Journal of Applied Ecology

British Ecological Society

*Journal of Applied Ecology* 2010, **47**, 377–385 doi: 10.1111/j.1365-2664.2010.01777.x

# The impact of proxy-based methods on mapping the distribution of ecosystem services

**Felix Eigenbrod[1], Paul R. Armsworth[1]†, Barbara J. Anderson[2], Andreas Heinemeyer[3], Simon Gillings[4], David B. Roy[5], Chris D. Thomas[2] and Kevin J. Gaston[1]***

[1]*Biodiversity & Macroecology Group, Department of Animal & Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK;* [2]*Department of Biology, University of York, PO Box 373, York YO10 5YW, UK;* [3]*Stockholm Environment Institute (York-Centre) and National Centre for Earth Observation (NCEO; York-Centre), Grimston House, Department of Biology, University of York, York YO10 5DD, UK;* [4]*British Trust for Ornithology, The Nunnery, Thetford, Norfolk IP24 2PU, UK; and* [5]*NERC Centre for Ecology and Hydrology, Wallingford, Crowmarsh Gifford, Wallingford, Oxfordshire OX10 8BB, UK*

## Summary

**1.** An increasing number of studies are examining the distribution and congruence of ecosystem services, often with the goal of identifying areas that will provide multiple ecosystem service 'hotspots'. However, there is a paucity of data on most ecosystem services, so proxies (e.g. estimates of a service for a particular land cover type) are frequently used to map their distribution. To date, there has been little attempt to quantify the effects of using proxies on distribution maps of ecosystem services, despite the potentially large errors associated with such data sets.

**2.** Here, we provide the first study examining the effects of using proxies on ecosystem service maps and the degree of spatial congruence of these maps with primary data, using England as a case study.

**3.** We show that land cover based proxies provide a poor fit to primary data surfaces for biodiversity, recreation and carbon storage, and that correlations between ecosystem services change depending on whether primary or proxy data are used for the analyses.

**4.** The poor fit of proxies to primary data was also evident when we selected hotspots of single ecosystem services, and consistency between raw and modelled surfaces was extremely low when considering the locations that were coincident hotspots for multiple services.

**5.** *Synthesis and applications*. Proxies may be suitable for identifying broad-scale trends in ecosystem services, but even relatively good proxies are likely to be unsuitable for identifying hotspots or priority areas for multiple services.

**Key-words:** biodiversity, carbon, conservation planning, ecosystem services, England, GIS, hotspots, natural capital assets, recreation, spatial value transfer

## Introduction

Awareness of the importance of key ecosystem services in maintaining human well-being has greatly increased since the publication of the Millennium Ecosystem Assessment (2005). Many conservation biologists, conservation NGOs and public agencies have embraced the 'ecosystem service based approach' [e.g. Daily & Matson 2008; the Natural Capital Project (http://www.naturalcapitalproject.org/); the UK's Department of Environment, Food and Rural Affairs (http://www.ecosystemservices.org.uk/)] as a way to protect biodiversity and other critical natural assets. The limited studies to date suggest that spatial variation in biodiversity and other ecosystem services are not necessarily positively correlated (e.g. Naidoo *et al.* 2008; Anderson *et al.* 2009). There is thus increasing interest in identifying small areas, or 'hotspots', that are important for both biodiversity and multiple ecosystem services (Chan *et al.* 2006; Turner *et al.* 2007; Naidoo *et al.* 2008; Egoh *et al.* 2009). In addition, studies that aim to predict

*Correspondence author. E-mail: k.j.gaston@sheffield.ac.uk
†Present address: Paul R. Armsworth, Ecology & Evolutionary Biology, University of Tennessee, Knoxville, TN 37996-1601, USA.

© 2010 The Authors. Journal compilation © 2010 British Ecological Society

ecosystem service distributions and tradeoffs in the face of climate change are beginning to appear (e.g. Metzger *et al.* 2006; Nelson *et al.* 2009).

Perhaps the greatest obstacle to substantial progress in assessing ecosystem services is a lack of data – there is simply none available for most services in most of the world. This has led to many maps of ecosystem services being based on crude estimates (cf. Naidoo *et al.* 2008), though the quality of data varies widely between studies and services. In general, the methods used to produce ecosystem service maps can be broadly divided into those that are based on at least some primary data from within the study region, and those that are not (proxies). The former category can be further subdivided into maps based on representative sampling across the whole study region and modelled surfaces based on primary data, while the latter can be broadly divided into land cover based proxies and prior knowledge driven modelled surfaces (Table 1).

Ecosystem service maps based on representative sampling of the entire study region are limited to very few services (biodiversity, recreation), and are either large-scale but coarse resolution (e.g. global distribution of birds; Orme *et al.* 2005), or limited to regional analyses of well studied areas (e.g. recreational use of the countryside in England; Eigenbrod *et al.* 2009). Ecosystem service maps created by modelling the relationship between samples of a service and readily measurable environmental variables (climate, land cover, soil types) are more common within the ecosystem service literature. They have been used in large-scale multi-service studies to map carbon storage (e.g. Milne & Brown 1997; as used by Eigenbrod *et al.* 2009), carbon fluxes (e.g. McGuire *et al.* 2001), and biodiversity priority areas (e.g. Chan *et al.* 2006). They are also sometimes used for studies examining a single service, such as to model pollination potential of forest patches in a portion of southern Madagascar (Bodin *et al.* 2006). Note that there is considerable debate (e.g., inside the Millennium Ecosystem Assessment process and the Natural Capital Project) as to whether biodiversity is an ecosystem service in its own right, or whether it should be considered separately, as elements of biodiversity may play an integral role in underpinning other ecosystem services. We consider biodiversity to be an ecosystem service in its own right here, but we recognize the need for further study to examine the role of different elements of biodiversity in supporting other ecosystem services.

