mammals richness was the response, excluding residuals autocovariate (RAC). Bold numbers denote relevant variables (% relevance greater than expected by chance, i.e. threshold = 100% divided by the number of variables included in each model). Significantly correlated variables (ρ≥ |0.7|) that

did not enter any model are not included in this table. Dashes represent variables excluded only from some models due to high correlation or because they were not pertinent (e.g. including trajectory-cluster as a predictor within specific cluster models).

| | **Global** | **Low-used** | **Recently-used** | **Steadily-used** |
|---|---|---|---|---|
| **No. trees** | 14000 | 10950 | 13550 | 7500 |
| **Residuals' Moran's I** | 0.63*** | 0.61*** | 0.53*** | 0.53*** |
| **% Deviance explained** | 78 | 84 | 76 | 81 |
| **Variables (relevance, %)** | | | | |
| Trajectory-cluster | 0.23 | - | - | - |
| Total richness | **51.63** | **52.50** | **40.68** | **25.07** |
| *Environmental* | | | | |
| AET | - | - | 7.75 | 5.38 |
| Temperature | 5.71 | 3.33 | 6.03 | **10.72** |
| Precipitation | - | - | - | **23.47** |
| Elevation | 4.72 | 2.94 | 8.61 | 4.41 |
| *Land use* | | | | |
| $LU_{BC6000}$ | 6.37 | 2.14 | **11.26** | - |
| $LU_{AD1000}$ | - | - | 3.96 | 1.89 |
| $LU_{AD1750}$ | - | - | - | 2.06 |
| $LU_{AD1900}$ | - | - | 3.46 | - |
| $LU_{AD2000}$ | 4.86 | 2.69 | 4.29 | 3.57 |
| *Land-use change* | | | | |
| Prehistoric[1] | - | 7.21 | - | - |
| Pre-industrialization[2] | **10.19** | **15.94** | **8.80** | **13.76** |
| Post-industrialization[3] | 6.83 | 2.24 | - | - |
| *Remarkable land-use changes* | | | | |
| Time break | | | | |
| $LU_{+20}$ | 9.11 | 2.89 | 2.74 | |
| $LU_{+50}$ | - | **7.96** | | **7.68** |
| Land use | | | | |
| $LU_{+20}$ | - | - | 1.83 | 1.51 |
| Duration | | | | |
| $LU_{max}$ | 0.36 | 0.16 | 0.58 | 0.48 |

[1]BC6000-AD0; [2]AD0-1750; [3]AD1750-AD1950

**Table S3.3.** Results of the boosted regression trees when proportion of threatened mammals over total mammal richness was the response, excluding residuals autocovariate (RAC). Bold numbers denote relevant variables (% relevance greater than expected by chance, i.e. threshold = 100% divided by the number of variables included in each model). Significantly correlated variables ($\rho \geq |0.7|$) that did not enter any model are not included in this table. Dashes represent variables excluded only from some models due to high correlation or because they were not pertinent (e.g. including trajectory-cluster as a predictor within specific cluster models).

| | Global | Low-used | Recently-used | Steadily-used |
|---|---|---|---|---|
| **No. trees** | 15000 | 14100 | 12300 | 6250 |
| **Residuals' Moran's I** | 0.50*** | 0.42*** | 0.34*** | 0.33*** |
| **% Deviance explained** | 68 | 74 | 65 | 75 |
| **Variables (relevance, %)** | | | | |
| Trajectory-cluster | 0.22 | - | - | - |
| *Environmental* | | | | |
| AET | **20.18** | 23.52 | **14.29** | **10.00** |
| Temperature | **24.01** | 23.93 | **19.48** | **33.20** |
| Precipitation | - | - | - | **14.45** |
| Elevation | **17.21** | 15.34 | **14.88** | **19.16** |
| *Land use* | | | | |
| $LU_{BC6000}$ | 8.14 | 6.42 | 8.67 | - |
| $LU_{AD1000}$ | - | - | **9.77** | 2.93 |
| $LU_{AD1750}$ | - | - | - | 2.10 |
| $LU_{AD1900}$ | - | - | 5.35 | - |
| $LU_{AD2000}$ | 3.49 | 1.91 | 6.22 | 2.28 |
| *Land-use change* | | | | |
| Prehistoric[1] | - | 7.83 | - | - |
| Pre-industrialization[2] | **10.91** | 8.77 | **14.61** | 8.94 |
| Post-industrialization[3] | 7.33 | 3.80 | - | - |
| *Remarkable land-use changes* | | | | |
| Time break | | | | |
| $LU_{+20}$ | 7.37 | 2.01 | 2.97 | - |
| $LU_{+50}$ | - | 5.37 | - | 4.60 |
| Land use | | | | |
| $LU_{+20}$ | - | - | 3.44 | 1.85 |
| Duration | | | | |
| $LU_{max}$ | 1.16 | 1.09 | 0.32 | 0.47 |

