



Agriculture rivals biomes in predicting global species richness

Laura Kehoe, Cornelius Senf, Carsten Meyer, Katharina Gerstner, Holger Kreft and Tobias Kueimmerle

L. Kehoe (laura.kehoe@geo.hu-berlin.de), C. Senf and T. Kueimmerle, Geography Dept, Humboldt-Univ. Berlin, Berlin, Germany. TK also at: Integrative Research Inst. on Transformations of Human-Environment Systems (IRI THESys), Humboldt-Univ. Berlin, Berlin, Germany. – C. Meyer and H. Kreft, Biodiversity, Macroecology and Conservation Biogeography, Georg August-Univ. of Göttingen, Göttingen, Germany. CM also at: German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany. – K. Gerstner, Dept of Computational Landscape Ecology, UFZ – Helmholtz Centre for Environmental Research, Leipzig, Germany.

Species–area relationships (SARs) provide an avenue to model patterns of species richness and have recently been shown to vary substantially across regions of different climate, vegetation, and land cover. Given that a large proportion of the globe has been converted to agriculture, and considering the large variety in agricultural management practices, a key question is whether global SARs vary across gradients of agricultural intensity.

We developed SARs for mammals that account for geographic variation in biomes, land cover and a range of land-use intensity indicators representing inputs (e.g. fertilizer, irrigation), outputs (e.g. yields) and system-level measures of intensity (e.g. human appropriation of net primary productivity – HANPP). We systematically compared the resulting SARs in terms of their predictive ability.

Our global SAR with a universal slope was significantly improved by the inclusion of any one of the three variable types: biomes, land cover, and land-use intensity. The latter, in the form of human appropriation of net primary productivity (HANPP), performed as well as biomes and land-cover in predicting species richness. Other land-use intensity indicators had a lower predictive ability.

Our main finding that land-use intensity performs as well as biomes and land cover in predicting species richness emphasizes that human factors are on a par with environmental factors in predicting global patterns of biodiversity. While our broad-scale study cannot establish causality, human activity is known to drive species richness at a local scale, and our findings suggest that this may hold true at a global scale. The ability of land-use intensity to explain variation in SARs at a global scale had not previously been assessed. Our study suggests that the inclusion of land-use intensity in SAR models allows us to better predict and understand species richness patterns.

The species–area relationship (SAR) is one of the most robust patterns found in ecology (Rosenzweig 1995) and is crucial to our understanding of biodiversity patterns (Rosenzweig 1995, Turner and Tjørve 2005, Drakare et al. 2006, Dengler 2009). By relating the number of species to the area of habitat, the application of SARs is central in predicting species loss in areas of habitat loss and land-use change (Ladle and Whittaker 2011, Keil et al. 2015). A key step in SAR analyses is to accurately estimate the slope of the relationship, i.e. the rate of species loss related to area loss. However, applying a universal (canonical) slope and treating human-dominated land as inhospitable (Pimm et al. 1995, Brooks et al. 2002, Thomas et al. 2004) may be overly simplistic since SAR slopes are known to vary geographically (Drakare et al. 2006, Gerstner et al. 2014) and since numerous factors may allow for species survival in the matrix surrounding remaining habitat patches.

A complex interplay of ecological, evolutionary, and environmental factors influences species richness in a given area. For example, the importance of energy availability and

ecosystem productivity in predicting species richness has led to the development of the species–energy theory (Wright 1983), whereby, (Gaston 2000) at broad scales, a positive relationship is generally found for terrestrial vertebrates, where higher energy availability results in higher species richness (Currie 1991, Gaston 2000, Cusens et al. 2012). Other factors that can influence species richness in a given area include dispersal ability (Storch et al. 2005), evolutionary history (Belmaker and Jetz 2015), disturbance frequency, climate (Hawkins et al. 2003a, b, Kreft and Jetz 2007) and environmental heterogeneity (Hawkins et al. 2003a, b, Stein et al. 2014).

Many of the natural bioclimatic factors at play in driving species distributions and species richness may be summed up by biome classifications. This is because biomes represent major types of potential natural vegetation originating from distinct climatic conditions (Olson and Dinerstein 1998, Ladle and Whittaker 2011). Land cover has many similarities with biome classifications, as land cover represents the biophysical attributes of the land surface (Lambin

et al. 2001) and is determined by the climate, topography, and soil. Land cover additionally includes areas predominantly influenced by human activity such as croplands.

Agricultural expansion leading to land-cover conversions is one of the main drivers of species loss on a global scale (Sala et al. 2000, Pereira et al. 2012), but species also respond differently to habitat loss and degradation (Pereira and Daily 2006). Recent studies reflect this, for instance, through the development of matrix-calibrated SARs which incorporate land-cover change (Koh and Ghazoul 2010), and SARs which include species specific habitat-affinity in human-modified landscapes (Pereira and Daily 2006). Countryside biogeography also provides better insights into species survival in complex agricultural landscapes and forest fragments (Mendenhall et al. 2014).

While currently available land-cover datasets (Channan et al. 2014) and SAR models incorporating land use (Pereira and Daily 2006, Koh and Ghazoul 2010) distinguish between natural or agricultural land-cover types, land-management practices can differ greatly in what we broadly describe as agricultural land. In parallel to agricultural expansion leading to land-cover conversions, agriculture has also rapidly intensified since the 1950s. For example, global irrigated areas have doubled in size (FAOSTAT 2010) and fertilizer application has increased fivefold (Matson et al. 1997, Tilman et al. 2001).

This is problematic because high agricultural land-use intensity (LUI) is generally detrimental to local species richness and abundance (Newbold et al. 2015). However, despite the global increase in LUI, most studies investigating land use and biodiversity are local in scale and either disregarded LUI completely or used only a single metric to measure it (Herzon et al. 2008, Kleijn et al. 2009, Geldmann et al. 2014). The latter approach has been shown to be simplistic as LUI is a multidimensional concept that embodies a wide variety of management practices that can have diverse effects on biodiversity. For instance, fertilizers and pesticides pose a substantial threat to terrestrial vertebrates (Kerr and Cihlar 2004, Gibbs et al. 2009, Kleijn et al. 2009). Long-term irrigation can salinize soils which can eventually become toxic to plants with potentially detrimental effects to entire ecosystems (Yamaguchi and Blumwald 2005). Intensive livestock grazing can have negative effects on biodiversity (Alkemada et al. 2012) and ecosystems, especially in the absence of remnant vegetation (Felton et al. 2010). All of these effects are of particular concern since different combinations of high LUI concordant with high biodiversity are spread heterogeneously across the globe (Kehoe et al. 2015) and may have region-specific effects on biodiversity. Therefore, while it is generally not accounted for, the intensity of agricultural land-use may improve predictions of SARs in human-modified landscapes.