Proxy-based maps are more common than maps based on primary data (e.g. Sutton & Costanza 2002; Chan *et al.* 2006; Troy & Wilson 2006; Turner *et al.* 2007; Egoh *et al.* 2008). These are often based on digital raster land cover maps due to the widespread availability of such data. These types of analyses include 'spatial value transfer' (Troy & Wilson 2006); otherwise known as 'benefits transfer' (Plummer 2009), where a

**Table 1.** Major approaches to producing maps of ecosystem services

| Methodology | Advantages | Disadvantages | Examples |
|---|---|---|---|
| **Requires primary data from within the study region** | | | |
| Representative sampling of entire study region (e.g. atlas data; region-wide survey) | Provides the best estimate of actual levels of ecosystem services<br>Well suited to heterogeneous ecosystem services | Expensive or difficult to obtain, so often unavailable<br>Degree of error will depend on sampling intensity | Recreation[1,2]<br>Biodiversity[3,4]<br>Reed and fish production[5] |
| Modelled surface based on sampling from within study region | May require far fewer samples than representative sampling<br>Smoothing will overcome sampling heterogeneity | Smoothing will mask true heterogeneity in the service<br>Error will depend on sample size and fit to modelled variables | Carbon storage[2]<br>Biodiversity[6]<br>Biodiversity 'hotspots'[7,8]<br>Carbon sequestration[9]<br>Agricultural production[10]<br>Pollination[11,12]<br>Water retention[13]<br>Recreation[14] |
| **Does not require primary data from within the study region** | | | |
| Land cover based proxy (e.g. benefits transfer) | Enables mapping of ecosystem services in regions where primary data are lacking | Fit of proxy to actual data may be very poor | Biodiversity (existence value and bioprospecting)[7,15,16,17]<br>Recreation[7,16,18]<br>Carbon storage[6,7,8,15,16]<br>Flood control; soil conservation[7,15,16] |
| Proxy based on logical combination of likely causal variables | Can offer a major improvement on performance of land cover based proxies alone, without the need for much additional data | Potential for large error is still high if assumed causal variables are not in fact good predictors | Recreation[6,16,19,20]<br>Flood control, water provision[6]<br>Soil accumulation[21] |

1. Larsen *et al.* 2008; 2. Eigenbrod *et al.* 2009; 3. Egoh *et al.* 2009; 4. Anderson *et al.* 2009; 5. Hein *et al.* 2006; 6. Chan *et al.* 2006; 7. Turner *et al.* 2007; 8. Naidoo *et al.* 2008; 9. Nelson *et al.* 2008; 10. Naidoo & Iwamura 2007; 11. Kremen *et al.* 2004; 12. Bodin *et al.* 2006; 13. Guo & Gan 2002; 14. Willemen *et al.* 2008 15. Eade & Moran 1996; 16. Sutton & Costanza 2002; 17. Naidoo & Ricketts 2006; 18. Metzger *et al.* 2005; 19. Troy & Wilson 2006; 20. Önal & Yanprechaset 2007; 21. Egoh *et al.* 2008.

monetary value for an ecosystem service is assigned to each land cover class based on estimates from prior studies. For example, the estimate of the value of nutrient cycling for tropical forests in Costanza *et al.*'s. (1997) seminal valuation of global ecosystem is based on a single Indian case study (Chopra 1993). Several authors have then used Costanza *et al.*'s biome-level estimates of the values of ecosystem service to produce maps (Kreuter *et al.* 2001; Sutton & Costanza 2002; Turner *et al.* 2007). Benefits transfer based mapping of ecosystem services has also been done over smaller spatial extents (Eade & Moran 1996; Naidoo & Ricketts 2006; Troy & Wilson 2006). Constant ecosystem service values may also be assigned to land cover classes in large-scale studies where more detailed information is lacking. For example, the carbon storage layer (Gibbs 2007) used by Naidoo *et al.* (2008) only gives a single estimate of carbon storage for each biome. In a second example, Metzger *et al.* (2006) considered all non-urban land cover in Europe except cropland to have equal recreational potential, and cropland and urban areas to have no recreational potential.

In addition to mapping ecosystem services using values from land cover based proxies alone, a number of studies have incorporated existing knowledge about causal relationships from multiple existing datasets into a proxy layer. For example, Chan *et al.* (2006) weighted land use by its proximity to human population centres when mapping the distribution of flood control and rural recreation services. Similarly, Troy & Wilson (2006) subdivided some land cover classes to account for proximity to settlements (e.g. beach vs. beach near human dwelling), and Egoh *et al.* (2008) combined maps of soil erosion potential and vegetation cover to create a map of soil retention for South Africa.

There is widespread recognition that proxy-based maps are crude estimates of actual distributions of ecosystem services (e.g. Turner *et al.* 2007; Naidoo *et al.* 2008), but, until recently, little discussion of the likely magnitude of such errors. Plummer (2009) suggests that the errors associated with ecosystem service mapping based on benefits transfer are likely to be very high, due primarily to 'generalization error'. Generalization error is that due to the assumption underlying benefits transfer based mapping that the value of an ecosystem service for a particular land cover type is (i) the same in the area being mapped as in the studies from which the value was obtained; and (ii) is constant across the entire area being mapped. Generalization error can be further subdivided into the error associated with extrapolating estimates of the *economic* value of an ecosystem service taken in one location and assuming these apply to a different location, and into the error associated with failing to account for the spatial variability in *biophysical* measurements of ecosystem services. We only consider the latter here. When considering biophysical variation in ecosystem properties, ecologists have long recognized the dangers of extrapolating from the results of a study at one location to wider geographical areas (e.g. Lawton 2000), and equally that spatial scaling is common in ecological systems (e.g. Wiens 1989); however, these insights have received insufficient attention in the ecosystem services literature.