[1]BC6000-AD0; [2]AD0-1750; [3]AD1750-AD1950

## APPENDIX S4. Mapping of relevant variables for each trajectory-cluster model

### *Low-used areas*

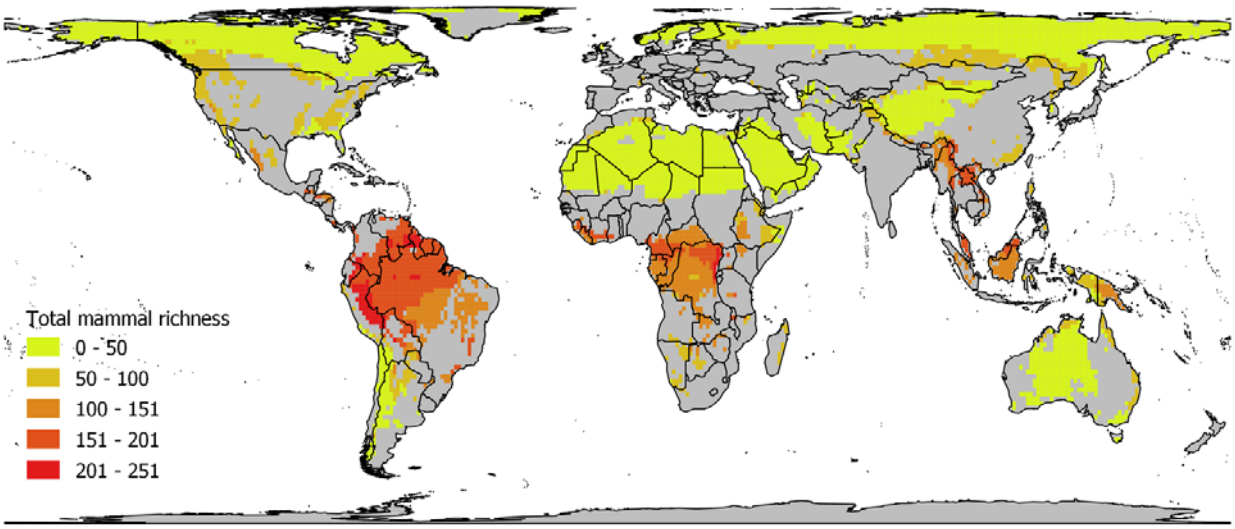



**Figure S4.1.** Total mammal richness per 110km x 110km grid-cell in *low-used* areas.

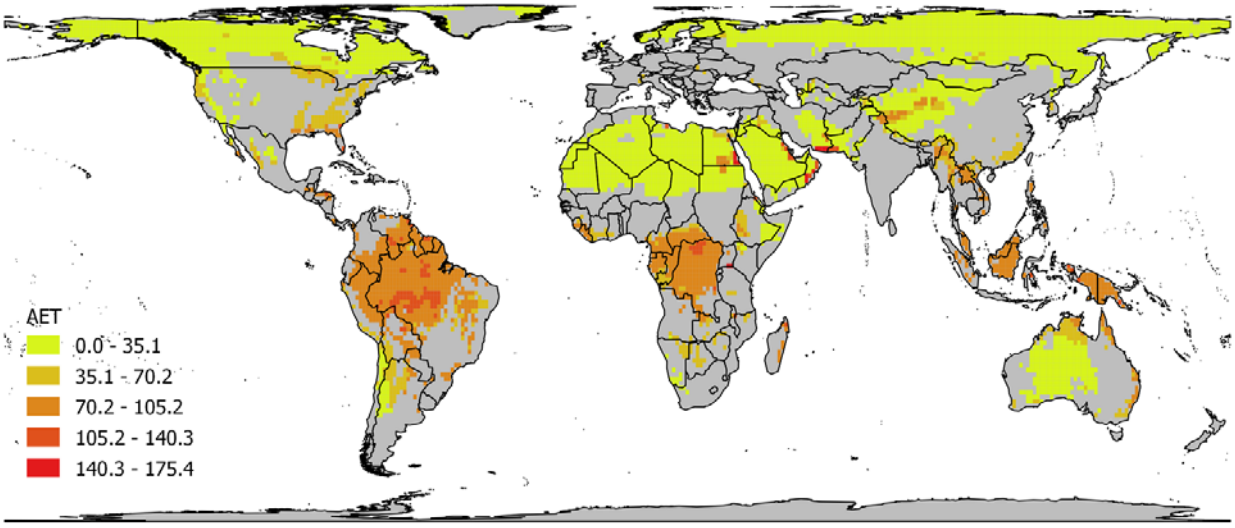



**Figure S4.2.** Mean actual evapotranspiration per 110km x 110km grid-cell in *low-used* areas.

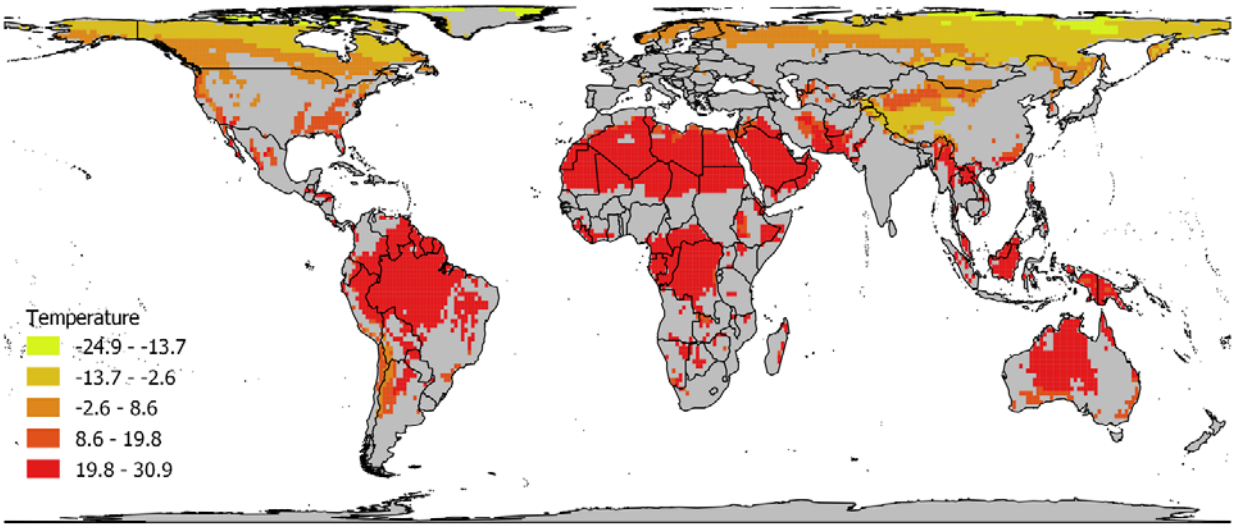