While the inclusion of biomes and land cover has recently been shown to improve SAR predictions for plants on a global scale (Gerstner et al. 2014), it remains unclear whether this extends to other taxa, and whether the inclusion of measures of land-use intensity improves global SAR models. Furthermore, the importance of human influence on species richness is often embraced at local grains (Dornelas et al. 2014, Newbold et al. 2015), however, recent research is emerging that indicates broader patterns of species richness

might also be related to human activities more than we suspect (Murray and Dickman 2000, Di Marco and Santini 2015).

Here, we first evaluated the extent to which the inclusion of agricultural activity and management in the form of land cover and land-use intensity improves global SAR models. To account for the multidimensionality of land-use intensity, we assessed three broad categories of agricultural management metrics, representing input (the intensity of land use along different input dimensions, e.g. fertilizer and irrigation), output (the ratio of outputs from agricultural production, e.g. yields, $\text{t ha}^{-1} \text{ yr}^{-1}$) and system metrics (the relationship between the inputs or outputs of land-based production to the overall system, e.g. human appropriation of NPP). Following this step, we compared whether this improvement is comparable to the inclusion of climate conditions and potential natural vegetation embodied by biome classifications. Therefore, we test a proxy for human factors in the form of land-cover and LUI, against a proxy for environmental factors, in the form of biomes, in their ability to predict SARs on a global scale.

Methods

Species data

We focused on terrestrial mammals due to their high endangerment status, 22% of mammals are currently threatened according to the IUCN (2013), and the availability of a recently updated global range maps (Schipper et al. 2008, IUCN 2013). We used extent-of-occurrence range maps provided by the IUCN (2013), which we overlaid with a grid to infer broad-scale species richness patterns. These range maps are currently considered the most comprehensive and detailed global dataset of mammal distributions (Di Marco and Santini 2015). Range maps are expert-based maps of mammal distributions that depict the extent of occurrence, i.e. areas containing all known species occurrences. However, like all global spatially explicit datasets, errors and gaps occur. For example, species' areas of occupancy can be overestimated at fine spatial resolutions by including uninhabited areas (Jetz et al. 2008). We therefore scaled the data to an equal area grid of approximately 110×110 km or 1 degree at the equator as finer resolutions lead to high levels of false presences (Hurlbert and Jetz 2007). We excluded all cells with $< 50\%$ land area to minimize confounding effects of coastal areas, predominantly marine species, and small oceanic islands.

Biome and land-use data

We used 14 biomes as defined by Olson and Dinerstein (1998, Fig. 1a). For land cover, we used 16 classes from the MODIS land cover map (Channan et al. 2014, Fig. 1b). To assess LUI, we explored three categories of metrics related to the intensity of a) inputs to agriculture, b) outputs from agriculture, and c) changes in system-level variables due to agriculture (Kuemmerle et al. 2013). Input metrics relate to the intensity of land management along input dimensions, such as fertilizer use and irrigation. Output metrics describe