Little is known about how the errors associated with proxy-based methods might affect the inferences drawn from analyses because quantifying the impacts of such errors is difficult without comparisons with primary data. These comparisons remain wanting, and shortcomings of proxy-based maps have only been addressed through sensitivity analyses (e.g. Turner *et al.* 2007; Nelson *et al.* 2008). Here, we compare patterns and spatial congruence of biodiversity (species of conservation concern) and two other ecosystem services (carbon storage and rural recreation) for England, based on proxy-based maps and maps of the same services based on primary data. Our study represents a first attempt to quantify the potential effects of using proxies on the results of studies mapping the spatial distribution and congruence of ecosystem services.

## Materials and methods

### CREATION OF ECOSYSTEM SERVICE MAPS

Our approach for comparing ecosystem service maps based on primary data with maps based on proxy data was deliberately simple, replicating approaches that are most commonly taken in the literature (Table 1). First, we constructed maps for the three services [biodiversity (species of conservation concern), carbon storage, and recreation] based on the best available primary data. We used the recorded presences of 326 terrestrial Biodiversity Action Plan (BAP) (Anonymous 1994) (birds, bryophytes, butterflies, mammals, herptiles and plants present in England) species measured at the $2 \times 2$ km grid square resolution as our measure of biodiversity. The biodiversity layer is thus effectively an accumulation of observer records, and could therefore be prone to observer bias due to higher densities of (volunteer) observers in the south and east of England. However, for taxonomic groups where such bias has been assessed (plants – Preston, Pearman & Dines 2002; butterflies – Asher *et al.* 2001) it is judged to be a minor issue for species with a restricted range and/or of conservation concern (e.g. UK BAP species), as such species are more intensively sampled than those which are common and widespread. Species richness derived from presence/absence within 2 km $\times$ 2 km grid squares is therefore a reasonable representation of true BAP biodiversity. The carbon storage layer is an estimate of combined organic soil and above ground vegetation carbon (in kg $C/m^2$) calculated at the $1 \times 1$ km resolution. We considered the number of day leisure visits ($n = 6279$ for all of England) to rural locations (where the main purpose was enjoyment of the landscape), to be representative of the recreation value of the landscape. Day leisure visits were obtained from a telephone-based survey of leisure trips of the entire English population - the England Leisure Visits Survey 2005 (main survey) (http://www.naturalengland.org.uk/ourwork/enjoying/research/monitor/leisurevisits/elvdownloads.aspx).

We then calculated the average values of each of these services at the $10 \times 10$ km grid square resolution. We used $10 \times 10$ km grids as this was the finest resolution that was suitable for all three data sets, and because earlier work (Anderson *et al.* 2009) showed that correlations between these ecosystem services are generally qualitatively similar whether $2 \times 2$ km or $10 \times 10$ km resolution data is used in the analyses. Detailed methods are available in Data S1. Second, we created land cover proxy maps from these three primary data maps by superimposing the data maps onto a land cover map, and then assigning a value to each of 11 land cover classes based on the average value of each service for a particular land cover class in the primary data

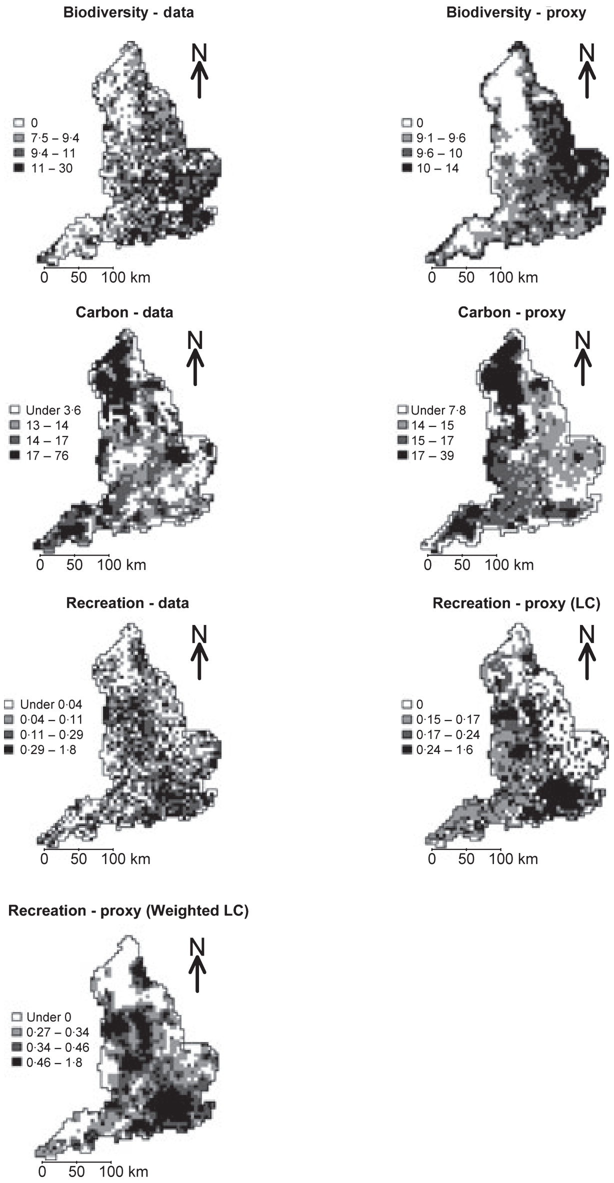


**Fig. 1.** Distribution of ecosystem services in England based both on proxies and primary data. The cut-off values are quartiles calculated separately for each map. Units are the average number of visits (recreation), number of BAP species (biodiversity) per $2 \times 2$ km square within each $10 \times 10$ km square, and the average number of kilograms of carbon stored per $m^2$ in each $1 \times 1$ km square within each $10 \times 10$ km square.

