LETTER

Species–area relationships always overestimate extinction rates from habitat loss

Fangliang He[1,2] & Stephen P. Hubbell[3,4]

Extinction from habitat loss is the signature conservation problem of the twenty-first century[1]. Despite its importance, estimating extinction rates is still highly uncertain because no proven direct methods or reliable data exist for verifying extinctions. The most widely used indirect method is to estimate extinction rates by reversing the species–area accumulation curve, extrapolating backwards to smaller areas to calculate expected species loss. Estimates of extinction rates based on this method are almost always much higher than those actually observed[2–5]. This discrepancy gave rise to the concept of an 'extinction debt', referring to species 'committed to extinction' owing to habitat loss and reduced population size but not yet extinct during a non-equilibrium period[6,7]. Here we show that the extinction debt as currently defined is largely a sampling artefact due to an unrecognized difference between the underlying sampling problems when constructing a species–area relationship (SAR) and when extrapolating species extinction from habitat loss. The key mathematical result is that the area required to remove the last individual of a species (extinction) is larger, almost always much larger, than the sample area needed to encounter the first individual of a species, irrespective of species distribution and spatial scale. We illustrate these results with data from a global network of large, mapped forest plots and ranges of passerine bird species in the continental USA; and we show that overestimation can be greater than 160%. Although we conclude that extinctions caused by habitat loss require greater loss of habitat than previously thought, our results must not lead to complacency about extinction due to habitat loss, which is a real and growing threat.

The Millennium Ecosystem Assessment[1] predicts that near-term extinction rates could be as high as 1,000 to 10,000 times background rates (see also ref. 7). Most predictions of species extinction rates, including those in the Millennium Ecosystem Assessment, are inferred from applying the SAR to rates of habitat loss[8–14]. The wide discrepancy between the rates of species extinction predicted by this method and the extinction rates actually recorded, has fuelled a continuing debate about how to explain the discrepancy[2,4,15–20]. The main issue is that, almost always, more species are left after a given loss of habitat than the number of species predicted to remain, based on the SAR. The most frequent interpretation is that the excess species are 'committed to extinction'. The term 'extinction debt' was coined to refer to species' populations that were no longer viable but were facing certain extinction due to habitat destruction that had already occurred[3,6,17]. The consensus on the most likely reason for the extinction debt is that there is a time lag for populations to go extinct after severe losses in population size[6,21].

Here we show that extinction rates estimated from the SAR are all overestimates. We define extinction rate as the fractional loss of species over a defined period accompanied by a given loss of habitat. These overestimates are due to the false assumption that the sampling problem for extinction is simply the reverse of the sampling problem for the SAR. The area that must be added to find the first individual of a species is in general much smaller than the area that must be removed to eliminate the last individual of a species (Fig. 1). Therefore, on average, it takes a much greater loss of area to cause the extinction of a species than it takes to add the species on first encounter, except in the degenerate case of a species having a single individual. We show mathematically that this is a necessary result of fundamental sampling differences between the SAR and the endemics–area relationship (EAR). Only in a very special and biologically unrealistic case, when all species are randomly and independently distributed in space, is it possible to derive the EAR from the SAR. Although this special case almost never occurs in nature, we examine this simple case first to clarify the nature of the problem. Then we relax these assumptions and consider the general case of aggregated species distributions.

The problem has gone unnoticed for so long because the traditional method for estimating extinction uses the power-law SAR, $S = cA^z$, which has no sampling theory relating it to species distributions (Supplementary Information A). To develop a sampling theory, we must consider the spatial distribution of species explicitly (Supplementary Information B and C). We derive the SAR and EAR from nearest-neighbour distances under two situations, random dispersion and clumped dispersion. We construct an SAR from the probability of encountering the first nearest neighbour of a species (a new species is added every time the sampling frame a encounters the first individual of the given species). In contrast, we construct the EAR from the probability of encountering the last neighbour of a species (a species is added only after all individuals are contained within frame a). We arrive at the species–area curve for randomly and independently distributed species as (Supplementary Information B):
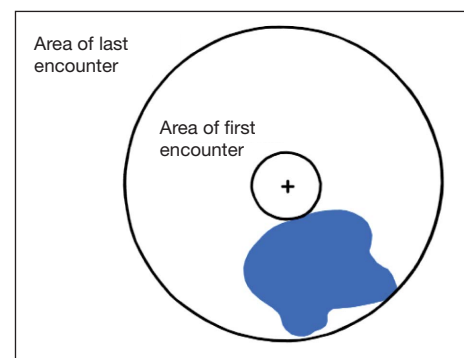
Figure 1 | Sampling differences for SAR and EAR. Range distribution of a species (blue area), and an arbitrary starting sample point, indicated by +. Regardless of the starting location, a sampling frame of arbitrary shape (here circular) with an area of a size sufficient to contact the species for the first time is always less than the sample area needed to encompass the entire range of the species. The SAR (species accumulation) is constructed from sample areas of first contact, and the EAR (species extinction) is constructed from areas of last contact.

