50 YEARS WITH IMPACT
Journal of Applied Ecology
1913 British Ecological Society 2013 CELEBRATING 100 YEARS

Journal of Applied Ecology 2014 doi: 10.1111/1365-2664.12335

Updating conservation priorities over 111 years of species observations

Austin W. Milt*[1], Sally R. Palmer[2] and Paul R. Armsworth[1]

[1]*University of Tennessee, 569 Dabney Hall, 1416 Circle Drive, Knoxville, TN 37996, USA; and* [2]*The Nature Conservancy, 2021 21st Ave, South Suite C-400, Nashville, TN 37212, USA*

Summary

1. Observations of species occurrences are often used to inform spatial prioritizations for the effective use of limited conservation resources. Additional species observations have the potential to change where a conservation group plans to invest. But by how much? How different would conservation priorities be if planners updated current observations with the information they will have next year?

2. We sought to address these questions using a 111-year data set that reflects commonly used collection and prioritization practices. We quantify changes in the ranking of Tennessee watersheds brought on by annual additions of species observations made between 1900 and 2010. We ranked watersheds by their complementary contribution to overall species richness. We examine the sensitivity of our results to the number of watersheds prioritized.

3. We expected the effect of new observations to diminish as the data set grew, and we found this to be the case. Importantly, however, additional observations may continue to significantly change conservation priorities in the future if current data collection trends continue.

4. We found that, overall, additional observations can greatly affect priorities and that this result is sensitive to the number of watersheds ranked. Thus, the extent of planning activities moderates the effect of including additional data.

5. *Synthesis and applications*. Long-term, opportunistically collected data of species locations are commonly used in conservation planning. We find that when using such data, additional species observations significantly affect subsequent priorities. This effect is most pronounced when data are sparse. As such, data collection should be a focus of very early conservation actions in new areas. Even in well-studied areas, however, additional observations may continue to change spatial priorities into the future, and so while data collection can decrease in well-studied areas, it should continue at a lower intensity. Our methods could also be used to determine the balance of data collection and conservation action in a new location.

Key-words: conservation planning, inventory, observed richness, reserve, site selection, surveys, uncertainty

Introduction

Spatial prioritization is prominent in the science (Brooks *et al*. 2006; Pressey *et al*. 2007; Watson *et al*. 2011; Coll *et al*. 2012; Rainho & Palmeirim 2013) and practice (Redford *et al*. 2003; Didier *et al*. 2009; Henson *et al*. 2009; Groves *et al*. 2012) of conservation. This is due to the long history of spatial conservation planning methods (Diamond 1975; Margules & Pressey 2000; Watson *et al*. 2011), to the recognition that conservation funding is limited (Ando 1998; Myers *et al*. 2000; Bottrill *et al*. 2008) and to the spatial nature of many conservation decisions (Pressey *et al*. 2007). Prioritization methods vary depending on the conservation goals, expertise and data available to planners. Priorities may be determined by local species richness or biodiversity uniqueness (Csuti *et al*. 1997; Myers *et al*. 2000), by metrics of threat (Pressey *et al*. 2007; Joseph, Maloney & Possingham 2009; Carwardine *et al*. 2012) or by many other factors. This variation in method, along with variation in the data used in a particular evaluation, can lead to very different decisions about where, when and how to take action (Wilson *et al*. 2005; Rondinini *et al*. 2006).

Data on species occurrences are commonly used in conservation prioritization (e.g. Zafra-Calvo *et al*. 2010; Mateo *et al*. 2013; Simaika *et al*. 2013). Often, such data

*Correspondence author. E-mail: austin.milt@utk.edu

sets change in extent, resolution, accuracy and coverage as more observations are added (e.g. Magurran *et al.* 2010; Ahrends *et al.* 2011; Felinks *et al.* 2011; Martin, Blossey & Ellis 2012). Spatial priorities will be affected by data characteristics, such as spatial resolution (Araujo *et al.* 2005; Arponen *et al.* 2012), type (e.g. presence/absence data vs. abundance data; Gaston & Rodrigues 2003) and bias (De Ornellas, Milner-Gulland & Nicholson 2011; Metcalfe *et al.* 2013). Past studies looking at the effect of changing data on the outcome of conservation planning have tended to stylize the spatio-temporal extent and resolution of data used in conservation (Freitag & van Jaarsveld 1998; Polasky & Solow 2001; Felinks *et al.* 2011) and ignore the somewhat opportunistic nature of data being used by many conservation practitioners.