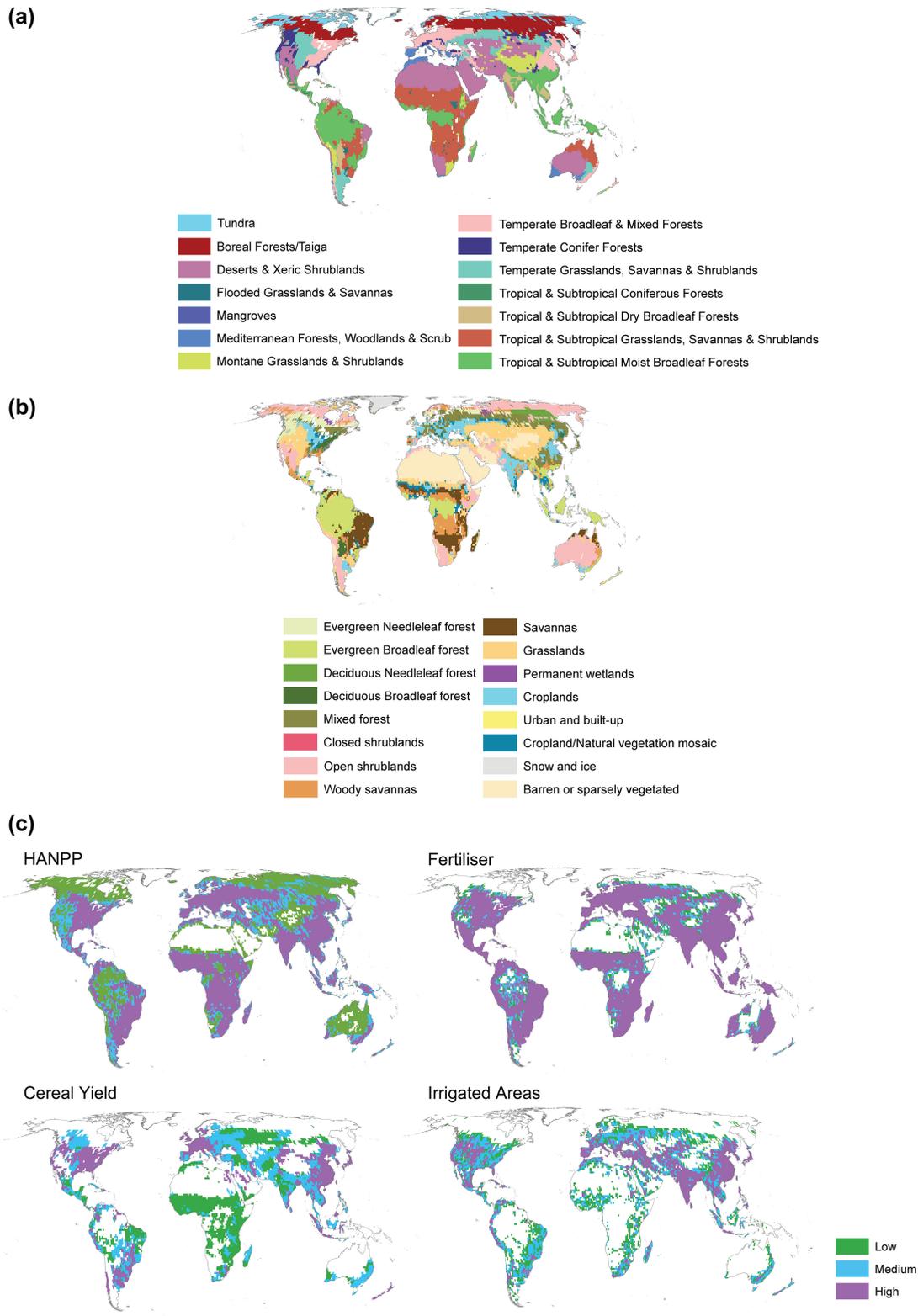


Figure 1. Maps of potential factors causing variation of species–area relationships (SARs): (a) biomes (Olson and Dinerstein 1998), (b) land cover (Channan et al. 2014) and (c) land-use intensity (LUI) split into high (shown in purple), medium (blue), low (green) and no use (white) levels for the following datasets: human appropriation of net primary productivity (HANPP; Haberl et al. 2007), fertiliser inputs (Potter et al. 2010), cereal yield (Monfreda et al. 2008) and areas equipped for irrigation (Siebert et al. 2005). Due to many 100% values, areas of high intensity are larger than other areas. The maps are projected using Eckert IV projection.

the ratio of inputs and outputs, for example, yields (harvests/area). System-level metrics refer to the relationship between land management and properties of the socio-ecological

system as a whole, such as the percentage of human appropriation of NPP (HANPP; Haberl et al. 2007), and can provide a general idea of the overall management intensity.

As input metrics, we chose areas equipped for irrigation measured in percentage of each grid cell (Siebert et al. 2005) and N-fertilizer application measured in percentage of each grid cell under fertilization (Potter et al. 2010). Output metrics included cereal yields measured in $\text{t ha}^{-1} \text{yr}^{-1}$ (Monfreda et al. 2008). As system metric, we chose an integrated measure of land-use pressure on the environment, namely, human appropriation of NPP, which entered the analyses as percentage of each grid cell where any level of NPP is appropriated (Haberl et al. 2007). The base cropland and land-cover maps used for the generation of the above datasets are given in the Supplementary material Appendix 1, Table A1.

Statistical analyses

To construct SARs, we took 500 samples with replacement across our global grid. Samples were chosen randomly in terms of the total land area they covered, and ranged from a square window size of 1×1 to 15×15 grid cells. Samples were randomly placed and non-nested, i.e. one sample was not necessarily contained within the previous sample but entirely random in location, therefore some overlap could occur (resulting in a type IIB SAR curve, Scheiner 2003).

Our models were based on the power law SAR, where $S = c \times A^z$ relates species richness (S) to the area (A) of habitat, ' z ' is the rate of change in species numbers, and ' c ' is the taxon- and region-specific constant of per unit area species richness (Arrhenius 1921). We systematically fitted different interactions to SARs that take into account the potential effects of land cover, biomes and each LUI metric, and compared their ability to predict large-scale species richness patterns. We tested two different model types. The first model fitted the species–area relationship with area as the only predictor. The equation for this universal global model takes the form of:

$$\log_{10}(S) = \log_{10}(c) + z * \log_{10}(A) \quad (1)$$

The second model included additional terms related to the percentage cover of either: biome, land cover or LUI, in each sampling unit. These were added as interaction effects to the area term in the model, as shown in Eq. 2:

$$\log_{10}(S) = \log_{10}c + \sum_{i=1}^n z_i \log_{10}A * R_i \quad (2)$$

' R_i ' refers to the proportional area for each class n (i.e. biome class, land cover class, land-use intensity class, etc.). The biome model included 14 biomes and the land cover model included 16 land cover classes (excluding water). In order to generate SARs for the LUI models, we generated four classes – no LUI (where there was no agricultural activity), followed by high, medium, and low LUI (split by terciles, Fig. 1c). A separate model was run for each LUI metric resulting in a total of seven models – one universal global model, one biome model, one land cover model, and four LUI models.

To estimate the predictive power of each of the seven models, we applied a 10-fold cross-validation and calculated the squared correlation coefficient between predicted and observed values (following abbreviated with r^2) (Harrell 2001). During initial model development we found spatial autocorrelation in the residuals (from Moran's I), we therefore

followed the approach taken by Gerstner et al. (2014) and employed simultaneous autoregressive models (Kissling and Carl 2008) using the R 3.1.2 statistical analysis software (R Core Development Team), function 'spautolm' in the package 'spdep' (Bivand et al. 2012). This method assumes spatial autocorrelation in a second error term which explicitly models spatial dependence in the residuals (Dormann 2007) and is an established method for accounting for spatial autocorrelation in SAR samples (Kissling and Carl 2008). We chose a neighbourhood structure based on the minimization of the residual spatial autocorrelation (Kissling and Carl 2008, Gerstner et al. 2014). We found an optimal neighbourhood distance of five grid cells ($550 \times 550 \text{ km}$).

Due to the spatial structure of the data, parameter estimates were strongly influenced by the random spatial configuration of the 500 samples, i.e. sample location had a large effect on the r^2 . We therefore re-ran our sampling approach 1000 times (each run contained 500 different random sampling locations). Our final results reported here are thus based on the average parameter estimates of 1000 sampling runs and associated model runs after trimming the most extreme 5% of model results. We tested linear, logarithmic and power models. In line with previous studies (Connor and McCoy 1979, Dengler 2009, Gerstner et al. 2014, Matthews et al. 2015), we report here only the power SAR (where area and species richness are \log_{10} -transformed prior to analysis), results for the linear and logarithmic models are reported only in the Supplementary material Appendix 1, Table A2.

Results

The inclusion of biomes, land cover or LUI all significantly improved the predictive ability of SARs compared to the universal global model (Table 1). The biome model and the HANPP model had the best predictive power, both with a cross-validated r^2 of 0.49 (compared to an r^2 of 0.15 for the global model). The land cover model had the third highest r^2 of 0.46 (Table 1). Thus, modelling according to one global relationship would lead to large over- or underestimations of species richness, depending on the bio-physical characteristics of the area of interest. We found a wide margin in

Table 1. Predictive ability of each simultaneous autoregressive model via 10-fold cross validation (results are averaged over 1000 model runs). The global model only included area as a predictor of species richness. Other models included either: biomes (Olson and Dinerstein 1998), land cover (Channan et al. 2014) or land-use intensity (LUI) split into high, medium, low and no-use levels for the following datasets: human appropriation of net primary productivity (HANPP; Haberl et al. 2007), fertiliser inputs (Potter et al. 2010), cereal yield (Monfreda et al. 2008) and areas equipped for irrigation (Siebert et al. 2005).