[1]State Key Laboratory of Biocontrol and School of Life Sciences, Sun Yat-sen University, Guangzhou 510275, China. [2]Department of Renewable Resources, University of Alberta, Edmonton, Alberta, T6G 2H1, Canada. [3]Department of Ecology and Evolutionary Biology, University of California, Los Angeles, California 90095, USA. [4]Center for Tropical Forest Science, Smithsonian Tropical Research Institute, Unit 0948, APO AA 34002-0948, Republic of Panama.

$$S_a^1 = S - \sum_{i=1}^{S} \left(1 - \frac{a}{A}\right)^{N_i} \quad (1)$$

and the endemics–area curve as:

$$S_a^N = \sum_{i=1}^{S} \left(\frac{a}{A}\right)^{N_i} \quad (2)$$

where N_i is the total abundance of species i and S is the total number of species in the region A. Equations (1) and (2), derived from nearest-neighbour distances, are identical to the classical random placement models[22–25].

Let the total area be A and let a sub-area a be lost. For randomly and independently distributed species, we can calculate the expected number of species lost with a loss of area a from the SAR (equation (1)) as $S_{\text{loss}} = S - S_{A-a}$. This is identical to the EAR calculated directly from equation (2): $S_{\text{loss}} = S - S_{A-a}^1 = \sum_{i=1}^{S} \left(\frac{a}{A}\right)^{N_i} = S_a^N$. This proves that, for the special case of species distributed randomly in space, extinction rates estimated from the backward random placement SAR and from the forward random placement EAR are the same, and the SAR and EAR are mirror images (Fig. 2 and Supplementary Fig. 1). This case is true because, under random placement, the total area A is equal to the sum of the areas of encountering the first individual and the last individual of a species. From the probability models of the nearest-neighbour distance, the expected area needed to sample the first individual is $a^1 = A/(N+1)$, and the expected area for the last individual is $a^N = NA/(N+1)$ (Supplementary Information B). Thus $a^1 + a^N = A$. Note that $a^N > a^1$ is always true except when $N = 1$.

This mirror-image relationship only holds for randomly distributed species, however. Almost all species in nature are clumped, not randomly distributed[26]. For aggregated species, one can show that $a^1 + a^N < A$ with $a^N \geq a^1$ remaining true (Supplementary Information C and Supplementary Fig. 2). This leads to $S - S_{A-a}^1 \neq S_a^N$. The more spatially aggregated species distributions are, the stronger the inequality $a^N \geq a^1$ becomes. These results are completely general and explain the discrepancy between the backward SAR and forward EAR methods as well as why the backward SAR method systematically overestimates extinction rates.

These results apply to sample areas on any spatial scale. We can assess the magnitude of overestimation by the backward SAR method precisely in cases where we know the species composition and spatial location of each individual of each species or spatial range of each species. To illustrate this, we use spatially explicit data from eight large stem-mapped plots from a global forest dynamics network. We also perform the analysis on biogeographical spatial scales for passerine species in the continental USA (see Methods). The results show that the classic power-law SAR model, $S = cA^z$, and its corresponding EAR model (Supplementary Information A),