In this paper, we examine how additional species occurrence records affect spatial conservation priorities. In so doing, we focus on species-centric conservation approaches, as opposed to focusing on conservation goals targeting priority habitats or whole ecoregions (Watson *et al.* 2011). Tennessee, which we use as a case study, is a centre of richness for freshwater fish species and molluscs and a region within the coterminous United States of particularly high species imperilment (Dobson *et al.* 1997; Stein, Kutner & Adams 2000). We examine how annual additions of species observations made from 1900 to 2010 would change the ranking of watersheds being prioritized for conservation action. Specifically, we rank watersheds by complementary richness. As a conservation objective, complementary richness rewards watersheds for covering species not found in other protected watersheds (Vane-Wright, Humphries & Williams 1991). We use a rank correlation statistic to quantify the change in priorities brought on by an additional year's observations. Further, we assess the magnitude, trend and consistency of priority changes over time. We examine the sensitivity of our results to the number of watersheds prioritized. In the Supporting Information, we also explore the sensitivity of our results to ranking method, spatial or taxonomic sampling bias, and changes in data reliability due to changing technology and organism or population persistence.

The data set we use is the one currently used by conservation planners in Tennessee. As is often the case with data sets built from historical occurrence records, this one has been collated in a piecemeal and somewhat opportunistic fashion. As a result, the data set suffers from more spatial, temporal and taxonomic sampling bias than in systematic surveys. Arguably, it still represents the best information available to conservation planners regarding the distribution of priority species in Tennessee today.

Previous studies with similar methodologies to ours focus on the effectiveness of conservation outcomes under different data quality and quantity scenarios (Freitag & van Jaarsveld 1998; Polasky & Solow 2001; Gaston & Rodrigues 2003; Gladstone & Davis 2003; Grantham *et al.* 2008; De Ornellas, Milner-Gulland & Nicholson 2011). The general approach in empirical studies has been to aggregate data over time, to simulate changes to the data (e.g. by subsampling to represent reduced sampling effort) and to evaluate conservation plans on the altered data set (but see Felinks *et al.* 2011). Doing this allows authors to cover a range of data collection scenarios in order to generalize across many situations that might be encountered. These studies conclude that data quality and quantity are important factors in taking effective conservation actions, but the details are data set specific. For example, Grantham *et al.* (2009) assess how switching from initial species surveys to habitat protection affects the long-term coverage and retention of proteas. They find that for their case study area, a shorter duration of surveying (*c.* 2 years) followed by longer protection is optimal. Their study has important implications for conservation planning since it indicates that long-term data collection need not preclude conservation actions.

We complement previous work in many ways. First, we use a much longer-term data set spanning 111 years. Secondly, we build the data set sequentially over the time period rather than subsampling without regard to time and thus, we follow the actual collection of species observations. As such, our analysis includes the co-variation of data characteristics over time. Thirdly, our analyses do not rely on the aggregate data set as the most accurate knowledge of species distributions over time. Rather, we focus on describing how changing knowledge over time affects conservation priorities. Finally, we use raw species observations as they were recorded rather than modelled data or controls for data biases. We briefly explore how reducing among-watershed sampling bias affects our results (see Appendix S1 in Supporting information).

Materials and methods

CASE STUDY AREA

Tennessee is one of the most biodiverse inland states in the United States, second only to Alabama in the diversity of freshwater fishes and possessing a comparatively high degree of species endemism (Stein 2002). Over 10% of the state's plant and animal species are considered at-risk, and Tennessee ranks seventh among all states in the number of documented extinctions, a fact largely attributed to the major modification of streams and river systems in the early to mid-20th century (Stein, Kutner & Adams 2000; Stein 2002). Widespread conversion of lands for agricultural purposes has also contributed to fundamental changes in hydrologic regimes in many subregions of the state, and excess nutrients and sedimentation from agricultural production contribute to degraded water and habitat quality (Tennessee Department of Environment & Conservation 2014). Increased urbanization within the state's Metropolitan Statistical Areas has resulted in destruction and fragmentation of terrestrial habitats and degradation of streams and wetlands.