	Mean r^2	SD
LUI – HANPP	0.49	0.11
Biome	0.49	0.14
Land cover	0.46	0.15
LUI – fertiliser	0.44	0.13
LUI – cereal yield	0.31	0.13
LUI – irrigation	0.26	0.12
Global	0.15	0.11

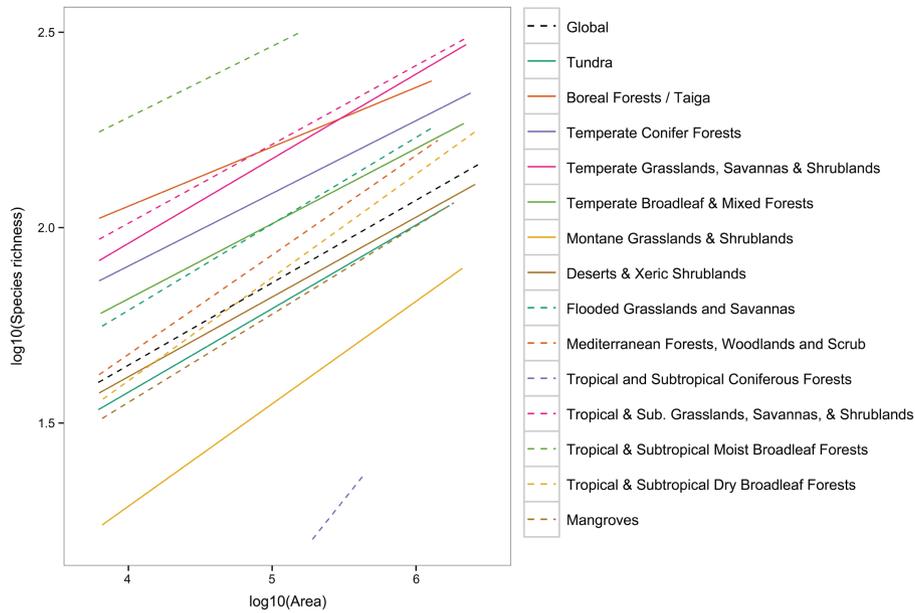


Figure 2. Species–area relationships (SARs) in log–log space (area relates to km²) for biomes.

the performance of LUI metrics – ranging from average r^2 values of 0.49 to 0.26 (Table 1), along with many different relationships with species richness in terms of high, medium and low LUI. HANPP, the only system metric investigated, out-performed all other LUI metrics (Table 1).

SARs by biome

The addition of biomes to the global model increased its predictive power from $r^2 = 0.15$ to $r^2 = 0.49$. Furthermore, SARs for individual biomes differed both in their intercept and z -values (Fig. 2). The boreal forest/taiga biome had the

lowest z -value estimate (0.14), which indicates low levels of species increase in larger areas. The highest z -value and thus highest rate at which species richness increases with area was found in the tropical and subtropical coniferous forests biome ($z = 0.49$).

SARs by land cover

Land-cover specific SARs also increased the model r^2 (0.46). As for biomes, we found a large range in SAR parameter estimates (Fig. 3, Table 2). The highest z -value and thus the highest rate of species gain with increasing area was found

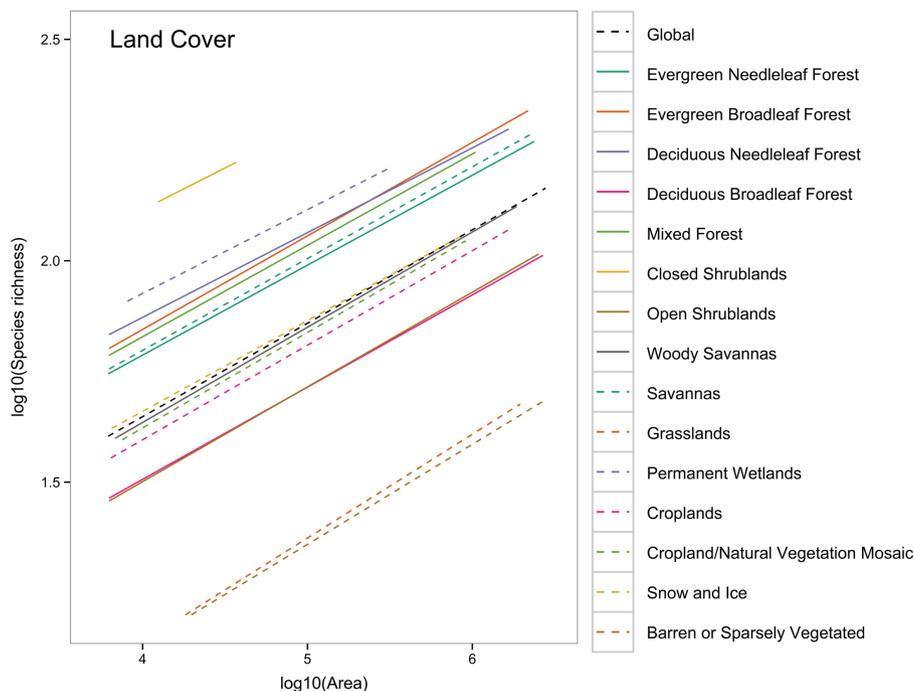


Figure 3. Species–area relationships (SARs) in log–log space (area relates to km²) for land cover.

Table 2. Parameter estimates for the species–area relationship (SAR): the slope z and intercept, $\log_{10}(c)$, of SARs in log–log space. Biome and land cover (LC) classes with less than 5% of the total land area are indicated with an *.

