



CrossMark  
click for updates

## Research

**Cite this article:** Amano T, Sutherland WJ. 2013 Four barriers to the global understanding of biodiversity conservation: wealth, language, geographical location and security. *Proc R Soc B* 280: 20122649. <http://dx.doi.org/10.1098/rspb.2012.2649>

Received: 8 November 2012

Accepted: 15 January 2013

### Subject Areas:

ecology, environmental science

### Keywords:

biodiversity information, geographical locations, global biodiversity databases, languages, scientific communication, security

### Author for correspondence:

Tatsuya Amano

e-mail: [amatatsu830@gmail.com](mailto:amatatsu830@gmail.com)

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspb.2012.2649> or via <http://rspb.royalsocietypublishing.org>.

# Four barriers to the global understanding of biodiversity conservation: wealth, language, geographical location and security

Tatsuya Amano and William J. Sutherland

Conservation Science Group, Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK

Global biodiversity conservation is seriously challenged by gaps and heterogeneity in the geographical coverage of existing information. Nevertheless, the key barriers to the collection and compilation of biodiversity information at a global scale have yet to be identified. We show that wealth, language, geographical location and security each play an important role in explaining spatial variations in data availability in four different types of biodiversity databases. The number of records per square kilometre is high in countries with high per capita gross domestic product (GDP), high proportion of English speakers and high security levels, and those located close to the country hosting the database; but these are not necessarily countries with high biodiversity. These factors are considered to affect data availability by impeding either the activities of scientific research or active international communications. Our results demonstrate that efforts to solve environmental problems at a global scale will gain significantly by focusing scientific education, communication, research and collaboration in low-GDP countries with fewer English speakers and located far from Western countries that host the global databases; countries that have experienced conflict may also benefit. Findings of this study may be broadly applicable to other fields that require the compilation of scientific knowledge at a global level.

## 1. Introduction

The world now faces a global biodiversity crisis and the consequent loss of ecosystem services. Understanding and predicting the fate of global biodiversity thus represents an urgent task for conservation scientists, practitioners and policy-makers. Global assessments of the distributions and changes in biodiversity in spatial and temporal contexts have played a crucial role in efficiently allocating limited resources to high-priority areas and species [1–5]. However, the limited amount of available information poses one of the toughest challenges for such global efforts [6,7]. For example, 18 per cent of animal species that have been evaluated to date by the IUCN have been categorized as data-deficient, because the information available for the assessment of population status is so limited [1]. Gaps and heterogeneity in the geographical and taxonomical coverage of existing information on biodiversity have also been recognized as critical problems in other global efforts to assess the status of global biodiversity [5,8,9]. A recent study demonstrated that existing biases in available information can lead to inaccurate inferences in ecological studies [10].

Earlier studies have shown that the availability of information is unevenly distributed across the globe [11], highlighting in particular the lack of data in species-rich tropics [8,12–15]. Also, the wealth of a country has been revealed to be positively associated with data availability [16,17]. However, there must be smaller-scale variations in data availability as well as the reported broad-scale contrast between the tropics and non-tropics. More importantly, many other factors, in addition to wealth, have been suggested as potential barriers

to the collection of biodiversity information. From poorer countries, it is suggested that problems include lack of adequate infrastructure, insufficient expertise, inaccessibility to research sites due to political upheaval, and difficulties in getting data published or made public [14]. Studies so far have rarely identified the multiple barriers to, and quantified their relative impact on, the collection and compilation of biodiversity information at a global scale. However, these steps are crucial for tackling efficiently the spatial biases in the available information that the conservation community now faces, and consequently obtaining an undistorted understanding of global biodiversity status.

Here, focusing on four potential barriers to the collection and compilation of information for global biodiversity conservation—wealth, language, geographical location and security—we aimed to quantify the contribution of each factor in explaining spatial variations in the amount of available data in ecological/conservation databases at global and continental scales. As well as the known role of country wealth [16,17], we also hypothesized that language, geographical location and security of a country could have an impact in explaining spatial biases in available information. English is now the common medium of international scientific communication [18], while geographical distance can be a barrier to face-to-face communication among scientists. Thus, both factors can affect the effective collection and compilation of biodiversity information. A low level of security can discourage scientific activities [19], potentially affecting the amount of scientific data available in the country. To test this hypothesis, this study analysed four databases that can potentially provide information that is essential for global biodiversity conservation: the Global Biodiversity Information Facility (GBIF; [www.gbif.org](http://www.gbif.org)), which collects records on the occurrence of organisms across the globe (over three hundred million records); the Global Population Dynamics Database (GPDD), the largest collection of time-series population data in the world (nearly 5000 records) [20], which was also used to assess the progress towards the 2010 Biodiversity Target as a part of the Living Planet Index [5]; MoveBank, a global data archive for animal movement data (records on 16 251 tagged animals from 405 studies) [21]; and the European Union for Bird Ringing Databank (EDB), which has compiled bird ringing recovery data through bird ringing schemes throughout Europe (about 4 million live and 1.4 million dead recoveries) [22]. These databases were chosen so as to cover different types of information needed for biodiversity conservation (distribution, population dynamics, behaviour and demographic parameters), and are not intended to be exhaustive, but rather indicative of each type of information (see more detail in §2).