$$\lambda = S_{\text{loss}}/S_A = 1 - (1 - a/A)^z \quad (3)$$

are not mirror-image curves. In equation (3), S_{loss} is the number of species lost (endemic) to destroyed sub-area a. Because of the difference in sampling procedure of encountering species and losing species, the slopes z of the power-law model $S = cA^z$ and EAR (3) are not the same. The fit of the power-law SAR and EAR to species–area and endemics–area data respectively lead to two very different slopes
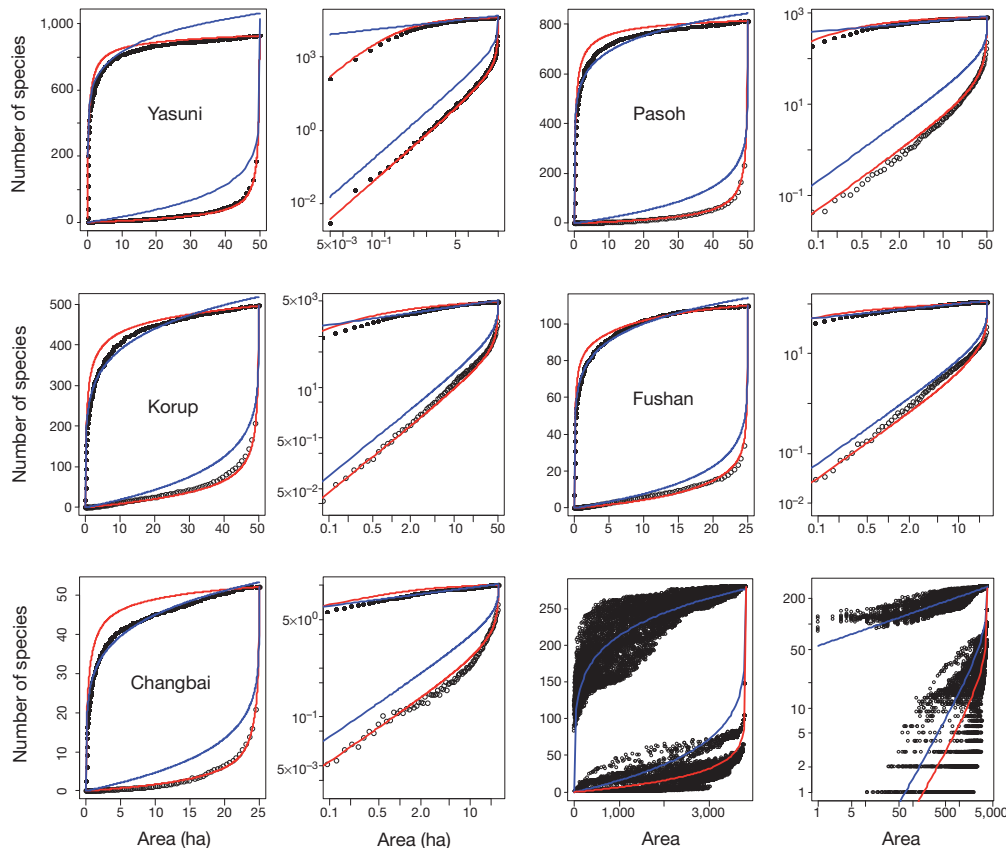
Figure 2 | Species– and endemics–area curves for six of the nine data sets in Table 1. The second and fourth columns are the plots on a log–log scale. The upper and lower blue curves are the fits of the power-law SAR and EAR (equation (3)), respectively. The upper and lower red curves are the predictions of the random placement SAR (equation (1)) and EAR (equation (2)), respectively. Unlike for the other data sets, the red curve for US passerine data

(cell size 0.48° latitude × 0.48° longitude) is the fit of equation (3) because the abundances of the passerine species are not known (so equation (2) cannot be used). The cloud of points represent 100 repeated random samples of the SAR and EAR. The SAR and EAR curves for the Barro Colorado Island plot are shown in Supplementary Fig. 1.

Table 1 | Eight stem-mapped forest plots across the world and distributions of passerine birds in the continental USA

Plot	Forest type	Size (ha)	Number of trees	Number of species	z_{SAR}	z_{EAR}	Bias$_{0.52}$ (%)	Bias$_{25}$ (%)
Barro Colorado Island, Panama	Lowland tropical forest	50	325,549	316	0.133 (0.00202)	0.0803 (0.000611)	65.61	64.38
Yasuni, Ecuador	Lowland tropical forest	50	307,279	1,128	0.126 (0.00473)	0.0623 (0.00189)	102.21	100.41
Pasoh, Malaysia	Lowland tropical forest	50	323,262	814	0.124 (0.00374)	0.0536 (0.00158)	131.30	129.02
Korup, Cameroon	Lowland tropical forest	50	328,973	496	0.179 (0.00369)	0.113 (0.00116)	58.38	56.92
Dinghu, China	Subtropical evergreen broad-leaved forest	20	71,617	210	0.274 (0.00180)	0.193 (0.000880)	41.94	40.34
Fushan, Taiwan	Subtropical evergreen broad-leaved forest	25	114,508	110	0.142 (0.00199)	0.0922 (0.000838)	53.99	52.92
Tiantong, China	Subtropical evergreen broad-leaved forest	20	94,603	152	0.200 (0.00214)	0.0994 (0.00175)	101.15	98.34
Changbai, China	Temperate forest	25	38,902	52	0.184 (0.00296)	0.0905 (0.00233)	103.27	100.62
USA	Passerine birds (0.24° × 0.24°)	14,904	–	279	0.187 (0.00101)	0.0766 (0.000516)	144.06	140.31
USA	Passerine birds (0.48° × 0.48°)	3,830	–	279	0.195 (0.00106)	0.0791 (0.000421)	147.39	143.39

The 'bias' is the overestimation calculated by comparing the extinction rates estimated from the z_{SAR} values with those from the endemic z_{EAR} values: $(\lambda_{SAR} - \lambda_{EAR})/\lambda_{EAR}$. We calculated percentage bias by assuming 0.52% and 25% habitat loss[1,29], respectively. Equation (3) gives λ_x. To analyse passerine distributions, we divided the lower 48 states of the USA into a grid of 14,904 cells with cell size of 0.24° latitude × 0.24° longitude and into 3,830 cells with cell size of 0.48° latitude × 0.48° longitude.