		z	Intercept
Global		0.22	0.75
Biome	Tundra	0.21	0.70
	Boreal forests/taiga	0.14	1.07
	Temperate conifer forests	0.20	0.88
	Temperate grasslands, savannas and shrublands	0.20	0.91
	Temperate broadleaf and mixed forests	0.20	0.83
	Montane grasslands and shrublands	0.27	0.42
	Deserts and xeric shrublands	0.21	0.71
	Flooded grasslands and savannas	0.23	0.76
	Mediterranean forests, woodlands and scrub	0.26	0.65
	Trop. subtrp. coniferous forests	0.49	0.00
	Trop. subtrp. grasslands, savannas and shrub	0.21	0.90
	Trop. subtrp. moist broadleaf forests	0.17	1.12
	Trop. subtrp. dry broadleaf forests	0.26	0.60
	Mangroves*	0.23	0.65
LC	Evergreen needleleaf forest	0.20	1.01
	Evergreen broadleaf forest	0.22	1.04
	Deciduous needleleaf forest	0.20	1.11
	Deciduous broadleaf forest	0.29	0.72
	Mixed forest	0.20	1.04
	Closed shrublands*	0.27	1.29
	Open shrublands	0.23	0.75
	Woody savannas	0.24	0.85
	Savannas	0.21	1.01
	Grasslands	0.31	0.40
	Permanent wetlands*	0.20	1.15
	Croplands	0.23	0.83
	Urban and built-up*	0.18	0.37
	Cropland/natural vegetation mosaic	0.24	0.84
	Snow and ice	0.05	0.97
	Barren or sparsely vegetated	0.23	0.44
HANPP	No LUI (0)	0.20	0.59
(% use/grid)	Low (0.6–91.3)	0.22	0.86
	Med (91.3–99.9)	0.22	0.72
	High (100)	0.23	0.72
Fert	No LUI (0)	0.21	0.73
(% use/grid)	Low (2.9–40)	0.24	0.81
	Med (40–80)	0.27	0.41
	High (80–100)	0.23	0.73
Irr	No LUI (0)	0.21	0.78
(%/grid)	Low (0.01–0.2)	0.23	0.70
	Med (0.2–1.7)	0.21	0.88
	High (1.7–82)	0.26	0.63
Cereal	No LUI (0)	0.21	0.77
(t ha ⁻¹ yr ⁻¹)	Low (0.2–1.6)	0.22	0.77
	Med (1.6–3.1)	0.23	0.80
	High (3.1–10.7)	0.23	0.81

for grasslands (0.31), snow and ice was found to have lowest z -values of 0.05. Results for land cover classes of less than 5% of the total area are not reported here as they tended towards extreme results due to their small area, and thus lack of samples, these comprise of closed shrublands, permanent wetlands and urban and built-up areas (see Supplementary material Appendix 1, Table A3 for standard deviations, and 5% and 95% percentile values of estimates).

SARs by LUI

While all LUI metrics improved the predictive ability of the models from the global baseline, there was a wide margin

in r^2 , from the HANPP model with an r^2 as high as that of biomes (0.49) to the irrigated areas model with an r^2 of 0.26. Furthermore, the relationship between different levels of LUI and species richness was not constant across LUI metrics (Fig. 4, Table 2).

Compared to the biome and land-cover parameter estimates, a relatively low range in z -values and species richness predictions for LUI models was found. The highest species increase with area was found for medium levels of fertilizer application ($z = 0.27$), the lowest species increase with area was found where there was no HANPP activity ($z = 0.20$). In terms of the overall relationship of species richness and LUI, for the HANPP model, low intensity was associated with highest species richness, followed by high intensity

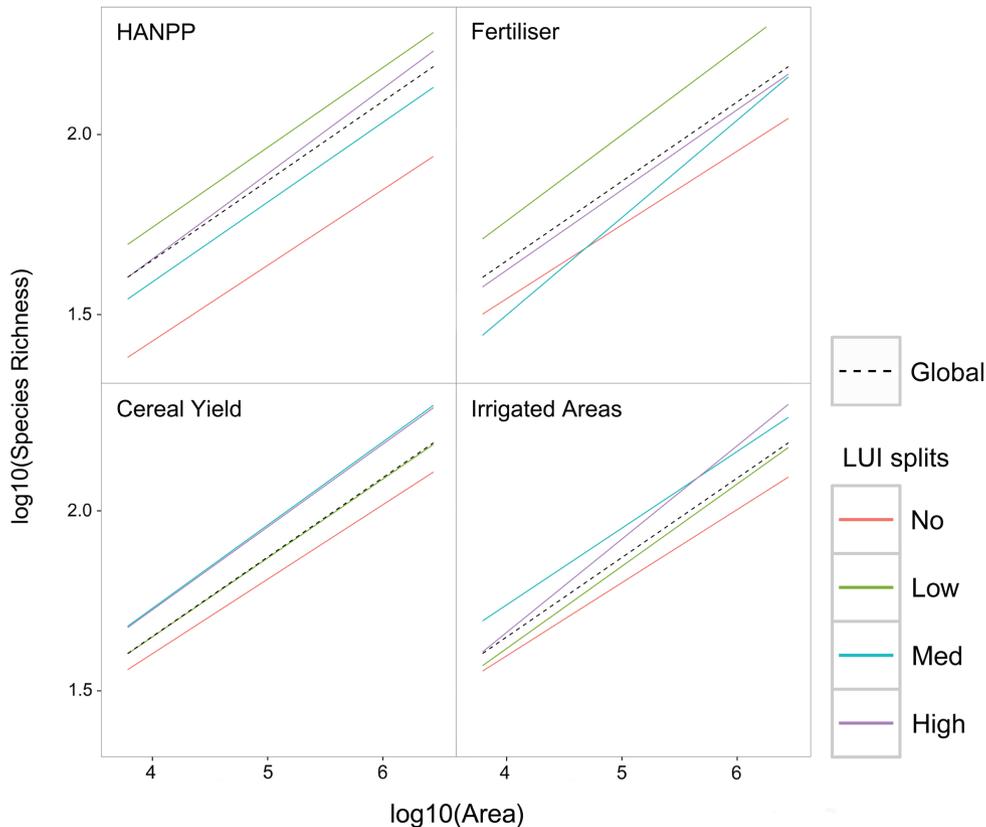


Figure 4. Species–area relationships (SARs) in log–log space (area relates to km²) for land-use intensity (LUI), split into high, medium, low and no-use levels for the following datasets: human appropriation of net primary productivity (HANPP; Haberl et al. 2007), fertiliser inputs (Potter et al. 2010), cereal yield (Monfreda et al. 2008) and areas equipped for irrigation (Siebert et al. 2005).

and then medium levels associated with the lowest levels of species richness (Fig. 4, Table 2). Fertilizer application exhibited the same overall relationship as HANPP, but with one distinct difference – due to a higher z-value in the SAR for medium-intensity fertilizer application, in larger areas, medium- and high-intensity fertilizer application were associated with similar levels of species richness (Fig. 4, Table 2). For cereal yields, medium- and high-intensity were associated with similarly high species richness regardless of area. Unlike HANPP and fertilizer application, higher species richness was associated with higher LUI for cereal yields and irrigated areas (Fig. 4). For all LUI metrics tested, species richness numbers were lowest in areas without any land-use, which generally represent ice-covered and desert lands.

Spatial arrangement of samples

Across all models, results varied substantially depending on the spatial location of the samples. When examining results from one single model run (with one sample set) we found that the model r^2 ranged from a minimum of zero to a maximum of 0.87 (Supplementary material Appendix 1, Table A4). Therefore, the random location and size of the samples alone, in extreme cases, could account for an r^2 that explained nothing or close to all variation in species richness.