Considering potential correlations among the four factors, a hierarchical partitioning [23] was performed to quantify how much variation in the number of records per square kilometre among countries can be explained by the independent and joint contributions of gross domestic product (GDP) per capita, the proportion of English speakers to the national population, the geographical distance (kilometres) from countries of the host organizations of the databases and the level of security measured by the Global Peace Index (GPI) [24]. The number of records per square kilometre in each country was also compared with bird species richness, after accounting for land area, as an index of biodiversity.

## 2. Material and methods

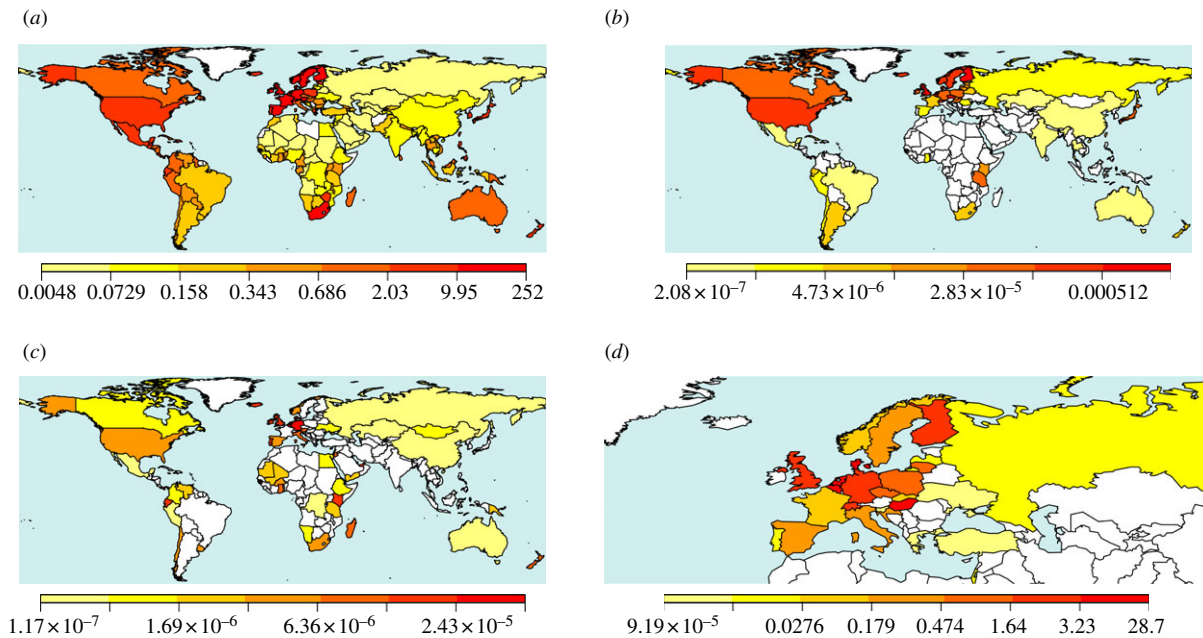
We first tested the relationship between the number of records per square kilometre in each country and the four explanatory variables: GDP per capita, the proportion of English speakers to the national population, the geographical distance from countries where the host organizations of the databases are based, and the GPI. The number of records per square kilometre, not the total number of records in each country, was used on the assumption that, all else being equal, more records should be collected in larger countries for effective conservation across the country. More detailed measures of wealth, such as the amount of budget spent on conservation science, might be a better predictor of data availability, but we could not obtain such information at a global scale and thus used the GDP per capita instead. The GPI was used to quantify a nation's level of security, which can affect accessibility to (and thus the amount of) data collected from the country [14]. The GPI comprises 23 indicators, which gauge three broad themes: the level of safety and security in society (10 indicators), the extent of domestic or international conflict (five), and the degree of militarization (eight), with data collated by the Economist Intelligence Unit [24]. Lower GPI scores indicate higher 'peacefulness'. While each of the 23 indicators could be used in the analysis, including all of them would greatly increase the number of models to compare, making the analysis unnecessarily complex. We decided to use the GPI because, to our knowledge, it is the only index that comprehensively quantifies a nation's level of security.

Second, the number of records per square kilometre in each country was also compared with bird species richness, after accounting for land area. Species richness in one taxon may not necessarily represent overall biodiversity, but bird species richness was highly correlated with both mammal (Kendall's  $\tau$  0.801) and amphibian (0.680) species richness (both derived from [1]). Thus, bird species richness was used as an index of biodiversity.

### (a) Data

The four databases (GBIF, GPDD, MoveBank and EDB) were not intended to be exhaustive, and there are clearly other conservation/ecological databases at global and continental scales, such as the IUCN Red List dataset [1], the World Wildlife Fund WildFinder [25], the ASEAN biodiversity information sharing system [26], and the International Legume Database and Information Service World Database of Legumes [27]. However, all of these databases only provide the distribution range of species by countries or regions, not actual observation records, making it impossible to assess spatial biases in available information within each range. Such information on species' distribution is represented by GBIF for the purpose of this study. To our knowledge, GPDD and MoveBank are the largest databases that are open to the public and provide information on population dynamics and behaviour, both of which are crucial in biodiversity conservation [5,28]. The EDB is a database at a European scale, but we could not find global databases that provide information on demographic parameters by country, and thus decided to use EDB in this study. Consequently, the amount of data stored is very different among the four databases; GBIF stores far more data than the other three. We do not consider that the unevenness of data availability among the selected databases is a drawback of this study, but rather that it reflects the information bias among different types of information needed for biodiversity conservation.