(the SAR z_{SAR} versus the EAR z_{EAR}) (Table 1). In some cases, z_{SAR} can be more than double z_{EAR}. This result is independent of the spatial scale of the data, as is evident for the passerine case shown in Table 1.

This analysis demonstrates that the most widely used method of estimating species extinction rates due to habitat loss, the backward SAR calculation, is not correct. For non-randomly distributed species, the SAR and EAR are not mirror images, so that one cannot be used to infer the other. This result holds regardless of how well the power-law SAR fits species–area data (Supplementary Information D). Even for randomly distributed species, the backward power-law SAR model is still not appropriate for estimating extinction rates because in this case equation (1) is the only correct SAR, not the power-law SAR (Supplementary Fig. 2), and equation (2) is the only correct EAR model. These results show that the concept of an 'extinction debt' (that is, the extinctions lost to biotic relaxation due to habitat destruction) based on the backward SAR model is not conceptually sound. Note that these results say nothing about whether an extinction debt exists, only that such a debt as might exist is not appropriately measured by the backward SAR method. To model the process of biotic relaxation will require a dynamic theoretical framework different from the current static SAR model. Currently, no such theory is available. The EAR curve is consistent with the concept of 'imminent extinction', which states that predictions of near-term extinctions due to habitat loss should focus on species endemic to the area of destroyed habitat[4,7,27,28].

Previous estimates of extremely high extinction rates, – for example, one species per hour to one species a day[8], 33–50% of all species between the 1970s and 2000 (ref. 9), from half to several million species by 2000 (refs 10, 12) or 50% of species by 2000 (ref. 11) – have not been observed. There is also reason to question the recent estimates of extinction rates made by the Millennium Ecosystem Assessment[1] and those by Thomas et al.[19]. In the latter case, the loss of habitat and the shift of species' ranges are driven by climate change. However, the use of the flawed backward SAR in Thomas et al. raises a legitimate question about the validity of their conclusion that 18–35% of species will be committed to extinction by 2050. We suggest that their estimated rates of extinction should be regarded as a high-end possibility rather than as supported by hard scientific evidence.

By how much have we overestimated extinction rates? Precise answers to this question require information about the EAR curve, which is generally not known. However, we can make a first approximation from the results shown in Table 1, for which we know the EAR curves in stem-mapped samples of forests and range distributions of passerines. We calculated the z_{EAR} and z_{SAR} averaged over the data in Table 1, leading to $z_{EAR} = 0.0940$ and $z_{SAR} = 0.174$. We then used two estimates of forest habitat loss, the annual deforestation rate of $(a/A)100\% = 0.52\%$ for humid tropical forests[29] and the estimated 25% conversion of original forest habitat into agricultural land[1]. The

SAR backward method (equation (3)) overestimates extinction rates by 85.07% and 83.00% in these respective cases, compared with the rates estimated by the forward EAR method. Conservation biologists often use a z value of 0.25 in cases where z values are not available[18,19]. Using this value inflates extinction rate estimates much more, being 165.85% and 160.10% for the two deforestation rates, respectively (see also Supplementary Fig. 1).

Are better methods available for estimating extinction rates? Our results show that the random placement EAR curve describes the empirical EAR curves for the forest plots very well. This result is remarkable and provides a simple method for estimating extinction (Supplementary Information E). Note that the theoretical random placement EAR for each plot is not data-fitting but a genuine prediction from equation (2).

These results might receive a mixed reaction from the conservation community. On the one hand, the good news is that all extinction rate estimates based on the backward SAR method are overestimates. Because it is derived from sample areas of first contact with each species, the backward SAR method makes the previously unrecognized assumption that any loss whatsoever of population due to habitat loss commits a species to extinction, which clearly is not true. On the other hand, there is likely to be concern that these results could jeopardize conservation efforts and be falsely construed in some quarters to imply that habitat loss is not a problem. Nothing could be further from the truth. There is no doubt whatsoever that the Millennium Ecosystem Assessment[1] has correctly identified habitat loss as the primary threat to conserving the Earth's biodiversity, and the sixth mass extinction might already be upon us or imminent[30]. Our results do indicate, however, that the backward SAR is not the correct way to estimate the magnitude of the current extinction event. To help mitigate contemporary extinctions and strengthen the science behind conservation planning, we need far better geographical data on endemism and species' distributions to improve forecasts of extinction rates[7]. Improving geographical databases on the distribution of biodiversity on Earth should be an urgent international priority.