Discussion

The objective of this study was to assess whether SARs are improved by better representing the geographic variation of its parameters. We found that the addition of biomes, land-cover and land-use intensity (LUI) all improve global predictions of species richness. Furthermore, some land-cover and LUI metrics perform as well as biomes in predicting species richness.

In terms of LUI metrics, we found diverse interactions with SARs both in predictive ability and relationship between high, medium and low LUI and species richness. This adds evidence to research suggesting that metrics of LUI have distinct global patterns (Kehoe et al. 2015) and relationships with biodiversity (Yamaguchi and Blumwald 2005, Felton et al. 2010, Alkemade et al. 2012).

Finally, we found that HANPP, our only overall metric of LUI, was the best predictor of species richness when compared to other LUI metrics. This shows that broader LUI metrics can better predict SARs, likely due to their comprehensive nature. The predictive ability of HANPP may also provide support for the species–energy hypothesis, since net primary productivity can be seen as a form of available energy (Wright 1990). The human appropriation of high levels of energy in the form of net primary productivity may result in a loss of species richness at a landscape scale (Haberl et al. 2014). We find that low HANPP levels are associated with higher species richness, however, high and medium HANPP

levels are less intuitive – with species richness higher in areas of high HANPP. Our analysis was not causal but predictive, thus we cannot provide strong evidence with regard to the shape of the species–energy or species–HANPP relationship. Overall, we show that at a broad spatial grain, factors related to human activity are on a par with biophysical factors in predicting species richness.

Geographic variability in SARs

We found that including spatially explicit variables in species richness models improves predictions of global SARs. Furthermore, we found a signal between land use and species richness that is equally strong as that between species richness and biomes. Thus, despite most research focusing on a local grain size when addressing the relationship between land use and species richness (see Newbold et al. 2015 for review), and global studies with larger grains generally focusing on natural biophysical drivers (Hawkins et al. 2003a, b, Field et al. 2009, Hortal et al. 2012), we show that human factors may play a more dominant role in predicting global biodiversity patterns than previously thought.

Our analyses do not provide a causal link of land use and biodiversity patterns. This link has been shown at local scales, where land use in the form of conversion from natural habitat and intensification of existing agricultural land results, on average, in decreased species richness (Newbold et al. 2015). Our results are the first to show how impacts may aggregate to affect species richness patterns at the global scale, which is important considering the acceleration of land-use change in recent times, and its importance in driving both current and future biodiversity loss (Sala et al. 2000, Pereira et al. 2012). However, land use itself depends on climate, soils, and productivity. Many of these same factors are the basis on which biomes are delineated, where species richness patterns are also closely related to climate and productivity variables (Hawkins et al. 2003a, b). Thus, attribution as to which factors are driving species richness patterns (land use vs. environmental factors) is challenging based on broad-scale analyses.

When compared with results from previous research, our biome model parameters performed similarly to those found by Gerstner et al. (2014) for plant species richness. In both our results and those of Gerstner et al. (2014), the boreal forest/taiga biome had the lowest z estimate (0.14 and 0.08 respectively). Furthermore, the biome with the largest z estimate was the tropical and subtropical coniferous forests biome (0.49 and 0.45). This indicates that there is a high concordance of biome explicit SARs for plants and mammals on a global scale.

Regarding the predictive ability of the models, the main difference in results is that Gerstner et al. (2014) found that land cover had a relatively lower r^2 when compared to our land cover result. This may be due to five reasons. First, the biodiversity datasets for plants used by Gerstner et al. (2014) were only available for a set number of locations, thus producing limitations in global predictions. Second, the plant dataset used by Gerstner et al. (2014) was not as up to date as the mammal data (IUCN 2013) used here.

Third, plants and mammals have different responses to land-use and LUI (Gibson et al. 2011). Fourth, our study uses a different land cover map from 2014, not available at the time of Gerstner's study. Finally, both studies are global in extent and at this scale species extinctions are relatively rare, where local extinctions and range contractions are more common, such processes are often not reflected at our coarse spatial grain.

Diversity of LUI metrics

Understanding the relationship between LUI and biodiversity is important since LUI is set to further accelerate in the future as 'sustainable intensification' gains support (Foley et al. 2011). Previous studies focusing on small grain sizes have found that a higher proportion of agricultural land and higher LUI can have negative effects on biodiversity (Martins et al. 2014, Newbold et al. 2015). However, due to the diverse patterns of LUI metrics globally and their likely diverse effects on biodiversity, we expected that LUI metrics would have a variety of relationships with species richness. Our results suggest that this is the case and that LUI metrics have varied relationships with species richness – at least for a 1 degree grain size.

In terms of model performance, LUI metrics exhibited diverse predictive abilities in relation to SARs, ranging from 49% (HANPP) to 26% (areas equipped for irrigation). This again illustrates the non-uniformity in LUI metrics, not just in their spatial patterns but also in their ability to predict global patterns of species richness. Of LUI metrics, HANPP had the best predictive ability. This is likely because this metric covers a wider variety of potential agricultural land-uses, namely, wherever any form of activity related to appropriating NPP is present. It is thus logical that the LUI metrics that cover a broader spectrum of human land-use will naturally have the best predictive ability.

We show that there is a large diversity in the relationship between high, medium and low LUI, and species richness, however, our research is of coarse spatial grain, with relatively large distributions in parameter estimates (Table 2). We thus cannot provide the answers as to which forms of LUI and at what level may be most detrimental to biodiversity. For this, experimental and observational small-scale research and synthesis are needed. However, we do show that there is a large diversity in the relationship between LUI and species richness and that the global story is not as one dimensional as fine-scale studies often suggest, i.e. higher LUI results in lower species richness.

Importance of the spatial location of samples

We found that our modelling results were highly dependent on the spatial location and size of samples (Supplementary material Appendix 1, Table A4), while we controlled for this by running 1000 models and taking the average of the parameter estimates, many studies do not have this option and must work with the limited samples that are available. Our results have implications for studies which use incomplete datasets and often draw broad conclusions. Where

studies are not as fortunate to have a complete global dataset, caution should be taken in model results and their probable high reliance on sample size and spatial location.