We derived the total number of records in each country from the websites of the four databases. The sum of live and dead recoveries was used for EDB. We also collected the information on land area, population and GDP of each country from the World Factbook [29]. GDP per capita was calculated by dividing GDP by the national population. The number of English speakers



**Figure 1.** Maps showing the spatial variations in the number of records per square kilometre in four conservation/ecological databases: (a) the GBIF, (b) the GPDD, (c) MoveBank and (d) the EDB. Countries without any records are shown in white.

in each country was estimated using four different sources: Ethnologue [30], the World Factbook [29], the *Cambridge encyclopedia of Language* [31] and the Eurobarometer survey [32]. The total number of speakers of English as the first or second language was derived from the first three sources, either directly as absolute numbers or as a proportion of the national population, which was multiplied by the national population listed on the World Factbook to estimate actual numbers. For the Eurobarometer survey, the number of English speakers was estimated by multiplying the proportion of people who are aged 15 years or older and answered that they can use English well enough for conversations by the total population aged 15 years or older (i.e. the target population for the survey). The maximum value in the four databases was used for the analysis. The distance (kilometres) from the country in which the organization responsible for the database resides (Denmark for GBIF, UK for GPDD and EDB and Germany for MoveBank) was calculated using the package `cshapes` [33] in the program R v. 2.13.0 [34]. Here, the distance between the capitals was used as that between two countries. The GPI was derived from a report by the Institute for Economics and Peace [24]. The number of bird species in each country was derived from BirdLife International's World Bird Database ([www.birdlife.org/datazone](http://www.birdlife.org/datazone), accessed January 2012). Since countries vary considerably in size, we used bird species richness controlled for area. It is well known that species richness increases with area nonlinearly, and thus it is inappropriate to simply divide richness values by area because this implicitly assumes a linear relationship. Instead, bird richness was divided by  $A^z$ , where  $A$  is a country's area and  $z$  is the exponent of the species-area curve [35]. Here,  $z$  was estimated by fitting the species-area curve to the data (see the electronic supplementary material, figure S1).

### (b) Statistical analysis

For each database, the model selection procedure based on the Akaike information criterion with a correction for small sample sizes (AICc) was first performed using simple linear regressions with the log-transformed number of records per square kilometre in each country as the response variable, and the log-transformed GDP per capita, the proportion of English speakers, the distance from the host organization and the GPI as explanatory variables. The sample size was 102, 34, 30 and

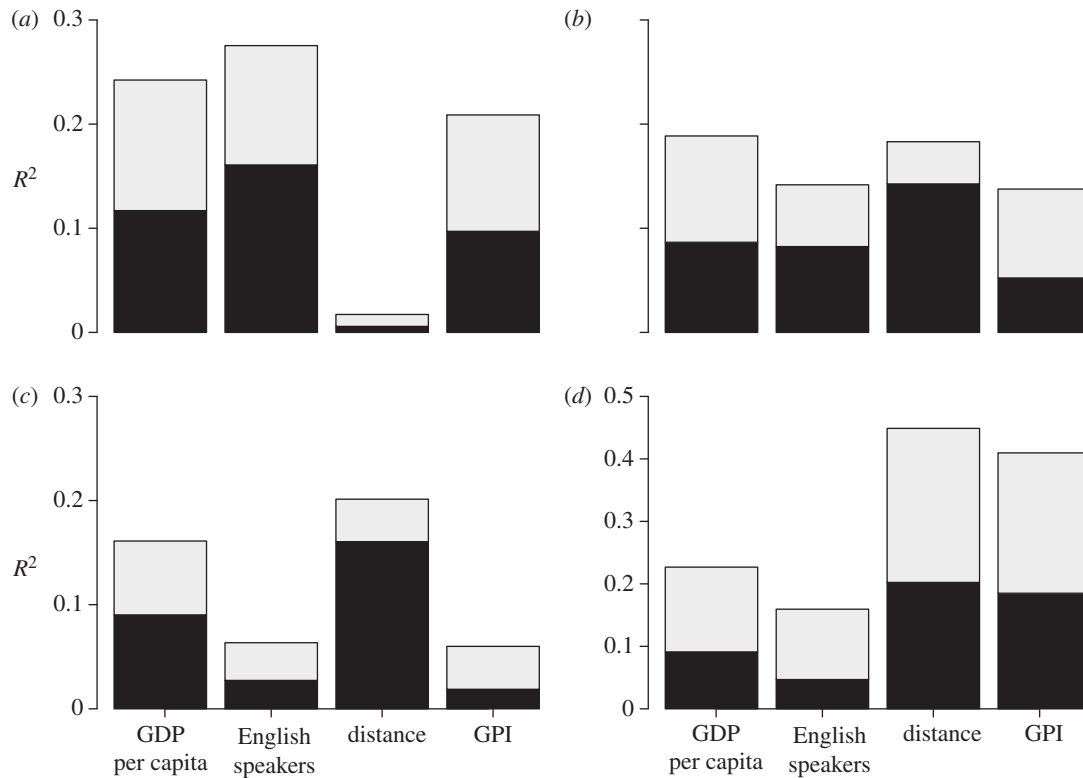
24 in GBIF, GPDD, MoveBank and EDB, respectively. To investigate the effect of spatial autocorrelation, Moran's  $I$  was calculated for the residuals from the full models, using the package `ncf` [36] in R. The calculated Moran's  $I$  was not significant up to the first 3000 km in all the databases, indicating no more than a weak autocorrelation. Thus, spatial autocorrelation was not considered explicitly in the model for the analysis.