map (Fig. 1). These land cover classes are: deciduous forest; coniferous forest; cropland; pasture/grassland; moorland heather; wetland; inland water; bare rock/quarry; urban/suburban; coastal rock; submerged/tidal rock (Table S1). Our land cover based proxy maps differ from existing studies in that we assigned values to each land cover type based on direct estimates from a single data set, rather than by combining estimates from multiple case studies as carried out by Costanza *et al.* (1997) and Troy & Wilson (2006). Our approach still makes the same key simplifying assumption that the ecosystem service value of a particular land cover type remains constant across the entire study area, and should thus be similarly affected by this component of generalization error (Plummer 2009) as studies based on benefits transfer. However, our approach should be a better reflection of the actual distribution of ecosystem services as it uses a consistent methodology and is based on data covering the entire study region. We also constructed a land cover based proxy map weighted by human population density and ease of access for recreation (Fig. 1) using the same approach as Chan *et al.* (2006) to allow us to compare the effectiveness of this approach to using a proxy based just on land cover. Detailed methods for the proxy map creation are available in Data S1. All GIS analyses were carried out in ArcGIS/ArcInfo 9.2 (ESRI, Redlands, CA, USA).

### ANALYSIS

We ran two types of analyses to quantify the effects of using proxies on inferences made from mapping ecosystem services. We compared the degree of congruence of the maps created from primary data and from proxies for England as a whole (Table 2; Fig. 1) and for 'hotspots' for each of the three ecosystem services (Table 3; Fig. 2). We defined hotspots to be the largest decile (approximately the top 10% of grid cells, depending on ties), the two largest deciles (20% of grid cells) or the three largest deciles (30% of grid cells) for each service (Fig. 2). We also calculated the bivariate relationships (Spearman correlation coefficient) between the three ecosystem services (Table 2) and the degree of congruence of the hotspots (Table 4; Fig. 3) using maps based on both the primary data and on proxies. The frequency distributions of the data limit options for controlling for spatial autocorrelation in residuals. Because of this and because we lack essential information regarding spatial processes (e.g. regarding their stationarity), we have taken the simpler course of presenting non-spatial statistics. All statistical analyses were carried out in R 2.7.2 (R Development Core Team 2008).

## Results

England-wide comparisons suggest that distribution maps based on proxies provide poor estimates of the distributions of ecosystem services based on primary data, with the proxy distributions at best only very broadly similar to maps based on primary data (Table 2; Fig. 1). Unsurprisingly, the highest correlation was for carbon storage - the only one of our services where the primary data map was itself a modelled surface partially based on land cover – although even there the correlation was not especially high (Spearman's *rho* = 0·57). Weighting land cover by population and proximity to roads improved the fit to the distribution based on representative sampling from *rho* = 0·42 to 0·50 (Fig. 1; Table 2) for recreation; the correlation between the two different recreation proxy surfaces was *rho* = 0·81.

The bivariate spatial correlations between the three ecosystem services were broadly similar whether we used proxy or primary data (Table 2). However, using the biodiversity proxy rather than primary biodiversity data meant that the weak positive association with recreation disappeared (*rho* = −0·02 vs. 0·18), and that there was a stronger negative association with carbon storage 'data' (*rho* = −0·36 vs. −0·19). Interestingly, using the land cover based proxy for recreation led to associations with the other ecosystem services that were generally more similar to those based on primary data than if we

**Table 2.** Correlations (Spearman's *rho*) ($n = 1485$) between ecosystem service maps based on primary data and proxy maps for England as a whole (Fig. 1)

| | Biodiversity – proxy | Carbon – data | Carbon – proxy | Recreation – data | Recreation – land cover alone | Recreation – land cover + population + access |
|---|---|---|---|---|---|---|
| Biodiversity – data | 0·37*** | −0·19*** | −0·23*** | 0·18*** | −0·13*** | 0·26*** |
| Biodiversity – proxy | | −0·36*** | −0·62*** | −0·02 | −0·12*** | 0·08** |
| Carbon – data | | | 0·57*** | −0·19*** | −0·24*** | −0·33*** |
| Carbon – proxy | | | | −0·15*** | −0·15*** | −0·32*** |
| Recreation – data | | | | | 0·42*** | 0·50*** |
| Recreation – land cover alone | | | | | | 0·81*** |

**Table 3.** Comparison of hotspots of ecosystem services based on primary data and land cover based proxy data. The possible number of cells (all $10 \times 10$ km grid cells in England) is 1485 for all analyses

| Ecosystem service | Hotspot criteria | Number of cells in hotspot: Primary data | Proxy data | Overlap between primary and proxy data | Percentage of cells that overlap compared with primary data |
|---|---|---|---|---|---|
| Biodiversity | Top 10% | 149 | 149 | 34 | 23 |
| Carbon | Top 10% | 149 | 149 | 93 | 62 |
| Recreation | Top 10% | 149 | 149 | 25 | 17 |
| Biodiversity | Top 30% | 446 | 446 | 187 | 42 |
| Carbon | Top 30% | 446 | 446 | 280 | 63 |
| Recreation | Top 30% | 438 | 407 | 240 | 54 |

used for recreation the proxy that included weighting by population and proximity to roads, despite the latter proxy having a better fit to primary data than the former (Table 2). Potential drivers of the correlations between the primary data surfaces for these three ecosystem services are discussed in Anderson *et al.* (2009) and are not considered further here.

The degree of spatial overlap between hotspots based on raw data versus proxies depended on the ecosystem service considered, and on the number of hotspot cells examined. For all three services, there was reasonable (42–63%) congruence between data- and proxy-based hotspots when these were defined as the top 30% of grid squares (Table 3). However, when the top 10% of cells was considered the overlap was much lower: 23% for biodiversity, and 17% for recreation. This indicates that the proxy maps were particularly poor predictors of the best (10%) areas in England for these two services (Fig. 2). However, this was not the case for carbon storage, where the degree of congruence of hotspots based on proxy and primary data was as high for the top 10% of cells (62%) as for the top 30% (63%) (Table 3; Fig. 2).