Local, national and international conservation organizations such as The Nature Conservancy and the World Wildlife Fund have invested in Tennessee for over 35 years in collaboration with many partners, including federal agencies such as the U.S.

Fish and Wildlife Service and state agencies such as the Tennessee Natural Heritage Program and the Tennessee Wildlife Resources Agency. Foundational to this work have been a series of conservation plans designed at ecoregional scales and using species occurrence data to set biodiversity conservation goals (Smith *et al.* 2002; The Nature Conservancy 2006). Beginning in 2005, all state wildlife agencies receiving federal State Wildlife Grant funding were required to submit a Comprehensive Wildlife Conservation Strategy, more commonly known as a State Wildlife Action Plan. The primary emphasis of State Wildlife Action Plans is to improve the habitat and population conditions of 'species of greatest conservation need' (SGCN) as defined by the state (Tennessee Wildlife Resources Agency 2005). Designing and executing these plans has resulted in an increased emphasis on the use of field-collected species occurrence and habitat data to identify priority conservation geographies and assess threats to these areas.

PRIORITIZATION DATA

We used species observation data collected between 1900 and 2010 by Tennessee Natural Heritage Program to test how species observations affect conservation priorities. The Natural Heritage data set is used in multiple forms of conservation planning in Tennessee, and the State Wildlife Action Plan in particular (Tennessee Wildlife Resources Agency 2005). The 2005 Tennessee State Wildlife Action Plan used these species occurrence data in combination with NatureServe global and state rarity rankings, U.S. Fish and Wildlife Service federal status listings, and other available population status data to assign SGCN status. The species occurrence data have been used as a key component in mapping local richness to understand where high SGCN concentrations occur across the state. These same data have been used to assess the complementarity of larger ecological units for terrestrial and freshwater species.

The Natural Heritage data are opportunistically recorded point observations (element occurrences, EOs) of individual species, most of which have a NatureServe Conservation Status rank higher than S3 and some of which are regularly monitored (TN Natural Heritage Program, pers. comm.). Rarely, observations are made as a result of premeditated prediction and collection efforts (TN Natural Heritage Program, pers. comm.). Because of the nature of how these data have been collected, there is no measure of sampling effort embedded in the data set. The data set contains both unique and repeat observations. A unique EO represents the spatial location, species identity and date of an observation of an individual or group. Subsequent observations of the same individual or group are here called repeat EOs. The data set contains 17 586 unique EOs or 25 838 EOs including