A hierarchical partitioning [23] was performed to estimate the independent and joint explanatory capacities of each of the explanatory variables separately, using the package `hier.part` [37] in R. The process of a hierarchical partitioning involves computation of the increase in the fit (measured in our case as  $R^2$ ) of all models with a particular variable compared with the equivalent model without that variable [23]. As a result, a hierarchical partitioning provides, for each explanatory variable separately, an estimate of the independent and conjoint contribution with all other variables. This approach has been successfully used to deal with multicollinearity among explanatory variables in ecological data [38,39].

## 3. Results

Records in all four databases were distributed unequally (figure 1). The four databases, although targeted to different types of biodiversity information, showed similar patterns in the distribution of data. Most notably, records were concentrated in western and northern Europe, and North America (figure 1). Even in other regions, however, some countries tended to show high densities for most of the databases, such as Panama, Ecuador, South Africa, Kenya, Tanzania, Ghana, Israel and New Zealand (figure 1). In Asia, both Japan and Taiwan provided relatively many records to GBIF and GPDD, but not to MoveBank (figure 1). There was a considerable variation in information availability even within each region, as well reflected in EDB, where countries such as the Netherlands, Hungary, Denmark and Belgium provided more data than others (figure 1d).

Correlations among the explanatory variables were not necessarily high ( $|\text{Kendall's } \tau| < 0.4$  for all combinations), but



**Figure 2.** The independent (dark grey) and joint (light grey) contributions (given as  $R^2$  values) of GDP per capita, the proportion of English speakers (English speakers), the distance from host organizations (distance) and the GPI for the number of records per square kilometre in (a) the GBIF, (b) the GPDD, (c) MoveBank and (d) the EDB, as estimated from hierarchical partitioning.

were still significant in most of the combinations (see the electronic supplementary material, table S1). Although the variance inflation factor calculated for the full models (see the electronic supplementary material, table S2) showed relatively low multicollinearity, these intercorrelations among predictors potentially require the use of hierarchical partitioning [23] to quantify the independent and joint explanatory power of each predictor on the response variable (see §2 for more details).

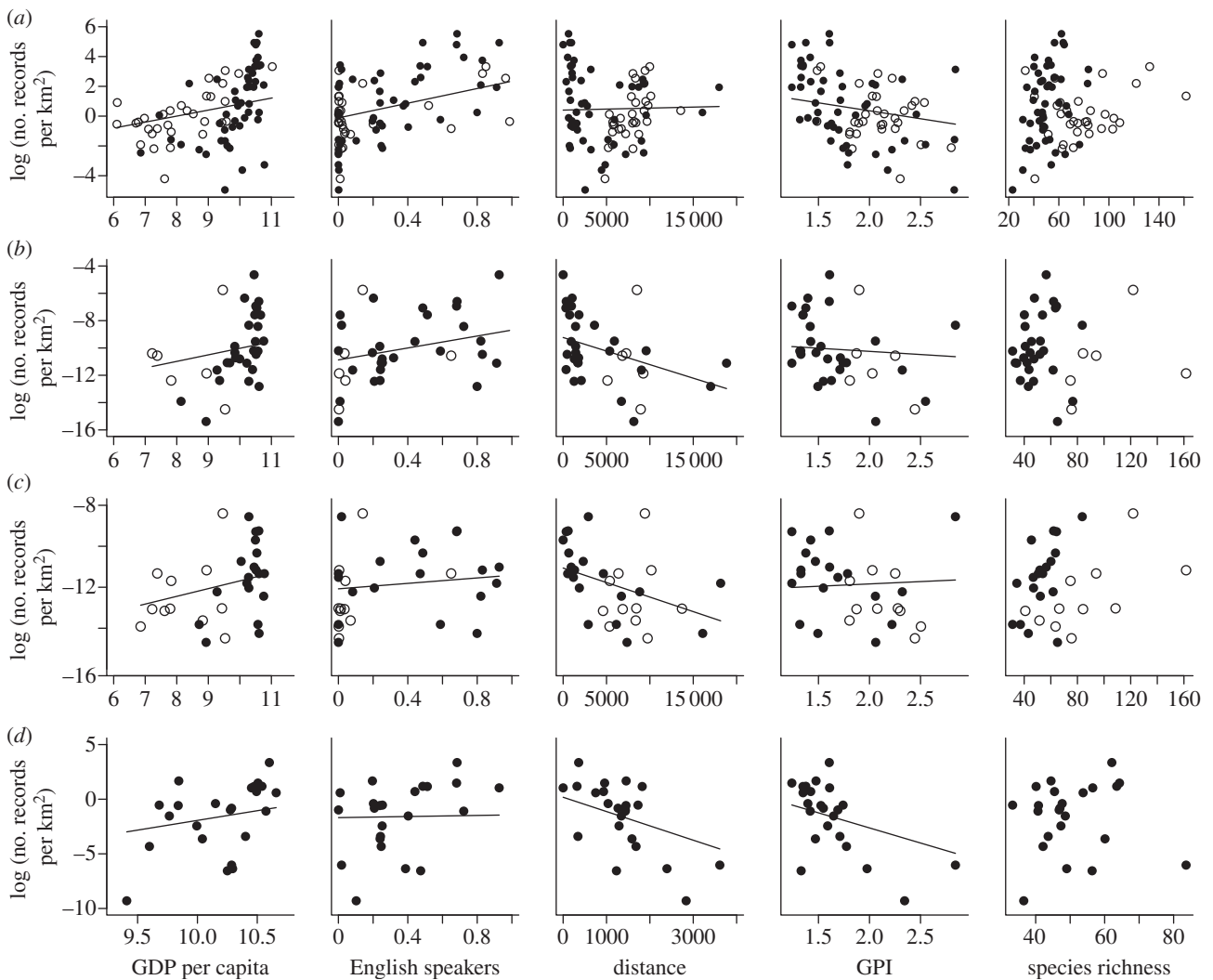
Wealth, language, geographical location and security all played an important role in explaining among-country variations in the amount of data, with varying importance among the four databases. Most of the four factors were included in at least one of the models with  $\Delta_i$  below 2.0 (see the electronic supplementary material, tables S3–S6), which provides substantial evidence that these models are the best models [40]. Exceptions were the distance from the host organization in GBIF, GPI in GPDD and MoveBank, and the proportion of English speakers in EDB (see the electronic supplementary material, tables S3–S6). A hierarchical partitioning revealed that each of the four predictors—GDP per capita, the proportion of English speakers, the distance from host organizations and GPI—explained up to 12–20 per cent independently, and 24–45 per cent jointly with other variables, of among-country variations in the number of records per square kilometre (figure 2). However, the explanatory power of each variable differed depending on databases. For example, the distance from host organizations in the GBIF and the proportion of English speakers and GPI in MoveBank explained less than 10 per cent of the variations even jointly with other variables (figure 2). Both GDP per capita and the proportion of English speakers were positively correlated with the number of records per square kilometre, while the distance from host organizations and the GPI generally