There was also very little spatial concordance in the cells in which two or more ecosystem service hotspots overlapped between maps created using primary data with those using proxy data (Table 4; Fig. 3). The highest degree of overlap between hotspots based on primary data with those based on proxies was 19% for biodiversity-recreation, when hotspots were defined as the top 10% of cells. Finally, there was at best 8% spatial concordance in the locations where hotspots of all three ecosystem services overlapped between the proxy- and primary data-based maps (Table 4).

## Discussion

Plummer (2009) calls for studies to assess the magnitude of potential errors in ecosystem service maps based on spatial value transfer. Ours is the first such study, and the results confirm Plummer's suspicion that the errors associated with spatial value transfer based proxies are considerable: proxies that are based on coarse or categorical input data (e.g. broad vegetation types) are likely to provide poor estimates of the actual distributions of ecosystem services. For biodiversity (species of conservation concern) and recreation – the two surfaces where the primary data maps were based on representative sampling – the correlations between the proxies and the primary data surfaces were less than 0·50, even using ranked data. Furthermore, correlations between two of the three bivariate combinations of ecosystem services were substantially different when proxies were used rather than primary data surfaces, and sets of hotspots selected using proxy data bore little resemblance to those selected using primary data. Finally, most proxies used in the literature are based on estimates for ecosystem services from either outside or within a small portion of the study area, and thus are likely to have even

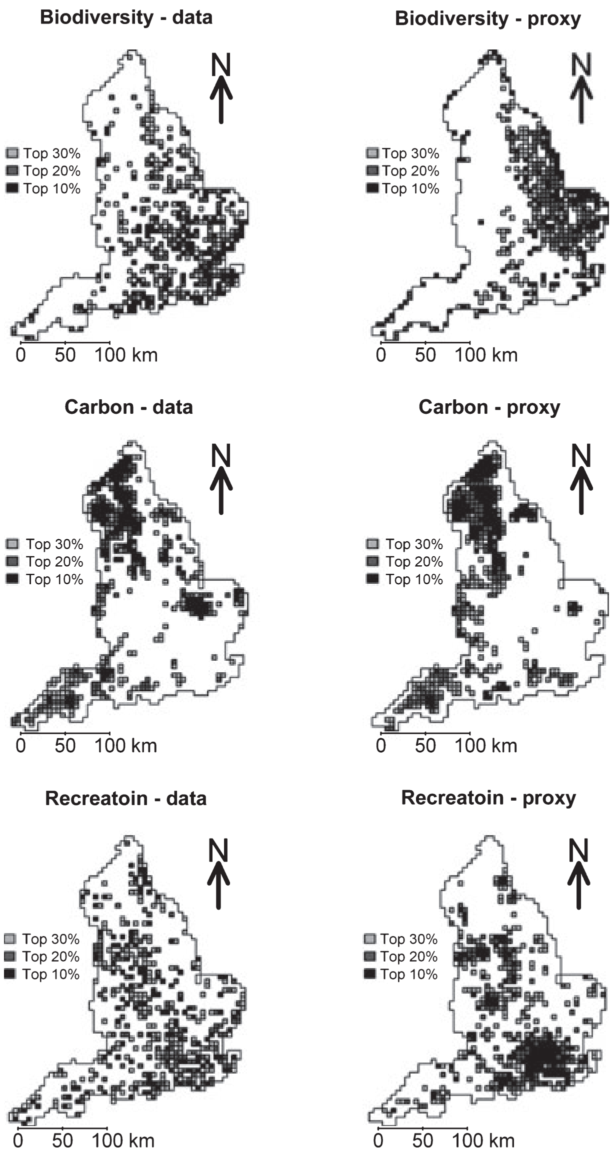


**Fig. 2.** Ecosystem service hotspots, as calculated using maps based on primary data and proxies. Hotspots are defined as being the top 30%, top 20% or top 10% of cells.

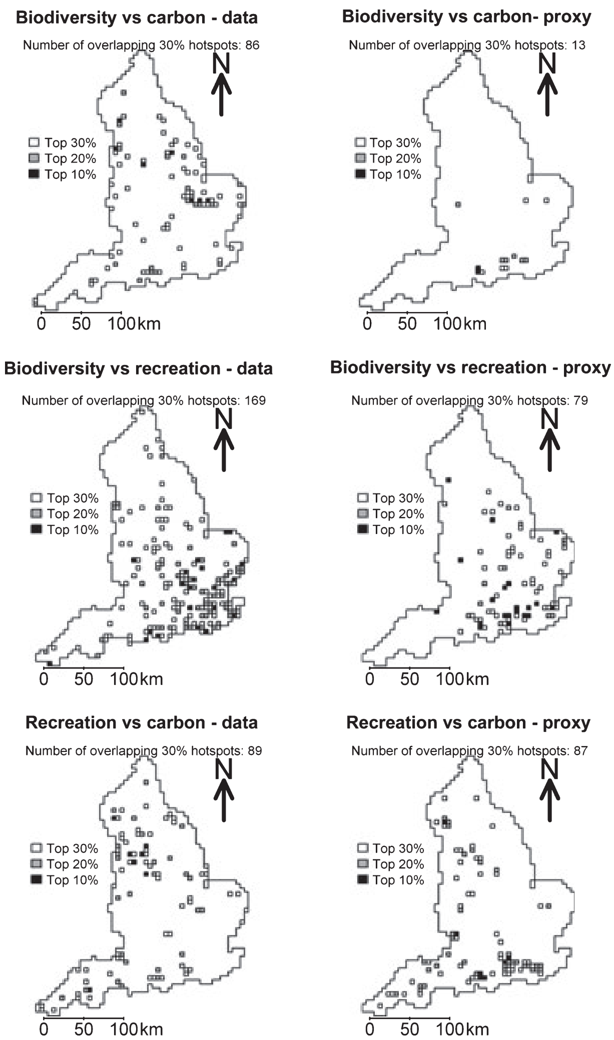


**Fig. 3.** Overlap of ecosystem service hotspots based on primary data and proxy data. Results are shown for the overlap of hotspots defined as being the top 30%, top 20% or top 10% of cells.