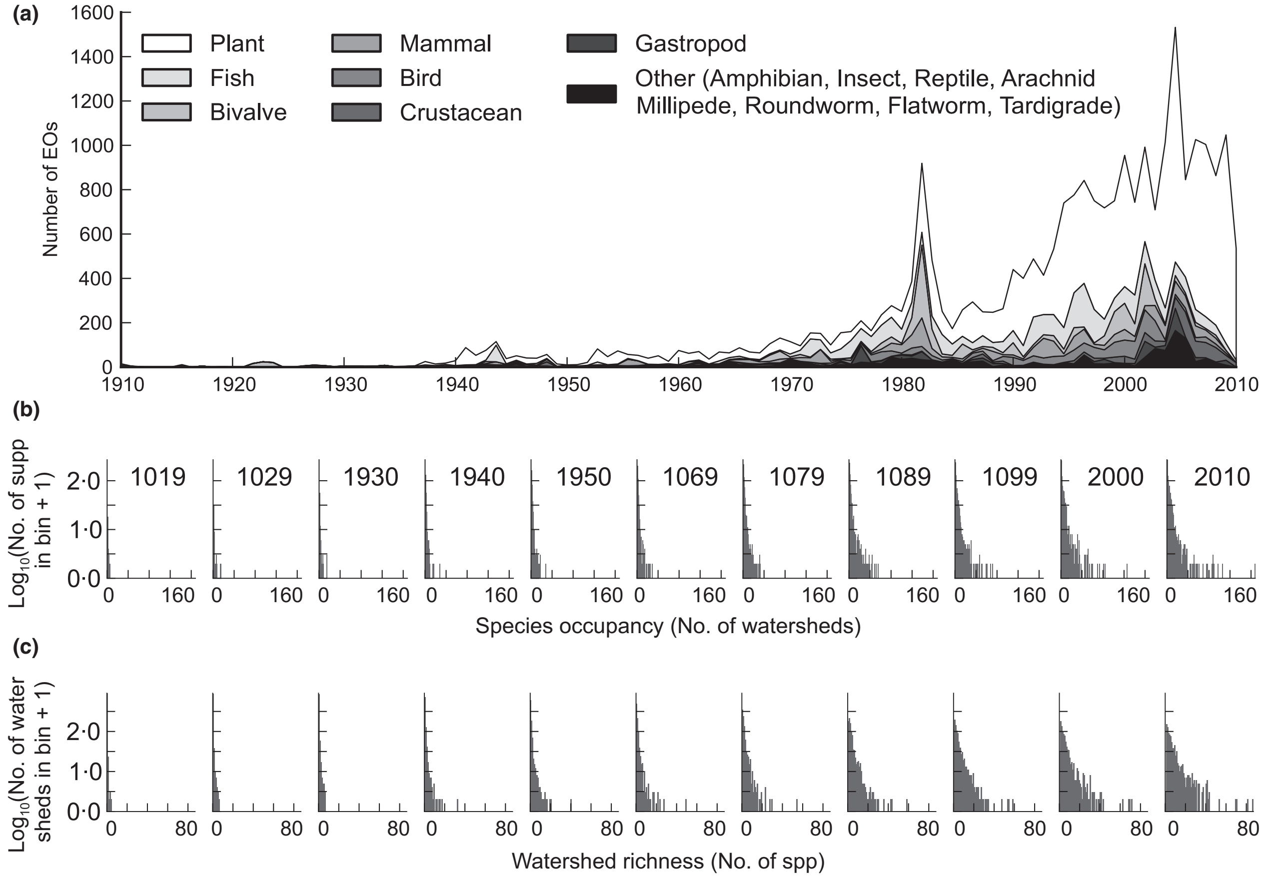

Fig. 1. Distribution of element occurrences (EOs) across watersheds and species over the 111-year period; *x*-axis is year. (a) Number of new EOs recorded throughout Tennessee each year broken down by taxonomic groups. An EO is an observed point location of an individual or population. There are 25 838 EOs in total, with the majority represented by Plant (15 001 or 58%). (b) and (c) show distributions as histograms of (b) species occupancies in Tennessee watersheds and (c) watershed richnesses at 11 time periods.

repeats. Fig. 1 shows the number of EOs recorded in each year including repeats. The data set contains both terrestrial and aquatic species. The majority of EOs recorded in the data set are of plants (15 001 records or 58%), although the data set also represents species from 14 other taxonomic groups recognized under the State Wildlife Action Plan. Approximately 51% of EOs were recorded after 1995.

Conservation organizations often make use of a mix of raw point observations and interpolated layers or modelled distributions (NatureServe, pers. comm.; The Nature Conservancy 2006; Schloss *et al.* 2011; Wilson 2011; Yorkshire Wildlife Trust, pers. comm.). Modelled data, such as those produced by species distribution models, are valuable because they estimate the unobserved range of a species and thus point out potential high-value areas not revealed by raw point occurrences. Also, species distribution models incur fewer omission errors and may reduce the effects of spatial sampling bias on conservation plans (Rondinini *et al.* 2006). On the other hand, raw point observations are simple to use and do not suffer as much from commission errors as species distribution models. Moreover, species distribution models cannot accurately estimate the ranges of very rare or under-sampled species (Olden, Jackson & Peres-Neto 2002; Wisz *et al.* 2008), a particular problem for our data set and similar contexts, because most species occur fewer than five times in the data (Fig. 1b, year 2010).