Limitations

The goal of this study was to assess whether or not the consideration of human influence in the form of land cover and LUI can improve predictions of SARs and if so, if it is comparable to that of environmental measures. Thus, we did not control for other factors at play in driving patterns of species richness, and models that include one LUI metric do not account for the many other potential land-use activities and environmental factors at play on the same landscape. Nor did we account for the collinearity inherent in our datasets where species richness, biomes, and agricultural suitability are closely tied to climate and topography. Furthermore, we do not know many species' tolerances to land-use change and even in the cases where tolerances to land-use are known, extent-of-occurrence range maps usually do not reflect such changes. In the knowledge that the SAR is affected by grain size, where different patterns emerge at different spatial grains, (Turner and Tjørve 2005), we chose $110 \times 110 \text{ km}^2$ grid cells as it is the minimum acceptable grain (Hurlbert and Jetz 2007). It is therefore expected that the relationships we found are scale-dependent and should not be extrapolated. Together, these issues present a challenge inherent in implying any form of causality between our predictor variables and our biodiversity distributions.

We compiled a set of land cover and LUI metrics with the highest quality currently available. Nevertheless, despite considerable recent progress, numerous gaps exist regarding the availability of alternative indicators and the difficulties in their measurement related to issues with data availability, accuracy and error propagation (Kuemmerle et al. 2013). Uncertainties in the accuracy of current LUI maps are often high due to inconsistent input data and limitations with processing algorithms and positional accuracy. Furthermore, there is a lack of formal validation for many of these datasets (Verburg et al. 2011). Systematically collected ground-based data only cover a few regions of the globe, statistical data are often only available at the national scale, and remote sensing cannot easily capture the subtle spectral effects of LUI changes (Kuemmerle et al. 2013). Furthermore, the fertilizer (Potter et al. 2010) and cereal yield (Monfreda et al. 2008) LUI maps used here all rely on one cropland map (Ramankutty et al. 2008), and inaccuracies in the base map can therefore propagate (Supplementary material Appendix 1, Table A1).

Conclusions

Human land-use has been shown to drive biodiversity loss at the local scale, however, its ability to predict variation in global SARs had not previously been assessed. This study adds evidence suggesting that human land use may be an important predictor of species richness. Great attention has previously been paid to the past and present biophysical

attributes at play in predicting patterns of species richness. Our findings suggest that human activity can better predict large-scale patterns of species richness than previously thought. This is useful information given that land-use is the most important driver of local biodiversity patterns, and that land-use change is expected to accelerate in the future, as human population and per-capita consumption soar. In order to better predict and understand biodiversity patterns using SARs, we need to adopt a more nuanced view, with both land-cover and the intensity of land-use taken as potentially important factors in explaining variation in global species richness.

Acknowledgements – LK and TK acknowledge funding by the Einstein Foundation Berlin (Germany). HK acknowledges funding from the German Research Foundation (DFG) through the German Excellence Initiative and CRC 990-EFForTS. CM acknowledges support by sDiv, the Synthesis Centre of the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig (DFG FZT 118). KG acknowledges funding by GLUES (Global Assessment of Land Use Dynamics, Greenhouse Gas Emissions and Ecosystem Services, BMBF). We are grateful for discussions on an earlier version with Ester Polaina, Lyndon Estes and David Wilcove's group. We appreciate two insightful reviews and editing work by Jeremy T. Kerr. This research contributes to the Global Land Project (<www.globallandproject.org>).

References

- Alkemada, R. et al. 2012. Assessing the impacts of livestock production on biodiversity in rangeland ecosystems. – *Proc. Natl Acad. Sci. USA* 110: 20900–20905.
- Arrhenius, O. 1921. Species and area. – *J. Ecol.* 9: 95–99.
- Belmaker, J. and Jetz, W. 2015. Relative roles of ecological and energetic constraints, diversification rates and region history on global species richness gradients. – *Ecol. Lett.* 18: 563–571.
- Bivand, R. S. et al. 2012. spdep: spatial dependence: weighting schemes, statistics and models. – R package ver. 0.5-53.
- Brooks, T. M. et al. 2002. Habitat loss and extinction in the hotspots of biodiversity. – *Conserv. Biol.* 16: 909–923.
- Channan, S. K. et al. 2014. Global mosaics of the standard MODIS land cover type data. – Univ. of Maryland and the Pacific Northwest National Laboratory, College Park, MD, USA
- Connor, E. F. and McCoy, E. D. 1979. The statistics and biology of the species–area relationship. – *Am. Nat.* 113: 791–833.
- Currie, D. J. 1991. Energy and large-scale patterns of animal- and plant-species richness. – *Am. Nat.* 137: 27–49.
- Cusens, J. et al. 2012. What is the form of the productivity–animal-species-richness relationship? A critical review and meta-analysis. – *Ecology* 93: 2241–2252.
- Dengler, J. 2009. Which function describes the species–area relationship best? A review and empirical evaluation. – *J. Biogeogr.* 36: 728–744.
- Di Marco, M. and Santini, L. 2015. Human pressures predict species' geographic range size better than biological traits. – *Global Change Biol.* 21: 2169–2178.
- Dormann, C. F. 2007. Effects of incorporating spatial autocorrelation into the analysis of species distribution data. – *Global Ecol. Biogeogr.* 16: 129–138.