weaker correlations with the actual distributions of ecosystem services than in our study, where the proxies are based on the mean values of the primary data for the entire study area. In addition, we only consider the implications of failing to

**Table 4.** Comparisons of multi-ecosystem service hotspots based on primary data and proxy data

| Number of Congruent cells | | | | | |
|---|---|---|---|---|---|
| Ecosystem services | Hotspot criteria | Primary data | Proxy data | Overlap between primary and proxy | Percentage of cells that overlap compared with primary data |
| Biodiversity and carbon | Top 10% | 7 | 2 | 0 | 0 |
| Biodiversity and recreation | Top 10% | 21 | 20 | 4 | 19 |
| Recreation and carbon | Top 10% | 7 | 7 | 0 | 0 |
| Carbon and recreation and biodiversity | Top 10% | 0 | 2 | 0 | 0 |
| Biodiversity and carbon | Top 30% | 86 | 13 | 4 | 5 |
| Biodiversity and recreation | Top 30% | 169 | 79 | 21 | 12 |
| Recreation and carbon | Top 30% | 89 | 87 | 16 | 18 |
| Carbon and recreation and biodiversity | Top 30% | 25 | 12 | 2 | 8 |

account for biophysical variation in ecosystem services in our study; these errors will be equally problematic when considering spatial variation in economic valuation estimates of ecosystem services.

The effect that the errors in ecosystem service proxies have on inferences will depend on the grain of the analysis that is attempted. For example, if the goal of our study had been to identify whether the southeast of England had more biodiversity than the northwest, then we would have obtained the same answer using our proxies as using our primary data surfaces for biodiversity. However, proxies were completely unsuitable for selecting the top 10% of land area in England for biodiversity or for recreation. Moreover, proxies were highly inaccurate for selecting areas important for multiple hotspots, even when 'important' was defined very broadly as the top 30% of grid squares. This has major implications for management given the large number of recent studies that identify small and/or overlapping areas of high ecosystem service value (Chan *et al*. 2006; Turner *et al*. 2007; Naidoo *et al*. 2008; Egoh *et al*. 2009; Nelson *et al*. 2009).

In general, modelled surfaces – of which proxies are extreme examples – should not be used to select hotspots (although areas identified in this way may represent useful locations for further examination). This is because hotspots are, by definition, a small number of sites selected from the tail end of the distribution of a variable. Modelled surfaces, which are based primarily on values around the mean of the variable, will invariably be poor at predicting the true locations of these extreme values. For example, in our land cover based proxy surface for biodiversity, cropland is assigned a higher biodiversity value than forest cover. While this is generally true (due to the large numbers of species of conservation concern within small fragments of natural vegetation within such areas and not the crops *per se*), some squares have much lower values than others, and indeed much lower values than the highest-value forested squares due to differences in the quality of the sites that are not apparent from land cover alone. It is such 'unusual' squares that are driving the distribution of the 10% hotspots for biodiversity in the raw data (Fig. 2). The problems associated with selecting hotspots based on modelled surfaces can be minimized by defining hotspots very broadly (e.g. Egoh *et al*. 2009). This is because as the proportion of the total area defined as a hotspot is increased, and fewer and fewer 'unusual' squares are included, the distribution of 'hotspots' will begin to look more like the broad pattern for the ecosystem service, reducing the importance of precisely identifying small-scale heterogeneity in the underlying data.

The use of proxies for identifying areas important for multiple ecosystem services is also problematic, due to error propagation. For example, if the odds of correctly identifying hotspots of two ecosystem services are 40% and 30%, respectively, then the expected probability of correctly identifying overlapping hotspots is only 12% ($0{\cdot}4 \times 0{\cdot}3 = 0{\cdot}12$), provided that the errors underlying the two ecosystem service layers are independent. The actual degree of congruence may be even lower. Proxy-based hotspots of carbon storage and biodiversity identified using the 30% cut-off in our analyses were 63% and 42% congruent, respectively, with the primary data based hotspots. However, hotspots of these two ecosystem services based on proxy data were only 5% congruent with the overlap obtained from primary data, rather than the 26% expected from simple probability theory ($0{\cdot}63 \times 0{\cdot}42 = 0{\cdot}26$). In this case, the low congruence is because the grid cells where the hotspots for these two services overlapped were disproportionately those where the proxy and primary based hotspots for these two services *did not* overlap (Figs 2 and 3). Finally, once multiple ecosystem services are considered, confidence in the correct identification of coincident hotspots becomes extremely low.

Our results also illustrate that using proxies based on strong causal drivers is still unlikely to result in a good fit to primary data. The two variables we used to create the weighted recreation surface – access and population density – are both known to be strong drivers of recreational use (e.g. Hörnsten & Fredman 2000; Natural England 2006), yet the fit to primary data was still only $rho = 0{\cdot}50$. This approach will be even more problematic where the associations among perceived causal drivers and a particular ecosystem service are less well understood, which is likely to be the case for most services.

Of course, as with all such data, the primary data surfaces we used in these analyses are themselves imperfect measures of the ecosystem services we consider here. However, errors in our primary data sets should have little qualitative impact on the major finding of this study – that land cover based proxies are unlikely to capture finer-scale variation in ecosystem services and hence result in a poor fit to actual distributions. Indeed, it is likely that better primary data would show even greater variability of ecosystem services within land cover types.