**Figure S4.3.** Mean annual temperature per 110km x 110km grid-cell in *low-used* areas.

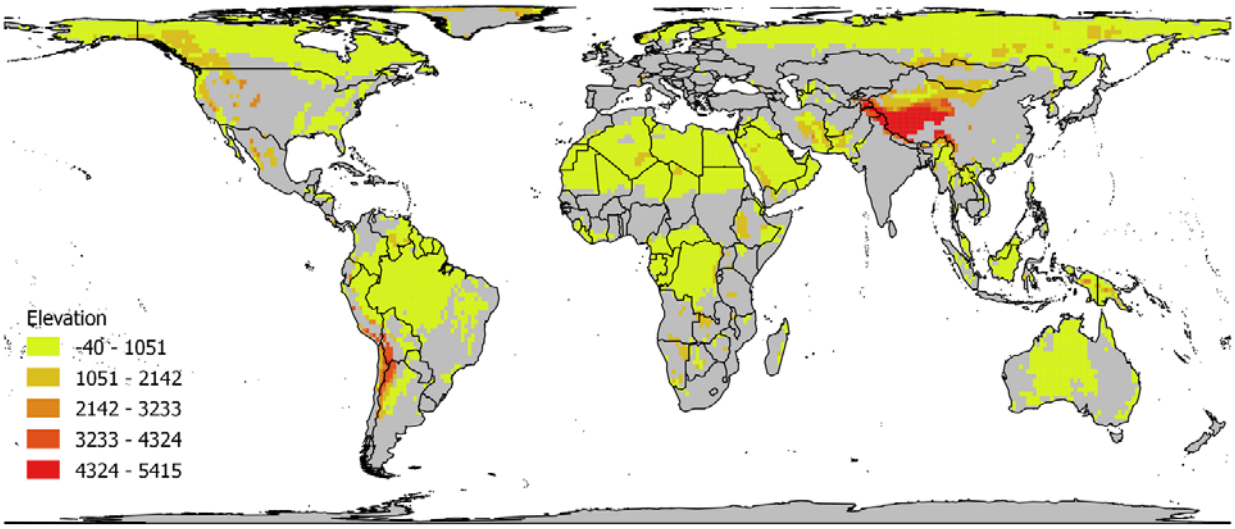



**Figure S4.4.** Mean elevation per 110km x 110km grid-cell in *low-used* areas.

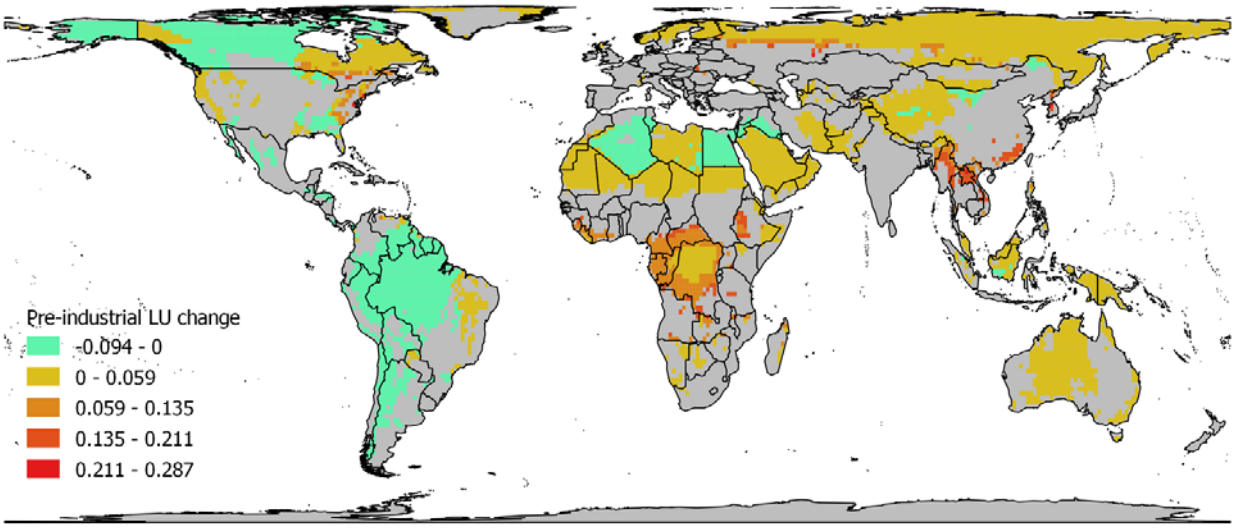



**Figure S4.5.** Mean pre-industrial rate of change per 110km x 110km grid-cell in *low-used* areas (from c.A.D.0 to c.A.D. 1750