showed a negative effect (figure 3). Note here that high GPI scores indicate low levels of security. These four factors seemed to explain not only the well-known tropical–temperate gradient in data availability (white versus black circles in figure 3) but also within-region variations. Finally, the number of records per square kilometre did not show a significant correlation with bird species richness controlled for area, an index of biodiversity in each country—Kendall's  $\tau$  ( $p$ -values): 0.044 (0.562), 0.045 (0.724), 0.195 (0.135) and 0.091 (0.565) in GBIF, GPDD, MoveBank and EDB, respectively (figure 3).

## 4. Discussion

The main findings of this study are twofold. First, records in all four databases were distributed unequally across the globe; there were fine-scale spatial biases as well as the reported broad-scale contrast between the tropics and non-tropics. Second, the results revealed that language, geographical location and security, relative to wealth, all explained similar or even larger proportions of among-country variations in available data in global/continental databases. These four factors seemed to explain not only the well-known tropical–temperate gradient in data availability but also within-region variations. It seems fair to say that this is a novel finding, because earlier studies have focused almost entirely on the wealth of a country as a driver of spatial biases in available biodiversity information [16,17], and never quantified the relative importance of multiple drivers.

Although GDP presumably reflects various aspects of a nation, it is most likely to represent the wealth of a country. Wealth inevitably affects the budget for education and science, which is one of the potential determinants of the



**Figure 3.** Scatterplots showing the relationships between GDP per capita, the proportion of English speakers (English speakers), the distance from host organizations (distance), the GPI and bird species richness that was controlled for area (species richness), and the number of records per square kilometre in each country (black circles, non-tropics; white circles, tropics) in (a) the GBIF, (b) the GPDD, (c) MoveBank and (d) the EDB. Regression lines were based on the estimated coefficients in the full model using the actual range of values for the explanatory variable of interest, and mean values for the other variables. Note here that high GPI scores indicate low levels of security.

quantity [41] and quality [42] of scientific communities, consequently regulating infrastructure and expertise for collecting data and getting data made public. Thus, wealth and associated factors seem to largely explain the macro-scale contrast in available biodiversity information between tropical and temperate regions, which have repeatedly been reported by earlier studies [8,12–15].

Nevertheless, GDP per capita still left 76–84 per cent of among-country variation unexplained, a considerable proportion of which was explained by the other three factors: language, geographical location and security. The positive association of the proportion of English speakers with data availability in global datasets can be created through two possible processes: a delay in the development of biodiversity science in countries with fewer English speakers, and failure to collect and compile existing information in such countries. The proportion of English speakers reflects, to some degree, the level of globalization of science in each country, since English is now the common medium of international scientific communication [18]. Low English skills can cause severe intellectual isolation [18], potentially delaying the progress of scientific research, such as the establishment of systematic monitoring surveys on biodiversity. Furthermore, even in

countries with well-developed biodiversity sciences, low English skills could impede active communication with other countries, and the impact is particularly detrimental when communication with countries with strong scientific influence, such as the USA and the UK, is blocked, resulting in missed opportunities to contribute, for instance, to global scientific databases. Note that this is not to say that the information has not been collated in local databases (see below an example in Japan). Contributing to global data compilation may not necessarily be essential for biodiversity conservation at national or smaller scales, but is crucial for the effective understanding and conservation of global biodiversity.