These results have major implications for future research. Good-quality proxy maps can be useful for mapping broad-scale patterns in ecosystem services, *sensu* Egoh *et al*. (2008). However, error propagation means that any attempt to select priority areas for multiple ecosystem services based on proxies – even if based on broad definitions of ecosystem services (e.g. Egoh *et al*. 2009) – is likely to result in a poor fit to actual data.

More generally, our results raise an interesting question: should we initiate policies to conserve natural capital assets based on available – but very imprecise – estimates of ecosystem services from proxy-based maps, or should we instead prioritize mapping the actual distribution of ecosystem services? Balmford & Gaston (1999) raised this question in the context of reserve selection for conserving biodiversity. They argue that, in most cases –provided it can be done quickly relative to the rate of any environmental degradation – it is more cost-effective to increase survey effort rather than use poor quality data for reserve selection. This is because the costs of surveys are usually far less than the costs of purchasing the extra land that is needed to compensate for the large error in the distribution of biodiversity that comes from using poor-quality data. It seems likely that the same is true for ecosystem services, and that the benefits of improved mapping will far outweigh the costs. This is because maps based on primary data will allow robust selection of priority areas for multiple services, whereas – as shown here – proxy-based maps can only be used for very coarse scale planning. More work is needed to explore the tradeoffs between the benefits of improving knowledge of ecosystem services versus the costs of inaction or the costs of misidentifying important areas for ecosystem services.

## Acknowledgements

This work was conducted during a U.K. Population Biology Network (UKPopNet) project ('Linking biodiversity and ecosystem services: processes, priorities and prospects') which was funded by the Natural Environment Research Council and Natural England. K.J.G. holds a Royal Society-Wolfson Research Merit Award. Many thanks to A. Darlow, M. Hill, G. Hinton, J. Hopkins, P. Ineson, D. Martin, and C. Swanwick for valuable discussions and comments. Many thanks also to R. Holland and M. Parnell for their help with the analyses, and to the many volunteers who provided the data that made up the biodiversity layer. Soil data were supplied under licence to the Centre for Terrestrial Carbon Dynamics by the National Soil Research Institute under a DEFRA funded project 'UK soil data base for modelling soil carbon fluxes and land use for the national carbon dioxide inventory'. This work is based on data provided through EDINA UKBORDERS with the support of the ESRC and JISC and uses boundary material which is copyright of the Crown. Finally, we thank the Associate Editor and two anonymous reviewers, whose comments greatly improved this manuscript.

## References

Anderson, B.J., Armsworth, P.R., Eigenbrod, F., Thomas, C.D., Gillings, S., Heinemeyer, A., Roy, D.B. & Gaston, K.J. (2009) Spatial covariance between biodiversity and other ecosystem service priorities. *Journal of Applied Ecology*, **46**, 888–896.

Anonymous (1994) *Biodiversity: the UK Action Plan*. HMSO Cm, London, UK.

Asher, J., Warren, M., Fox, R., Harding, P., Jeffcoate, G. & Jeffcoate, S. (2001) *The Millennium Atlas of Butterflies in Britain and Ireland*. Oxford University Press, Oxford.

Balmford, A. & Gaston, K.J. (1999) Why biodiversity surveys are good value. *Nature*, **398**, 204–205.

Bodin, O., Tengo, M., Norman, A., Lundberg, J. & Elmqvist, T. (2006) The value of small size: Loss of forest patches and ecological thresholds in southern Madagascar. *Ecological Applications*, **16**, 440–451.

Chan, K.M.A., Shaw, M.R., Cameron, D.R., Underwood, E.C. & Daily, G.C. (2006) Conservation planning for ecosystem services. *Plos Biology*, **4**, 2138–2152.

Chopra, K. (1993) The value of non-timber forest products – an estimation for tropical deciduous forests in India. *Economic Botany*, **47**, 251–257.

Costanza, R., d'Arge, R., deGroot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., O'Neill, R.V., Paruelo, J., Raskin, R.G., Sutton, P. & van den Belt, M. (1997) The value of the world's ecosystem services and natural capital. *Nature*, **387**, 253–260.

Daily, G.C. & Matson, P.A. (2008) Ecosystem services: from theory to implementation. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 9455–9456.

Eade, J.D.O. & Moran, D. (1996) Spatial economic valuation: benefits transfer using geographical information systems. *Journal of Environmental Management*, **48**, 97–110.

Egoh, B., Reyers, B., Rouget, M., Richardson, D.M., Le Maitre, D.C. & van Jaarsveld, A.S. (2008) Mapping ecosystem services for planning and management. *Agriculture Ecosystems & Environment*, **127**, 135–140.

Egoh, B., Reyers, B., Rouget, M., Bode, M. & Richardson, D.M. (2009) Spatial congruence between biodiversity and ecosystem services in South Africa. *Biological Conservation*, **142**, 553–562.

Eigenbrod, F., Anderson, B.J., Armsworth, P.R., Heinemeyer, A., Jackson, S.E., Parnell, M., Thomas, C.D. & Gaston, K.J. (2009) Ecosystem service benefits of contrasting conservation strategies in a human-dominated region. *Proceedings of the Royal Society B-Biological Sciences*, **276**, 2903–2911.

Gibbs, H.K. (2007) *Olson's Major World Ecosystem Complexes Ranked by Carbon in Live Vegetation: An Updated Data base Using the GLC2000 Land Cover Product*. Oak Ridge National Laboratory, Oak Ridge, TN, USA.

Guo, Z.W. & Gan, Y.L. (2002) Ecosystem function for water retention and forest ecosystem conservation in a watershed of the Yangtze River. *Biodiversity and Conservation*, **11**, 599–614.

Hein, L., van Koppen, K., de Groot, R.S. & van Ierland, E.C. (2006) Spatial scales, stakeholders and the valuation of ecosystem services. *Ecological zEconomics*, **57**, 209–228.