***Recently-used areas***

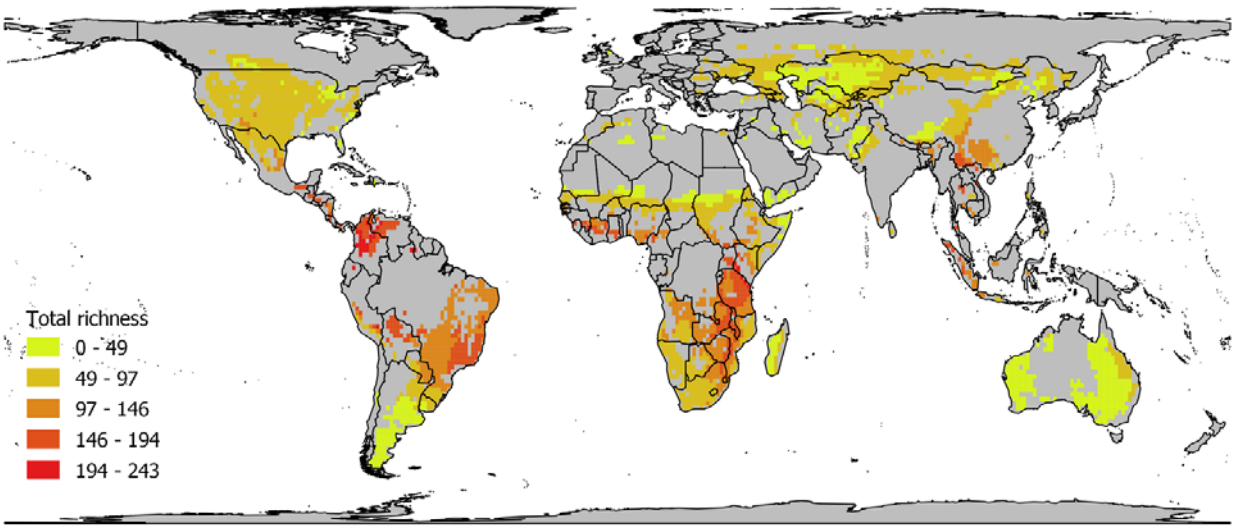



**Figure S4.6.** Total richness per 110km x 110km grid-cell in *recently-used* areas.

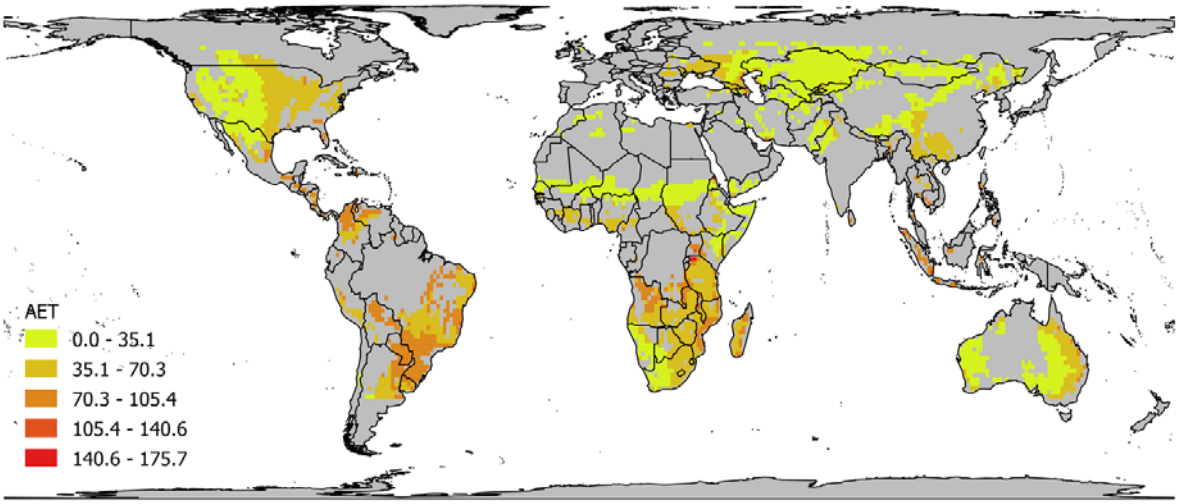



**Figure S4.7.** Mean actual evapotranspiration per 110km x 110km grid-cell in *recently-used* areas.

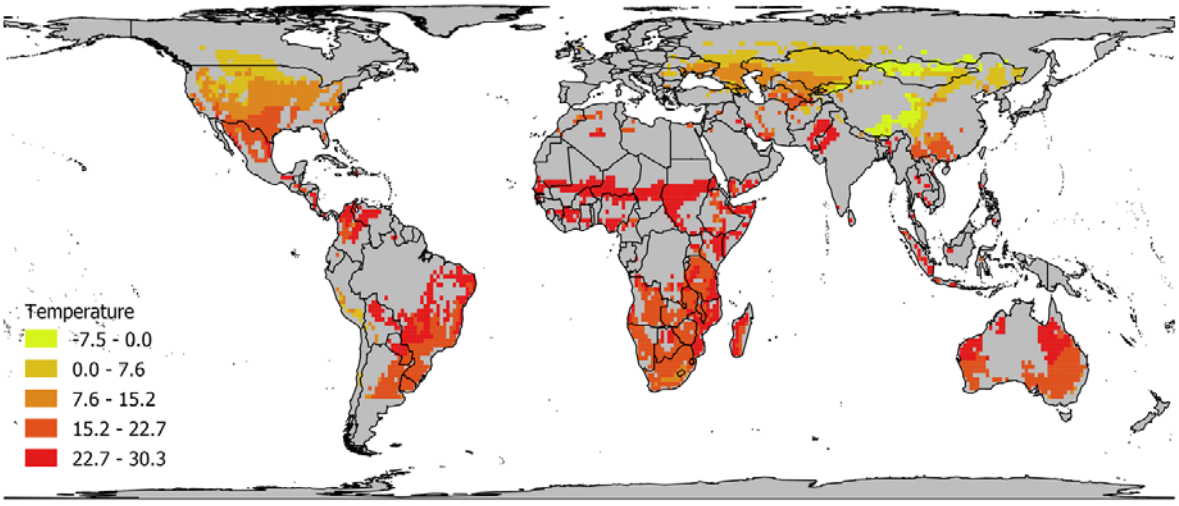