METHODS SUMMARY

We analysed data from eight 20–50 ha (1 hectare (ha) = 10^4 m^2), stem-mapped plots of the Center for Tropical Forest Science global plot network to construct SAR and EAR curves (http://www.ctfs.si.edu/). These data sets are suitable because (1) our analysis is independent of spatial scale, (2) they are among the few data sets in which individuals are mapped on a landscape scale and (3) the EAR curve, which must be known, cannot be calculated from SAR curves (see text).

We obtained the SAR and EAR curves as follows: (1) grid the plot into cells of some minimum size (for example 5 m × 5 m); (2) count the number of species and the number of endemic species (species completely confined to the sample area) in each cell; (3) average the number of species per cell and the number of endemic species across all cells of a given size; and (4) construct species–area and endemics–area

curves by repeating steps 1–3, increasing cell size from 5 m × 5 m, 10 m × 5 m, 10 m × 10 m, etc. up to the entire plot.

We estimated z_{SAR} by nonlinear fit of the power-law SAR model to the observed species–area data. We limited fitting to areas of at least 0.2 ha because the power-law model is not considered applicable at small spatial scales (including them inflates z values and worsens overestimation). We estimated z_{EAR} by directly fitting equation (3) to the observed endemics–area data (see Table 1).

We analysed SAR and EAR curves for 279 passerine species in the lower 48 states of the USA using individual species' range maps from Natureserve (http://www.natureserve.org/getData/birdMaps.jsp). We divided the USA into grids at two respective cell sizes, 0.24° latitude × 0.24° longitude (14,904 cells) and 0.48° latitude × 0.48° longitude (3,830 cells), to confirm that our analysis is robust to scale change, as predicted by the analytical results. We computed SAR and EAR curves using presence–absence data following the above procedure.

1. Millennium Ecosystem Assessment. *Ecosystems and Human Well-Being: Biodiversity Synthesis* (World Resources Institute, 2005).
2. Simberloff, D. in *Tropical Deforestation and Species Extinction* (eds Whitmore, T. C. & Sayer, J. A.) 75–89 (Chapman & Hall, 1992).
3. May, R. M., Lawton, J. H. & Stork, N. E. in *Extinction Rates* (eds Lawton, J. H. & May, R. M.) 1–24 (Oxford Univ. Press, 1995).
4. Pimm, S. L. & Askins, R. A. Forest losses predict bird extinctions in eastern North America. *Proc. Natl Acad. Sci. USA* **92,** 9343–9347 (1995).
5. Rosenzweig, M. L. *Species Diversity in Space and Time* (Cambridge Univ. Press, 1995).
6. Tilman, D., May, R. M., Lehman, C. L. & Nowak, M. A. Habitat destruction and the extinction debt. *Nature* **371,** 365–366 (1994).
7. Pimm, S. L., Russell, G. J., Gittleman, J. L. & Brooks, T. M. The future of biodiversity. *Science* **269,** 347–350 (1995).
8. Myers, N. *The Sinking Ark: A New Look at the Problem of Disappearing Species* (Pergamon, 1979).
9. Lovejoy, T. E. in *The Global 2000 Report to the President: Entering the Twenty-First Century* (study director Barney, G. O.) 328–331 (Council on Environmental Quality, U.S. Government Printing Office, 1980).
10. National Research Council. *Research Priorities in Tropical Biology* (National Academy of Sciences, 1980).
11. Ehrlich, P. R. & Ehrlich, A. H. *Extinction: The Causes and Consequences of the Disappearance of Species* (Random House, 1981).
12. Myers, N. in *Tropical Rain Forest Ecosystems, Structure and Function* (ed. Golley, F. B.) 325–334 (Elsevier, 1983).
13. Lugo, A. E. in *Biodiversity* (ed. Wilson, E. O.) 58–70 (National Academy Press, 1988).
14. WCMC. *Global Biodiversity: Status of the Earth's Living Resources* (Chapman & Hall, 1992).
15. Heywood, V. H. & Stuart, S. N. in *Tropical Deforestation and Species Extinction* (ed. Whitmore, T. C. & Sayer, J. A.) 91–117 (Chapman & Hall, 1992).
16. Dial, R. & Budiansky, S. Extinction or miscalculation? *Nature* **370,** 104–105 (1994).
17. Heywood, V. H., Mace, G. M., May, R. M. & Stuart, S. N. Uncertainties in extinction rates. *Nature* **368,** 105 (1994).
18. Brooks, T. & Balmford, A. Atlantic forest extinctions. *Nature* **380,** 115 (1996).
19. Thomas, C. D. et al. Extinction risk from climate change. *Nature* **427,** 145–148 (2004).
20. Harte, J., Ostling, A., Green, J. & Kinzig, A. Biodiversity conservation: Climate change and extinction risk. *Nature* **430,** doi:10.1038/nature02718 (2004).
21. Brooks, T. M., Pimm, S. L. & Oyugi, J. O. Time lag between deforestation and bird extinction in tropical forest fragments. *Conserv. Biol.* **13,** 1140–1150 (1999).
22. Arrhenius, O. Species and area. *J. Ecol.* **9,** 95–99 (1921).
23. Coleman, B. D. Random placement and species-area relations. *Math. Biosci.* **54,** 191–215 (1981).
24. He, F. & Legendre, P. Species diversity patterns derived from species–area models. *Ecology* **83,** 1185–1198 (2002).
25. Green, J. & Ostling, A. Endemics–area relationships: the influence of species dominance and spatial aggregation. *Ecology* **84,** 3090–3097 (2003).
26. Condit, R. et al. Spatial patterns in the distribution of tropical tree species. *Science* **288,** 1414–1418 (2000).
27. Kinzig, A. P. & Harte, J. Implications of endemics–area relationships for estimates of species extinctions. *Ecology* **81,** 3305–3311 (2000).
28. Pimm, S. L. & Raven, P. Extinction by numbers. *Nature* **403,** 843–845 (2000).
29. Achard, F. et al. Determination of deforestation rates of the world's humid tropical forests. *Science* **297,** 999–1002 (2002).
30. Barnosky, A. D. et al. Has the Earth's sixth mass extinction already arrived? *Nature* **471,** 51–57 (2011).