The negative impact of geographical location detected in this study is surprising given that most data can now be communicated easily by electronic means, even between remotely separated areas. This result seems to indicate that the level of face-to-face communication is still critical in the collection and compilation of biodiversity information at a global scale. Even in the modern age, travelling 10 000 km, such as between east Asia and Europe, is time-consuming and costly, inevitably impeding face-to-face communication among scientists on a daily basis. Thus, it is much more difficult to

create an effective network that is necessary for information exchange from a far-away location. There is, nonetheless, a ray of light in the results. In the GBIF database, the most well-known worldwide database, the distance from host organizations had little effect on the amount of data collected from each country. This result indicates that the effect of geographical location can be overcome by active efforts to advertise databases and promote information exchange. This result also indicates that wealth, language and security, all of which explained non-negligible proportions of variations in data availability even in GBIF, had a stronger impact than geographical location. The effect of geographical location would also be influenced by when records were collected in each database; most of the data in GPDD were collected before 1999, when Internet access was still limited, possibly leading to a relatively strong effect of geographical location in this database.

The result also supports our hypothesis that a nation's security explains, at least to some degree, the amount of data compiled in the global databases. A low level of security discourages or forces researchers to stop scientific activities [19], affecting the amount of scientific data collected from such countries. Military expenditures may preclude spending on environmental science and management [43], and political instability can also reduce external financial support [44]. Though the potential effect of security on biodiversity science has been suggested by earlier studies [14], few studies have provided quantitative evidence, particularly at a global scale. The level of security can also affect the status of biodiversity itself. For example, warfare has widespread ecological consequences, both positive and negative, such as severe habitat destruction, accumulation of pollutants and increased poaching, but also creation of undisturbed habitats [19, 43]. This finding points to security as one of the serious barriers to the understanding of biodiversity conservation, particularly in areas with high levels of threats to biodiversity.

A drawback of this study is that the results cannot separate the impact of these drivers on the actual amount of data from the impact on the ability to collect existing information. For example, Japan has a relatively thorough coverage of data for both spatial and temporal dynamics of biodiversity, providing an important basis for conservation science within the country ([45,46] and references therein), but those data have not necessarily been used to contribute to important global biodiversity assessments [5,8]. This example well reflects the case where language (few English speakers) and geographical location (long distance from host organizations) pose a barrier to effective communication at a global level. On the other hand, these four barriers can also affect the actual amount of existing data through processes discussed earlier. Thus, separating these two impacts of the revealed barriers will be a next step for effectively tackling spatial biases in global biodiversity information. The interacting effects of the four factors might also be worth pursuing in future studies. Furthermore, as is always the case in regression-based studies, this study only shows the associations between variables and does not necessarily demonstrate that the four factors actually drive the spatial distribution of biodiversity information. Having said that, considering that earlier studies have rarely quantitatively identified potential drivers of biodiversity information, the findings of this study can serve as a basis for future efforts to obtain an undistorted understanding of global biodiversity status.

Finally, this study revealed that the four major barriers of wealth, language, geographical location and security have caused the under-collection of biodiversity data from biodiversity-rich countries in global/continental databases. This clearly indicates that an undistorted view of global biodiversity requires enhancing ecological education, research and collaboration not only in low-GDP countries but also countries with fewer English speakers, located far from Western countries. Although many research funds have been dedicated specifically to research in developing countries, care should be paid to ensure funds are not just given to countries with relatively many English speakers, but to distribute widely also among countries with few English speakers. Focusing more efforts on countries far from host organizations would be a good strategy when collecting available information in global databases. It is obviously not easy to enhance scientific activities in countries with low security levels, but we should at least recognize that security can have a serious consequence on science at a global scale and take opportunities when security has improved. Furthermore, although European and North American countries have often taken initiatives in global efforts to compile scientific knowledge, developed countries in other regions could also play an initiative role in efforts to collect information in their own regions. For example, in the case of biodiversity information in Asia, developed countries in east Asia and Oceania, such as Japan, South Korea, Australia and New Zealand, are in a position to take initiatives in collecting and compiling biodiversity information in southeast Asia, a region that is far from both Europe and North America, with high biodiversity, yet little information.

The findings of this study should be broadly applicable to other research areas. The result was essentially the same in EDB as in the other global databases, indicating that the conclusions are applicable at a continental scale, at least in Europe. Particularly in an era of global environmental change, compiling scientific knowledge at the global level offers a great potential for tackling many scientific challenges and problems [47,48]. Based on the above, we conclude that wealth, language, geographical location and security of countries should be taken into account equally when trying to understand the drivers of spatial heterogeneity in scientific knowledge, and efforts to overcome these barriers should be encouraged for the development of science at the global scale. Though not intended to be exhaustive, examples include the production of textbooks in the local language with local examples through co-authorship with local biologists (as pioneered by Richard Primack [49]), donation of books (as in the Gratis Book Scheme [50]), the creation of global partnerships of collaborating organizations (as with BirdLife International; [www.birdlife.org](http://www.birdlife.org)), the creation of tropical training courses on fieldwork and workshops in which participants convert their previous research into submitted papers (as with the Tropical Biology Association; [www.tropical-biology.org](http://www.tropical-biology.org)) or the creation of a global network of student conferences (as in the Student Conference in Conservation Science; [www.sccs-cam.org](http://www.sccs-cam.org)).