**Figure S4.8.** Mean annual temperature per 110km x 110km grid-cell in *recently-used* areas.

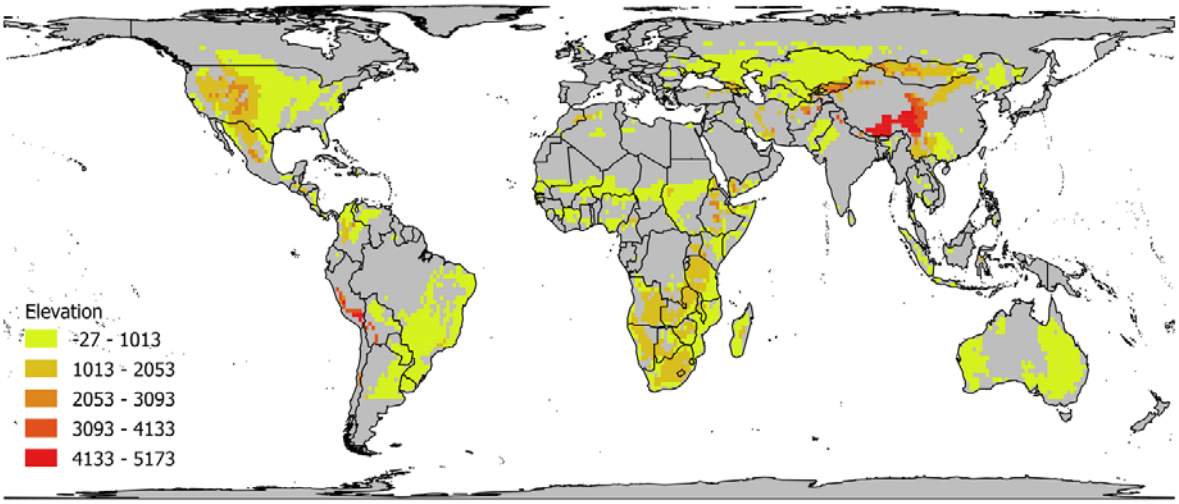



**Figure S4.9.** Mean elevation per 110km x 110km grid-cell in *recently-used* areas.

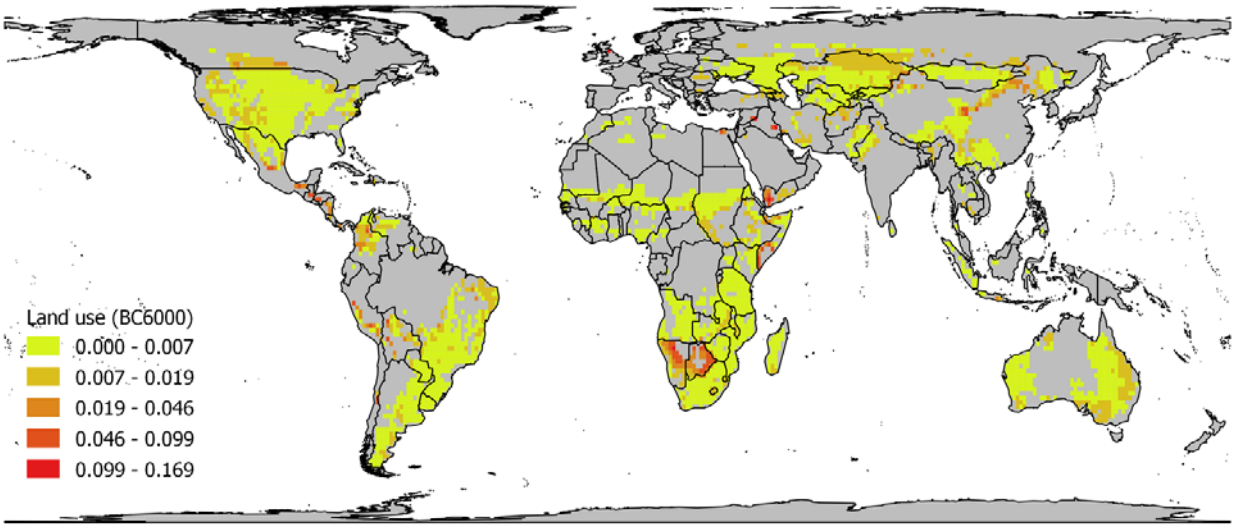



**Figure S4.10.** Mean proportion of land intended for human use c. B.C.6000 (per 110 x 110 km grid-cell) in *recently-used* areas.