Author Contributions Both authors designed the study, analysed and interpreted data, and wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to F.H. (fhe@mail.sysu.edu.cn; fhe@ualberta.ca) or S.P.H. (shubbell@eeb.ucla.edu).

BIODIVERSITY

Species loss revisited

Conservationists predict massive extinctions as a result of habitat loss. Habitat loss undoubtedly does drive extinctions, but dealing with an unmet assumption that underlies these predictions yields much lower estimates. SEE LETTER P.368

CARSTEN RAHBEK & ROBERT K. COLWELL

Scientists generally agree that Earth is facing a biodiversity crisis, losing species 100 to 1,000 times faster than the normal background rate of extinction[1] and resulting in the sixth period of mass extinction in Earth's history. On page 368 of this issue, He and Hubbell[2] provide a fresh perspective on predictions of the rate of this species loss.

Previous periods of mass extinction were driven by global changes in climate and in atmospheric chemistry, bolide impacts and volcanism[3]. This time, species extinction is a result of interaction and competition for resources with another species — humans. We are immensely successful. Our numbers are many times higher than ecological theory would predict for a species with our life history and body mass. We explore, populate and drastically alter almost all corners of the Earth and modify the global climate. Loss of habitat is predicted by various studies to cause the extinction of 20–50% of all species in just half a century[3]. These estimates began to surface decades ago, but sceptics have repeatedly demanded evidence of widespread extinction, asking 'Where are the bodies?'. If proof is not forthcoming, they argue, then politicians and decision-makers should denounce the biodiversity crisis as a myth[4].

He and Hubbell[2] question the way that extinction rates attributed to habitat loss have most often been estimated. Biologists have struggled for decades to estimate how many species are going extinct. Traditionally, the answer has relied on estimates based on an almost universal ecological relationship — when we inventory the species in an area of natural habitat, the list grows as the area is increased. Using theoretical or empirically derived functions to describe this species–area relationship (SAR), it has been assumed that, by working backwards along the SAR, one can estimate the number of species that would be lost to extinction if a larger area were reduced by habitat loss.