- Dornelas, M. et al. 2014. Assemblage time series reveal biodiversity change but not systematic loss. – *Science* 344: 296–299.
- Drakare, S. et al. 2006. The imprint of the geographical, evolutionary and ecological context on species–area relationships. – *Ecol. Lett.* 9: 215–227.
- FAOSTAT 2010. FAOSTAT database. – <<http://faostat3.fao.org/home/E>>.
- Felton, A. et al. 2010. A meta-analysis of fauna and flora species richness and abundance in plantations and pasture lands. – *Biol. Conserv.* 143: 545–554.
- Field, R. et al. 2009. Spatial species–richness gradients across scales: a meta-analysis. – *J. Biogeogr.* 36: 132–147.
- Foley, J. A. et al. 2011. Solutions for a cultivated planet. – *Nature* 478: 337–342.
- Gaston, K. J. 2000. Global patterns in biodiversity. – *Nature* 405: 220–227.
- Geldmann, J. et al. 2014. Mapping change in human pressure globally on land and within protected areas. – *Conserv. Biol.* 1–13.
- Gerstner, K. et al. 2014. Accounting for geographical variation in species–area relationships improves the prediction of plant species richness at the global scale. – *J. Biogeogr.* 41: 261–273.
- Gibbs, K. E. et al. 2009. Human land use, agriculture, pesticides and losses of imperiled species. – *Divers. Distrib.* 15: 242–253.
- Gibson, L. et al. 2011. Primary forests are irreplaceable for sustaining tropical biodiversity. – *Nature* 478: 378–381.
- Haberl, H. et al. 2007. Quantifying and mapping the human appropriation of net primary production in earth's terrestrial ecosystems. – *Proc. Natl Acad. Sci. USA* 104: 12942–12947.
- Haberl, H. et al. 2014. Human appropriation of net primary production: patterns, trends, and planetary boundaries. – *Annu. Rev. Environ. Resour.* 39: 363–391.
- Harrell, F. 2001. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. – Springer.
- Hawkins, B. A. et al. 2003a. Energy, water, and broad-scale geographic patterns of species richness. – *Ecology* 84: 3105–3117.
- Hawkins, B. A. et al. 2003b. Productivity and history as predictors of the latitudinal diversity gradient of terrestrial birds. – *Ecology* 84: 1608–1623.
- Herzon, I. et al. 2008. Intensity of agricultural land-use and farmland birds in the Baltic States. – *Agric. Ecosyst. Environ.* 125: 93–100.
- Hortal, J. et al. 2012. Integrating biogeographical processes and local community assembly. – *J. Biogeogr.* 39: 627–628.
- 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. USA* 104: 13384–13389.
- IUCN 2013. IUCN Red List of threatened species. – Version 2013.2.
- Jetz, W. et al. 2008. Ecological correlates and conservation implications of overestimating species geographic ranges. – *Conserv. Biol.* 22: 110–119.
- Keohoe, L. et al. 2015. Global patterns of agricultural land-use intensity and vertebrate diversity. – *Divers. Distrib.* 21: 1308–1318.
- Keil, P. et al. 2015. On the decline of biodiversity due to area loss. – *Nat. Commun.* 6: 8837.
- Kerr, J. T. and Cihlar, J. 2004. Patterns and causes of species endangerment in Canada. – *Ecol. Appl.* 14: 743–753.
- Kissling, W. D. and Carl, G. 2008. Spatial autocorrelation and the selection of simultaneous autoregressive models. – *Global Ecol. Biogeogr.* 17: 59–71.
- Kleijn, D. et al. 2009. On the relationship between farmland biodiversity and land-use intensity in Europe. – *Proc. R. Soc. B* 276: 903–909.
- 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.
- Kreft, H. and Jetz, W. 2007. Global patterns and determinants of vascular plant diversity. – *Proc. Natl Acad. Sci. USA* 104: 5925–5930.
- Kuemmerle, T. et al. 2013. Challenges and opportunities in mapping land use intensity globally. – *Curr. Opin. Environ. Sustainability* 5: 484–493.
- Ladle, R. J. and Whittaker, R. J. 2011. Conservation biogeography. – John Wiley and Sons
- Lambin, E. F. et al. 2001. The causes of land-use and land-cover change: moving beyond the myths. – *Global Environ. Change-Human Policy Dimensions* 11: 261–269.
- Martins, I. S. et al. 2014. The unusual suspect: land use is a key predictor of biodiversity patterns in the Iberian Peninsula. – *Acta Oecol.* 61: 41–50.
- Matson, P. A. et al. 1997. Agricultural intensification and ecosystem properties. – *Science* 277: 504–509.
- Matthews, T. J. et al. 2015. On the form of species–area relationships in habitat islands and true islands. – *Global Ecol. Biogeogr.* 25: 847–858.
- Mendenhall, C. D. et al. 2014. Predicting biodiversity change and averting collapse in agricultural landscapes. – *Nature* 509: 213–217.
- Monfreda, C. et al. 2008. Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. – *Global Biogeochem. Cycles* 22: GB1022.
- Murray, B. R. and Dickman, C. R. 2000. Relationships between body size and geographical range size among Australian mammals: has human impact distorted macroecological patterns? – *Ecography* 23: 92–100.
- Newbold, T. et al. 2015. Global effects of land use on local terrestrial biodiversity. – *Nature* 520: 45–50.
- Olson, D. M. and Dinerstein, E. 1998. The global 200: a representation approach to conserving the Earth's most biologically valuable ecoregions. – *Conserv. Biol.* 12: 502–515.
- Pereira, H. M. and Daily, G. C. 2006. Modeling biodiversity dynamics in countryside landscapes. – *Ecology* 87: 1877–1885.
- Pereira, H. M. et al. 2012. Global biodiversity change: the bad, the good, and the unknown. – *Annu. Rev. Environ. Resour.* 37: 25–50.
- Pimm, S. L. et al. 1995. The future of biodiversity. – *Science* 269: 347–350.
- Potter, P. et al. 2010. Characterizing the spatial patterns of global fertilizer application and manure production. – *Earth Interactions* 14: 1–22.
- Ramankutty, N. et al. 2008. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. – *Global Biogeochem. Cycles* 22: GB1003.
- Rosenzweig, M. L. 1995. Species diversity in space and time. – Cambridge Univ. Press.
- Sala, O. E. et al. 2000. Global biodiversity scenarios for the year 2100. – *Science* 287: 1770–1774.
- Scheiner, S. M. 2003. Six types of species–area curves. – *Global Ecol. Biogeogr.* 12: 441–447.
- Schipper, J. et al. 2008. The status of the world's land and marine mammals: diversity, threat, and knowledge. – *Science* 322: 225–230.
- Siebert, S. et al. 2005. Development and validation of the global map of irrigation areas. – *Hydrol. Earth Syst. Sci.* 9: 535–547.

- Stein, A. et al. 2014. Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. – *Ecol. Lett.* 17: 866–880.
- Storch, D. et al. 2005. The species–area–energy relationship. – *Ecol. Lett.* 8: 487–492.
- Thomas, C. D. et al. 2004. Extinction risk from climate change. – *Nature* 427: 145–148.
- Tilman, D. et al. 2001. Forecasting agriculturally driven global environmental change. – *Science* 292: 281–284.
- Turner, W. R. and Tjørve, E. 2005. Scale-dependence in species–area relationships. – *Ecography* 28: 721–730.
- Verburg, P. et al. 2011. Challenges in using land use and land cover data for global change studies. – *Global Change Biol.* 17: 974–989.
- Wright, D. H. 1983. Species–energy theory: an extension of species–area theory. – *Oikos* 41: 496–506.
- Wright, D. H. 1990. Human impacts on energy flow through natural ecosystems, and implications for species endangerment. – *Ambio* 19: 189–194.
- Yamaguchi, T. and Blumwald, E. 2005. Developing salt-tolerant crop plants: challenges and opportunities. – *Trends Plant Sci.* 10: 615–620.

Supplementary material (Appendix ECOG-02508 at <www.ecography.org/appendix/ecog-02508>). Appendix 1.