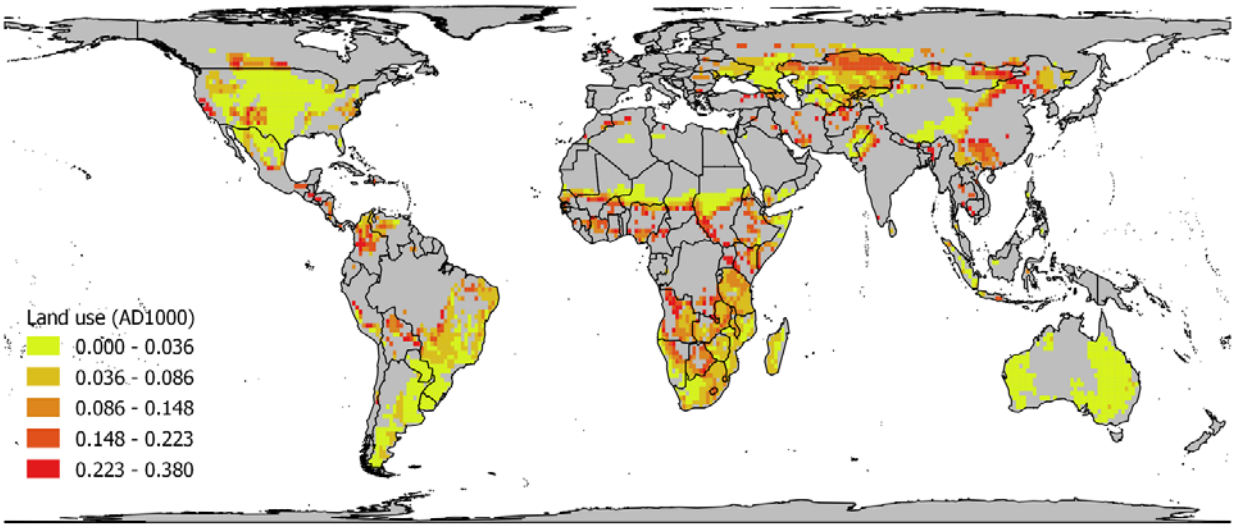



**Fig S4.11.** Mean proportion of land intended for human use c. A.D.1000 (per 110 x 110 km grid-cell) in *recently-used* areas.

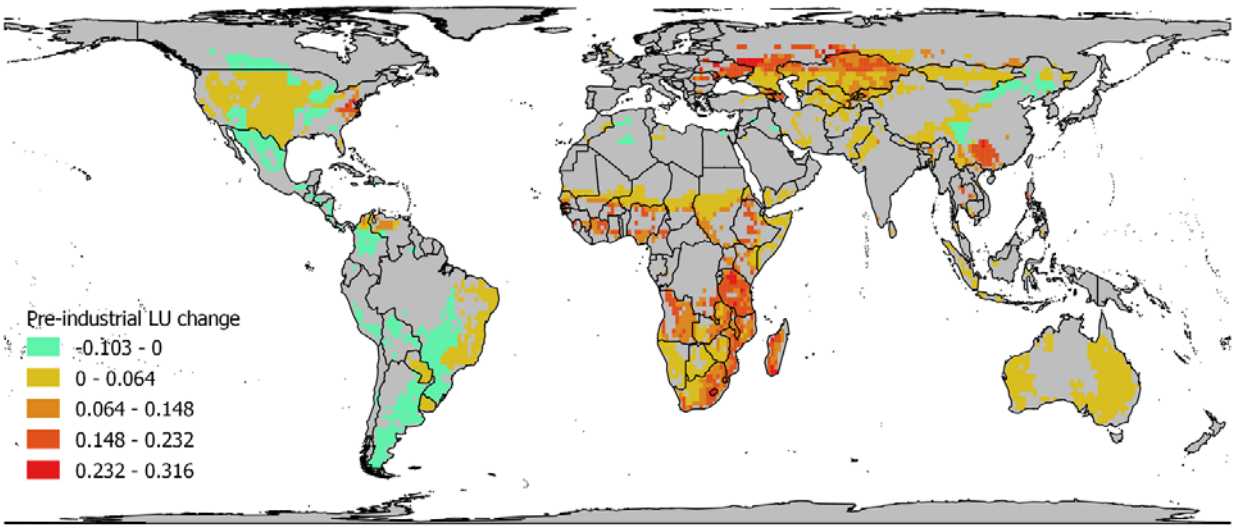



**Figure S4.12.** Mean pre-industrial rate of change per 110km x 110km grid-cell in *recently-used* areas (from c.A.D.0 to c.A.D. 1750).