Hörnsten, L. & Fredman, P. (2000) On the distance to recreational forests in Sweden. *Landscape and Urban Planning*, **51**, 1–10.

Kremen, C., Williams, N.M., Bugg, R.L., Fay, J.P. & Thorp, R.W. (2004) The area requirements of an ecosystem service: crop pollination by native bee communities in California. *Ecology Letters*, **7**, 1109–1119.

Kreuter, U.P., Harris, H.G., Matlock, M.D. & Lacey, R.E. (2001) Change in ecosystem service values in the San Antonio area, Texas. *Ecological Economics*, **39**, 333–346.

Larsen, F.W., Petersen, A.H., Strange, N., Lund, M.P. & Rahbek, C. (2008) A quantitative analysis of biodiversity and the recreational value of potential national parks in Denmark. *Environmental Management*, **41**, 685–695.

Lawton, J.H. (2000) *Community Ecology in a Changing World*. Ecology Institute, Oldendorf/Luhe, Germany.

McGuire, A.D., Sitch, S., Clein, J.S., Dargaville, R., Esser, G., Foley, J., Heimann, M., Joos, F., Kaplan, J., Kicklighter, D.W., Meier, R.A., Melillo, J.M., Moore, B., Prentice, I.C., Ramankutty, N., Reichenau, T., Schloss, A., Tian, H., Williams, L.J. & Wittenberg, U. (2001) Carbon balance of the terrestrial biosphere in the twentieth century: Analyses of CO2, climate and land use effects with four process-based ecosystem models. *Global Biogeochemical Cycles*, **15**, 183–206.

Metzger, M.J., Rounsevell, M.D.A., Acosta-Michlik, L., Leemans, R. & Schrotere, D. (2006) The vulnerability of ecosystem services to land use change. *Agriculture Ecosystems & Environment*, **114**, 69–85.

Millennium Ecosystem Assessment (2005) *Ecosystems and Human Well-Being: Synthesis*. Island Press, Washington DC.

Milne, R. & Brown, T.A. (1997) Carbon in the vegetation and soils of Great Britain. *Journal of Environmental Management*, **49**, 413–433.

Naidoo, R. & Iwamura, T. (2007) Global-scale mapping of economic benefits from agricultural lands: implications for conservation priorities. *Biological Conservation*, **140**, 40–49.

Naidoo, R. & Ricketts, T.H. (2006) Mapping the economic costs and benefits of conservation. *Plos Biology*, **4**, 2153–2164.

Naidoo, R., Balmford, A., Costanza, R., Fisher, B., Green, R.E., Lehner, B., Malcolm, T.R. & Ricketts, T.H. (2008) Global mapping of ecosystem services and conservation priorities. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 9495–9500.

Natural England (2006) *England Leisure Visits – Report of the 2005 Survey*. Natural England Publications, Wetherby, UK.

Nelson, E., Polasky, S., Lewis, D.J., Plantingall, A.J., Lonsdorf, E., White, D., Bael, D. & Lawler, J.J. (2008) Efficiency of incentives to jointly increase carbon sequestration and species conservation on a landscape. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 9471–9476.

Nelson, E., Mendoza, G., Regetz, J., Polasky, S., Tallis, H., Cameron, D.R., Chan, K.M.A., Daily, G.C., Goldstein, J., Kareiva, P.M., Lonsdorf, E., Naidoo, R., Ricketts, T.H. & Shaw, M.R. (2009) Modeling multiple ecosystem services, biodiversity conservation, commodity production, and tradeoffs at landscape scales. *Frontiers in Ecology and the Environment*, **7**, 4–11.

Önal, H. & Yanprechaset, P. (2007) Site accessibility and prioritization of nature reserves. *Ecological Economics*, **60**, 763–773.

Orme, C.D.L., Davies, R.G., Burgess, M., Eigenbrod, F., Pickup, N., Olson, V.A., Webster, A.J., Ding, T.S., Rasmussen, P.C., Ridgely, R.S., Stattersfield, A.J., Bennett, P.M., Blackburn, T.M., Gaston, K.J. & Owens, I.P.F. (2005) Global hotspots of species richness are not congruent with endemism or threat. *Nature*, **436**, 1016–1019.

Plummer, M.L. (2009) Assessing benefit transfer for the valuation of ecosystem services. *Frontiers in Ecology and the Environment*, **7**, 38–45.

Preston, C.D., Pearman, D.A. & Dines, T.D. (2002) *New Atlas of the British and Irish Flora*. Oxford University Press, Oxford.

R Development Core Team (2008) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Sutton, P.C. & Costanza, R. (2002) Global estimates of market and non-market values derived from nighttime satellite imagery, land cover, and ecosystem service valuation. *Ecological Economics*, **41**, 509–527.

Troy, A. & Wilson, M.A. (2006) Mapping ecosystem services: practical challenges and opportunities in linking GIS and value transfer. *Ecological Economics*, **60**, 435–449.

Turner, W.R., Brandon, K., Brooks, T.M., Costanza, R., da Fonseca, G.A.B. & Portela, R. (2007) Global conservation of biodiversity and ecosystem services. *BioScience*, **57**, 868–873.
Wiens, J.A. (1989) Spatial scaling in ecology. *Functional Ecology*, **3**, 385–397.
Willemen, L., Verburg, P.H., Hein, L. & van Mensvoort, M.E.F. (2008) Spatial characterization of landscape functions. *Landscape and Urban Planning*, **88**, 34–43.

*Received 4 May 2009; accepted 12 January 2010*
*Handling Editor: Jeremy Wilson*

## Supporting Information

Additional supporting information may be found in the online version of this article:

**Data S1**. Detailed methods outlining the creation of the primary data and proxy maps.

**Table S1.** Assignment of LC2000 land cover map to 12 land cover classes

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.