A classic rule of thumb says that if habitat area is reduced by 90% (comparable to actual habitat loss in many regions), roughly one-half of its species will be lost. He and Hubbell cite studies using SAR that predicted the loss of 50% of all species by the year 2000 — predictions
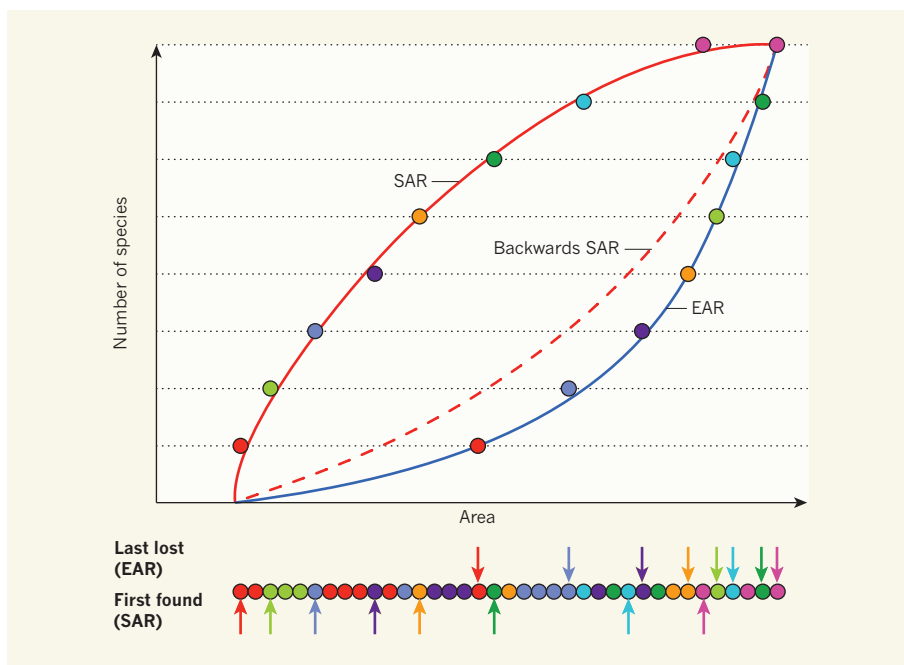
Figure 1 | Estimating species extinctions due to habitat loss. This hypothetical example shows the contrast between use of the backwards species–area relationship (SAR), traditionally used to predict extinctions, and the true endemics–area relationship (EAR) that correctly estimates extinctions with increasing area lost. The coloured circles under the graph represent the spatial ordering of 37 individuals (each occupying one unit of area) of 8 species along a transect through a habitat, each species indicated by a different colour. The total area surveyed increases with each individual encountered. As the first individual of each species is found, the SAR rises by one species, whereas the EAR is incremented only when the last individual of a species is accounted for along the transect. The backwards SAR mirrors the loss of species as area is reduced by moving right-to-left along the SAR. He and Hubbell[2] demonstrate mathematically and with examples for trees and birds that, for realistic (aggregated) spatial patterns of individuals and species, the backwards SAR always lies above the true EAR, thus overestimating expected rates of extinction. Species aggregation is simulated here by placing dots of the same colour closer to one another than expected at random.

that clearly have not been fulfilled. The discrepancy is well known and has often been explained as 'extinction debt', a time-lag before populations reduced in numbers by habitat loss actually become extinct. Individuals of long-lived species may continue to reproduce or simply live on without reproducing, even if the current living space for the species cannot sustain viable populations over time.

The authors[2] explain why this traditional 'backwards' use of SAR is fundamentally flawed for typical spatial diversity patterns, and show that this approach can produce drastic overestimation of extinction rates.

The problem with the traditional approach

is surprisingly simple. With increasing habitat area, the SAR rises by one species unit each time the first individual of a species new to the inventory is encountered (Fig. 1). Additional individuals of a species already encountered add nothing to the species count. By contrast, with decreasing habitat area an extinction does not occur until the last individual of a species is encountered. The authors show that, for the aggregated spatial patterns characteristic of species in real communities, the predicted number of extinctions rises more gradually with increasing habitat loss than predicted by the 'backwards' SAR (Fig. 1). The curve that correctly describes the rate of extinction as habitat

area decreases is called the endemics–area relationship (EAR). This was proposed more than a decade ago by Harte and Kinzig[5] and, they persuasively argued[6], is more appropriate than the SAR for estimating species extinctions, especially under non-random spatial distributions[7]. (A species is endemic if it is found only within some specified area.)

In their novel conceptualization of the problem, He and Hubbell[2] show that both the classic SAR and the EAR can be derived from a sampling theory based on spatially explicit patterns of individuals. Applying this approach to empirical data for woody plants in the rainforest and North American birds, which show typical patterns of spatial aggregation, they quantify the substantial discrepancy between backwards-SAR-based and EAR-based extinction rate predictions (finding overestimation as high as 160% for the plants). Importantly, the authors also justify the use of a simple approximation for the EAR that is robust to variation in species' spatial patterns and scale.

He and Hubbell, then, strongly question the use of SAR to estimate extinction rates not only from direct habitat loss, but also from projected species-range contractions expected to occur under climate change (see ref. 8 for an example). But they emphasize that their results do not in any way imply that there is not an ongoing mass extinction of species, nor that extinction debt is not a genuine biological phenomenon. Even with a better way to estimate rates of future species extinctions, there is still a need to obtain the data required to use the EAR to make more rigorous estimates. There is also the daunting problem of rigorously inferring extinction — showing that the last individual of a species has indeed died.