***Steadily-used areas***

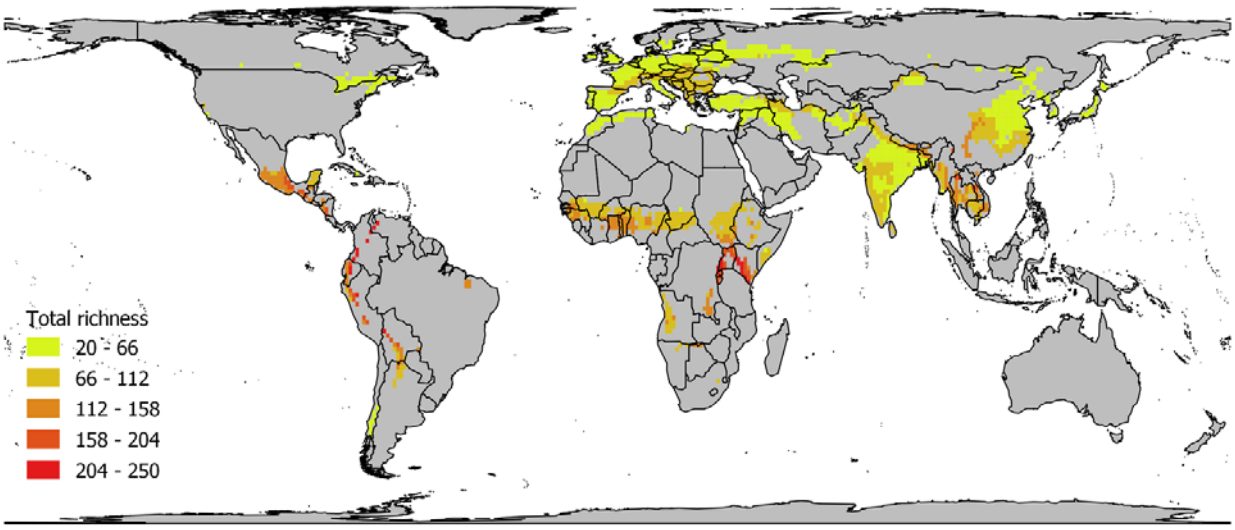



**Figure S4.13.** Total richness per 110km x 110km grid-cell in *steadily-used* areas.

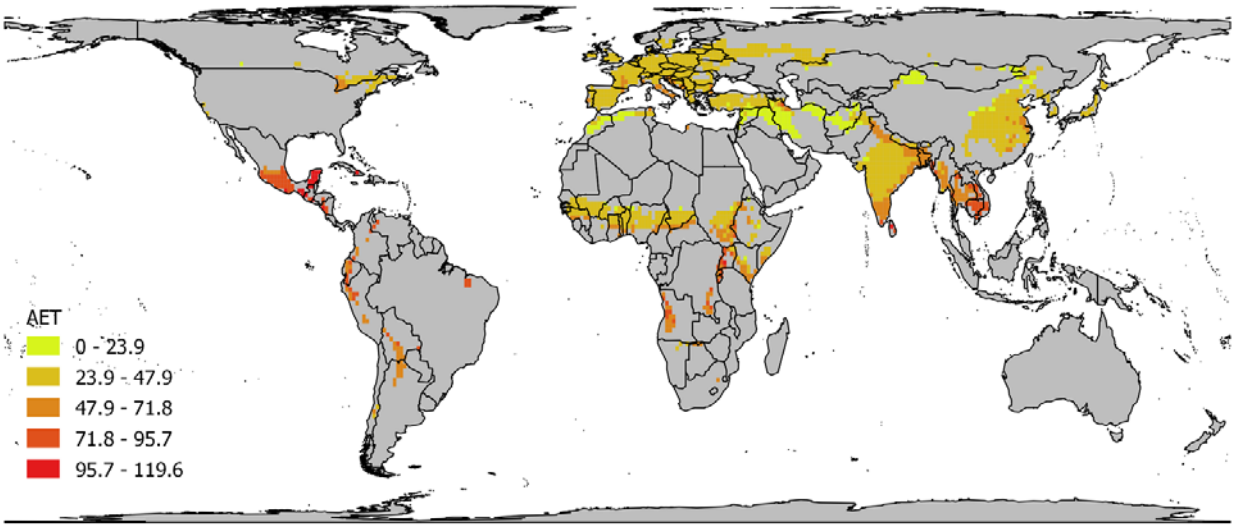



**Figure S4.14.** Mean actual evapotranspiration per 110km x 110km grid-cell in *steadily-used* areas.

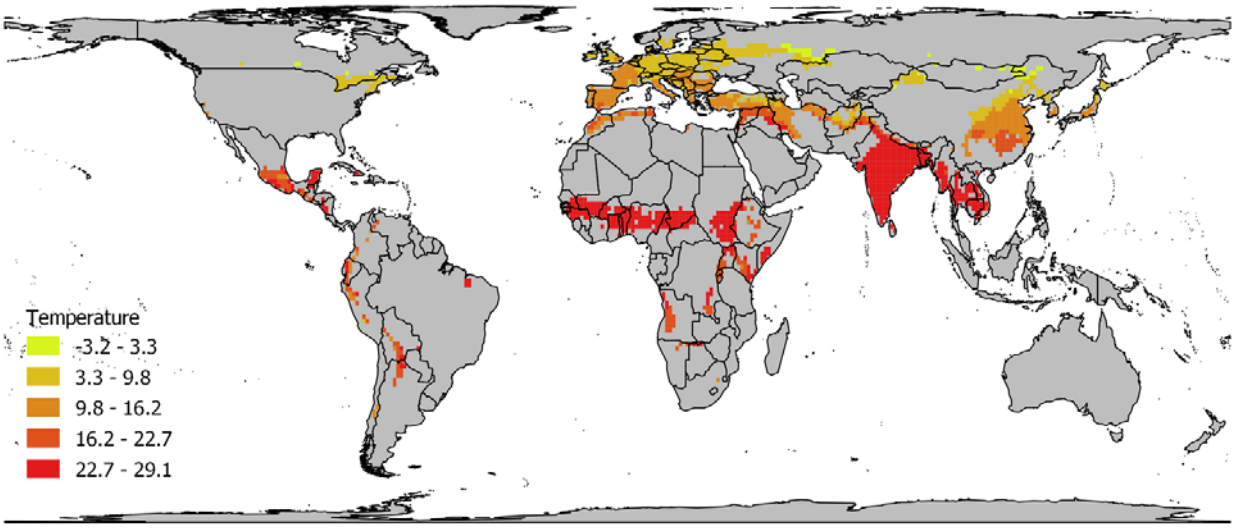



**Figure S4.15.** Mean annual temperature per 110km x 110km grid-cell in *steadily-used* areas.

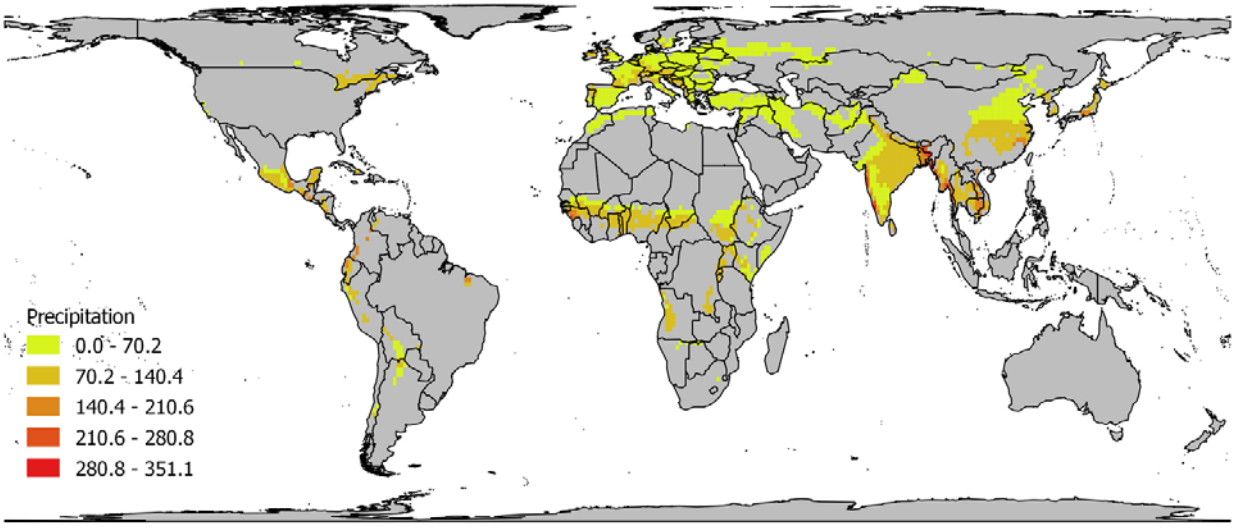