We thank all the people who developed and contributed to the four databases, without which this work was impossible. We also thank A. Symes and P. Taylor at BirdLife International for providing information on bird species richness and M. Amano for all the help. T.A. was supported by the JSPS Post-doctoral Fellowships for Research Abroad and W.J.S. by the Arcadia Fund. We thank two referees for their comments on an earlier version of this paper.

- IUCN (International Union for Conservation of Nature). 2011 IUCN Red List of Threatened Species. Version 2011.1. See <http://www.iucnredlist.org> (accessed 12 October 2011).
- Grenyer R *et al.* 2006 Global distribution and conservation of rare and threatened vertebrates. *Nature* **444**, 93–96. (doi:10.1038/nature05237)
- Tittensor DP, Mora C, Jetz W, Lotze HK, Ricard D, Vanden Bergh E, Worm B. 2010 Global patterns and predictors of marine biodiversity across taxa. *Nature* **466**, 1098–1101. (doi:10.1038/nature09329)
- Brooks TM, Mittermeier RA, da Fonseca GAB, Gerlach J, Hoffmann M, Lamoreux JF, Mittermeier CG, Pilgrim JD, Rodrigues ASL. 2006 Global biodiversity conservation priorities. *Science* **313**, 58–61. (doi:10.1126/science.1127609)
- Butchart SHM *et al.* 2010 Global biodiversity: indicators of recent declines. *Science* **328**, 1164–1168. (doi:10.1126/science.1187512)
- Walpole M *et al.* 2009 Tracking progress toward the 2010 biodiversity target and beyond. *Science* **325**, 1503–1504. (doi:10.1126/science.1175466)
- Pennisi E. 2010 Filling gaps in global biodiversity estimates. *Science* **330**, 24. (doi:10.1126/science.330.6000.24)
- Collen B, Loh J, Whitmee S, McRae L, Amin R, Baillie JEM. 2009 Monitoring change in vertebrate abundance: the Living Planet Index. *Conserv. Biol.* **23**, 317–327. (doi:10.1111/j.1523-1739.2008.01117.x)
- Millennium Ecosystem Assessment. 2005 *Ecosystems and human well-being: current state and trends*, vol. 1. Washington, DC: Island Press.
- González-Suárez M, Lucas PM, Revilla E. 2012 Biases in comparative analyses of extinction risk: mind the gap. *J. Anim. Ecol.* **81**, 1211–1222. (doi:10.1111/j.1365-2656.2012.01999.x)
- Boakes EH, McGowan PJK, Fuller RA, Ding C-q, Clark NE, O'Connor K, Mace GM. 2010 Distorted views of biodiversity: spatial and temporal bias in species occurrence data. *PLoS Biol.* **8**, e1000385. (doi:10.1371/journal.pbio.1000385)
- Boitani L, Maiorano L, Baisero D, Falucci A, Visconti P, Rondinini C. 2011 What spatial data do we need to develop global mammal conservation strategies? *Phil. Trans. R. Soc. B* **366**, 2623–2632. (doi:10.1098/rstb.2011.0117)
- Yesson C *et al.* 2007 How global is the global biodiversity information facility? *PLoS ONE* **2**, e1124. (doi:10.1371/journal.pone.0001124)
- Collen B, Ram M, Zamin T, McRae L. 2008 The tropical biodiversity data gap: addressing disparity in global monitoring. *Trop. Conserv. Sci.* **1**, 75–88.
- Feeley KJ, Silman MR. 2011 The data void in modeling current and future distributions of tropical species. *Glob. Change Biol.* **17**, 626–630. (doi:10.1111/j.1365-2486.2010.02239.x)
- Martin LJ, Blossey B, Ellis E. 2012 Mapping where ecologists work: biases in the global distribution of terrestrial ecological observations. *Front. Ecol. Environ.* **10**, 195–201. (doi:10.1890/110154)
- Newbold T. 2010 Applications and limitations of museum data for conservation and ecology, with particular attention to species distribution models. *Prog. Phys. Geogr.* **34**, 3–22. (doi:10.1177/0309133309355630)
- Montgomery S. 2004 Of towers, walls, and fields: perspectives on language in science. *Science* **303**, 1333–1335. (doi:10.1126/science.1095204)
- Kanyambwa S. 1998 Impact of war on conservation: Rwandan environment and wildlife in agony. *Biodivers. Conserv.* **7**, 1399–1406. (doi:10.1023/A:1008880113990)
- NERC Centre for Population Biology, Imperial College. 2010 The Global Population Dynamics Database, version 2. See <http://www3.imperial.ac.uk/cpb/databases/gpdd> (accessed 25 January 2013).
- Wikelski M, Kays R. 2011 Movebank: archive, analysis and sharing of animal movement data. See <http://www.movebank.org> (accessed 16 November 2011).
- du Feu CR, Joys AC, Clark JA, Fiedler W, Downie IS, van Noordwijk AJ, Spina F, Wassenaar R, Baillie SR. 2009 EURING Data Bank geographical index 2009. See <http://www.euring.org/edb> (accessed 4 June 2012).