We invest heavily in infrastructure to store and make accessible the data we have, but by and large we have all but halted investment in discovering and describing the diversity of species with which we share the Earth. At best we have described only about 10% of all living multicellular species. If we 'fog' a tropical tree, literally hundreds of insect species unknown to science fall to the ground. Every year, many new species of even the best-known groups, the mammals and birds, are described. For only a fraction of the known species do we have even a rough idea of their entire geographical distributions.

Most of Earth's biodiversity occurs in tropical regions where species occur at low density and tend to have tiny geographical ranges. The first individual of such a species encountered in a brief inventory is not far from the last to go when extinction threatens, compared with populous, widespread species at higher latitudes. Thus, when modifying tropical habitat through forestry, mining or agriculture, we rarely have an idea which species inhabit the environment we are about to affect, nor the exact consequences of our action. The 'body bags' are rarely counted. ∎

Carsten Rahbek *is at the Center for Macroecology, Climate and Evolution, Department of Biology, University of Copenhagen, DK-2100 Copenhagen, Denmark.* Robert K. Colwell *is in the Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, Connecticut 06269-3043, USA.*
e-mails: crahbek@bio.ku.dk;
robert.colwell@uconn.edu

1. Pimm, S. L., Russell, G. J., Gittleman, J. L. & Brooks, T. M. *Science* **269,** 347–350 (1995).
2. He, F. & Hubbell, S. P. *Nature* **473,** 368–371 (2011).
3. Barnosky, A. D. *et al. Nature* **471,** 51–57 (2011).
4. Lomborg, B. *The Skeptical Environmentalist: Measuring the Real State of the World* (Cambridge Univ. Press, 2001).
5. Harte, J. & Kinzig, A. P. *Oikos* **80,** 417–427 (1997).
6. Kinzig, A. P. & Harte, J. *Ecology* **81,** 3305–3311 (2000).
7. Green, J. L. & Ostling, A. *Ecology* **84,** 3090–3097 (2003).
8. Thomas, C. D. *et al. Nature* **427,** 145–148 (2004).

ASTRONOMY

Bound and unbound planets abound

Two teams searching for extrasolar planets have jointly discovered a new population of objects: ten Jupiter–mass planets far from their host stars, or perhaps even floating freely through the Milky Way. SEE LETTER P.349

JOACHIM WAMBSGANSS

Two decades ago, we had no idea whether planets orbiting stars other than the Sun existed at all. Today, more than 500 exoplanets have been discovered, and the field of exoplanet research has advanced to become one of the most captivating branches of astronomy. Observational techniques now aim to address questions such as what the atmosphere and weather are like on some of these planets, and to determine their global statistical properties. On page 349 of this issue, the MOA and OGLE research teams[1] provide an exciting result for exoplanetary science: the discovery of a population of planets that have roughly the mass of Jupiter and separations from their putative host stars of at least ten times Earth's distance to the Sun.

The teams' finding[1] is based on gravitational microlensing, an established technique for detecting exoplanets that is well placed for statistical studies of exoplanets. There are two particularly exciting aspects to the discovery of this new exoplanetary population. The first is the authors' conclusion that, on average, there is more than one Jupiter-mass planet per Milky Way star. The second is the evidence that these planetary-mass objects could be at great distances from their host stars. Some of them could even be floating freely through the Milky Way — that is, they might not be gravitationally bound to any star at all.

Gravitational microlensing is one of a suite of planet-search techniques. The methods are truly complementary to one another, each probing different planetary properties and having its own particular strengths[2]. But most of them detect and explore nearby exoplanets. By contrast, microlensing probes more distant planets, using the host star–planet system as a magnifying glass. When a foreground star (the lens) passes in front of a distant, background star, the latter is magnified and displays a characteristic 'light curve'[3]. The two observables that characterize such a microlensing event are the height of the light curve's magnification peak and the duration of the magnification, which depends, among other parameters, on the mass of the lens: the lower the mass, the shorter the duration. Originally proposed as a way of searching for dark matter, it soon became clear that microlensing could also be used to detect planetary systems[4]: a planet orbiting the foreground star would produce a secondary peak in the light curve (Fig. 1).

Microlensing offers two advantages over other methods: it has the potential to yield the most representative statistical sample of Milky Way planets and it is, in principle, sensitive enough to detect Earth-mass objects[5,6] with current technology. However, the downside is that microlensing events are rare: fewer than one in a million stars in the central part of the Milky Way are microlensed at any given time by a foreground lensing star. And even if every such lensing star had a Jupiter-mass planet at a few times the Earth–Sun distance, only about 1% of these planets would be detected, owing to the exact geometric alignment required between the background star, the planetary system and an observer on Earth. So discovering such microlensing events is akin to finding a needle in a haystack.

To tackle these statistical challenges, a handful of independent research teams have developed advanced techniques to monitor the brightness of about 100 million Milky Way stars every few days. These techniques have allowed the teams to routinely find about