**Figure S4.16.** Mean annual precipitation per 110km x 110km grid-cell in *steadily-used* areas.

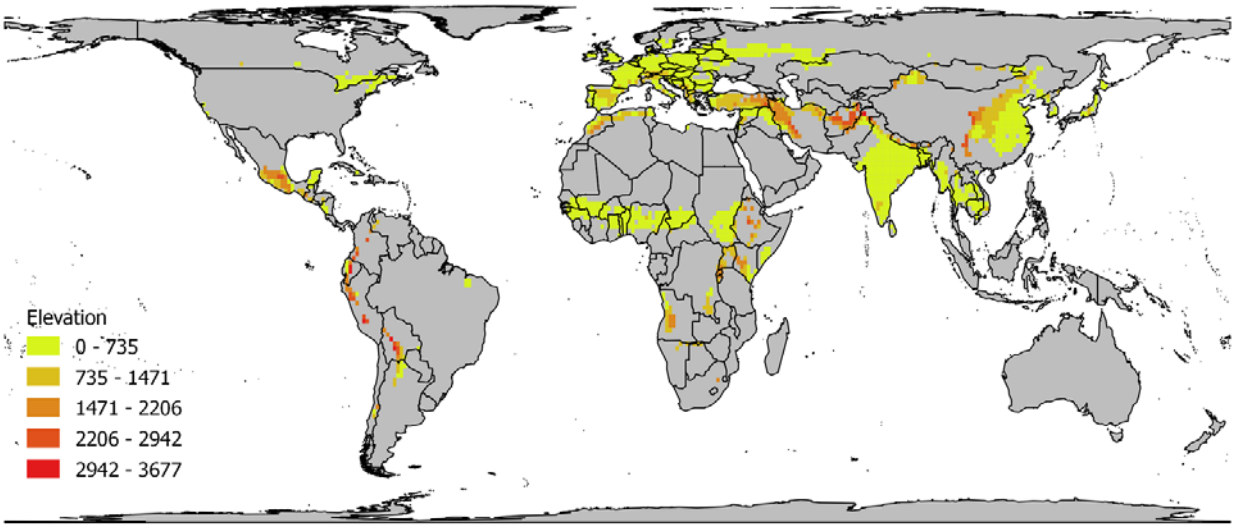



**Figure S4.17.** Mean elevation per 110km x 110km grid-cell in *steadily-used* areas.

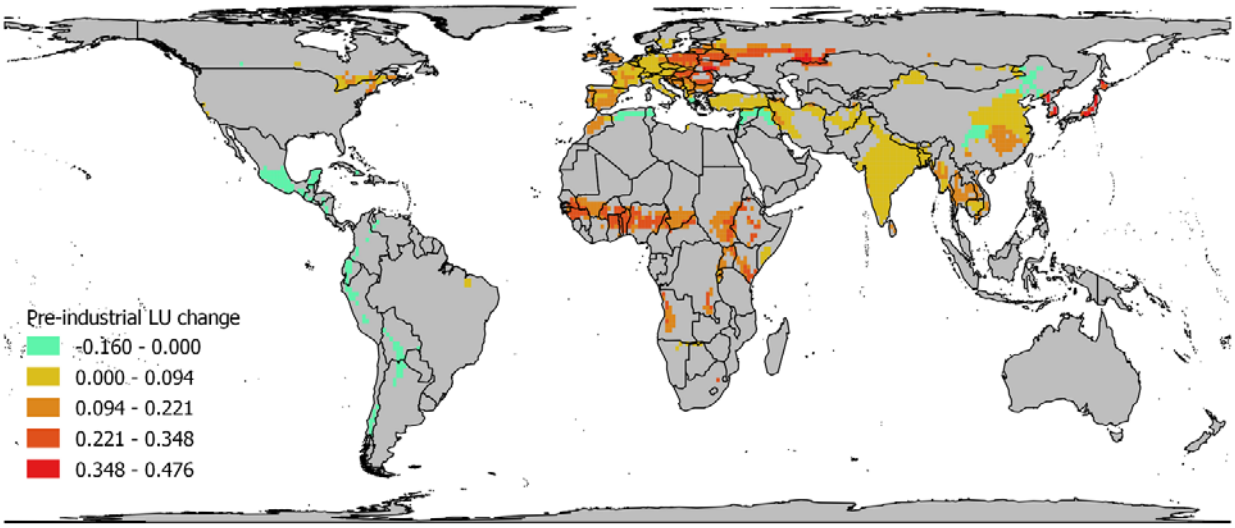



**Figure S4.18.** Mean pre-industrial rate of change per 110km x 110km grid-cell in *steadily-used* areas (from c.A.D.0 to c.A.D. 1750).