- Mac Nally R. 1996 Hierarchical partitioning as an interpretative tool in multivariate inference. *Aust. J. Ecol.* **21**, 224–228. (doi:10.1111/j.1442-9993.1996.tb00602.x)
- Institute for Economics and Peace. 2012 *Global peace index 2012*. Sydney, Australia: Institute for Economics and Peace.
- World Wildlife Fund. 2006 WildFinder: online database of species distributions, v. 01.06. See [www.wwfus.org/wildfinder](http://www.wwfus.org/wildfinder) (accessed 6 June 2012).
- Dekker AJFM. 2001 Biodiversity information sharing service (BISS). *ASEAN Biodiversity* **1**, 64–65.
- Roskov YR, Bisby FA, Zucchini JL, Schrire BD, White RJ. 2005 ILLDIS World Database of Legumes, 10th edn. See <http://www.ildis.org> (accessed 6 June 2012).
- Sutherland WJ. 1998 The importance of behavioural studies in conservation biology. *Anim. Behav.* **56**, 801–809. (doi:10.1006/anbe.1998.0896)
- Central Intelligence Agency. 2011 The World Factbook. See <https://www.cia.gov/library/publications/the-world-factbook/index.html> (accessed 4 June 2012).
- Lewis MP. 2009 *Ethnologue: languages of the world*, 16th edn. See <http://www.ethnologue.com>. Dallas, TX: SIL International.
- Crystal D. 2010 *The Cambridge encyclopedia of language*, 3rd edn. Cambridge, UK: Cambridge University Press.
- European Commission. 2006 Special Eurobarometer 243: Europeans and their languages. Brussels, UK: European Commission.
- Weidmann NB, Gleditsch KS. 2012 cshapes dataset and utilities. R package. See <http://cran.r-project.org/web/packages/cshapes/index.html> (accessed 25 October 2012).
- R Development Core Team. 2011 *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. See <http://www.R-project.org/> (accessed 4 June 2012).
- Moore J, Balmford A, Allnutt T, Burgess N. 2004 Integrating costs into conservation planning across Africa. *Biol. Conserv.* **117**, 343–350. (doi:10.1016/j.biocon.2003.12.013)
- Bjørnstad ON. 2005 ncf: spatial nonparametric covariance functions. R package. See <http://cran.r-project.org/web/packages/ncf/index.html> (accessed 1 November 2012).
- Walsh C, Mac Nally R. 2008 hier.part: hierarchical partitioning. R package. See <http://cran.r-project.org/web/packages/hier.part/index.html> (accessed 25 October 2012).
- Blackburn TM, Duncan RP. 2001 Establishment patterns of exotic birds are constrained by non-random patterns in introduction. *J. Biogeogr.* **28**, 927–939. (doi:10.1046/j.1365-2699.2001.00597.x)
- Heikkinen RK, Luoto M, Virkkala R, Rainio K. 2004 Effects of habitat cover, landscape structure and spatial variables on the abundance of birds in an agricultural-forest mosaic. *J. Appl. Ecol.* **41**, 824–835. (doi:10.1111/j.0021-8901.2004.00938.x)
- Burnham KP, Anderson DR. 2002 *Model selection and multimodel inference: a practical information-theoretic approach*, 2nd edn. New York, NY: Springer.
- United Nations Educational, Scientific and Cultural Organization. 2010 *UNESCO science report 2010: the current status of science around the world*. Paris: UNESCO Publishing.
- King DA. 2004 The scientific impact of nations. *Nature* **430**, 311–316. (doi:10.1038/430311a)
- Machlis GE, Hanson T. 2008 Warfare ecology. *BioScience* **58**, 729–736. (doi:10.1641/B580809)
- Inogwabini BI, Ilambu O, Gbanzi MA. 2005 Protected areas of the Democratic Republic of Congo. *Conserv. Biol.* **19**, 15–22. (doi:10.1111/j.1523-1739.2005.00181.x)
- Higuchi H, Primack RB. 2009 Conservation and management of biodiversity in Japan: an introduction. *Biol. Conserv.* **142**, 1881–1883. (doi:10.1016/j.biocon.2009.03.011)
- Amano T. 2010 Monitoring bird populations—a first step towards the effective conservation of biodiversity. *Ornith. Sci.* **9**, 1–2. (doi:10.2326/osj.9.1)
- Foley JA *et al.* 2011 Solutions for a cultivated planet. *Nature* **478**, 337–342. (doi:10.1038/nature10452)
- Root TL, Price JT, Hall KR, Schneider SH, Rosenzweig C, Pounds JA. 2003 Fingerprints of global warming on wild animals and plants. *Nature* **421**, 57–60. (doi:10.1038/nature01333)
- Primack RB. 2001 Publish again in another language. *Conserv. Biol.* **15**, 290–291.
- Sutherland WJ. 2001 A new approach to global book distribution. *Nature* **411**, 738. (doi:10.1038/35081292)