We used watersheds from the US Geological Survey HUC-12 Watershed Boundary Data set (http://datagateway.nrcs.usda.gov, accessed 9 May 2013) as our spatial unit of analysis (Fig. 2). Watersheds are an appropriate unit for spatial prioritization in conservation at the state level when focussing on terrestrial and aquatic species. At this scale, planners can target particular watersheds for further action. This action may come in the form of whole-watershed management (e.g. best practices by all farmers in the watershed), or it may call for more refined analysis to target protection within the watershed (e.g. protecting stream headwaters through forested-land conservation). For instance, watersheds have been used to delineate conservation priorities for known occurrences of freshwater species in the south-eastern United States (Smith *et al.* 2002) and elsewhere (Pryce *et al.* 2006). Watershed boundaries are also less changeable than other spatial units like land parcel boundaries and are therefore fitting for our century-spanning analysis. Of the 1152 watersheds in Tennessee, 925 have at least one observation by the Natural Heritage Program by 2010 (Fig. 2). These 925 watersheds acted as our candidate sites for selection and have a median area of 101 km^2 (1st quartile = 77 km^2, 3rd quartile = 142 km^2).

RANKING WATERSHEDS

We explored the situation in which a conservation group aims to cover as many species across the combined set of priority watersheds, rather than only the most species-rich watersheds. This requires consideration of how the species assemblages of watersheds complement one another. To implement this, we ranked watersheds based on their frequency in near-optimal solutions to the maximal coverage problem (MCP; Cabeza & Moilanen 2001). Given a watershed budget (b), the globally optimum solution to the MCP is the set of b watersheds that together cover more species than any other set of b watersheds. Solving the MCP gives a set of watersheds that together perform well as a conservation strategy. It does not automatically give a means to rank individual watersheds. However, Pressey, Johnson & Wilson (1994) introduced the idea of irreplaceability, defined as 'the frequency of occurrence of individual [watersheds] in the range of possible representative systems.' Irreplaceable watersheds are those that have a high potential to contribute to the conservation goal under many realized priority sets. We drew on this concept when deciding how to rank complementary watersheds.

The algorithm we used to rank watersheds has two parts. In the first part, we used a genetic algorithm optimizer to choose one set of b watersheds that maximizes the number of species covered. The difficulty of the MCP means that the genetic algorithm optimizer guarantees, at worst, locally optimum solutions. In conservation, the local optimality of solutions can be a strength, because finding many near-optimal solutions rather than the one best solution lends flexibility to the decision-making process. For a given budget, we ran the genetic algorithm optimizer 500 times and kept those solutions that achieved $\geq$95% of the richness of the best solution. In the second part, we ranked watersheds by the number of times they appeared across those top-scoring solutions (irreplaceability). In total, we examined 11 sensitivity tests corresponding to 11 watershed budgets (b): 1, 5, 10, 20, …, and 90 watersheds.

In Appendix S1 (Supporting information), we also explore the sensitivity of our results to other ranking methods. Namely, we assess prioritization based on the local richness of watersheds when ignoring complementarity in order to explore how sensitive our results are to the particular choice of conservation objective that we examine. We also assess variations on this local richness case to the number of watersheds prioritized, controls for data reliability and a control for sampling bias. All ranking was carried out in PYTHON v2.5.4, https://www.python.org/.

MEASURING CHANGES IN PRIORITIES

We measured by how much additional species observations cause watershed rankings to change. We used Spearman's rank correlation statistic, ρ, to measure the similarity of two different rankings of candidate watersheds for conservation, where the ranking of watersheds in a year is based on the observed assemblages of species in those watersheds. Our measure of priority change (V) is the difference between rankings, or

$$V_t = 1 - \rho\,(\text{ranking in year } t, \text{ranking in year } t-1).$$

We subtract the rank correlation from 1 to ease explanation such that larger values of V correspond to larger changes in priorities. Because ρ can take values between -1 and 1, V can vary between 0 and 2. We would typically expect V to take values between zero – where the rankings are identical – and one – where the two rankings have no relation to one another.

Below, we describe the entire process of ranking watersheds and calculating V with a budget of 10 watersheds:

1. Initialize: Using data from 1900 to 1909, the genetic algorithm optimizer chooses 10 watersheds that maximally cover present species. This is repeated 500 times. Watersheds are ranked by their frequency in the top 95% of the 500 solutions.
2. Update Records: Add records from the next year. In the first iteration, we added records from 1910 so that the irreplaceability and ranking of watersheds in 1910 were based on EOs from 1900

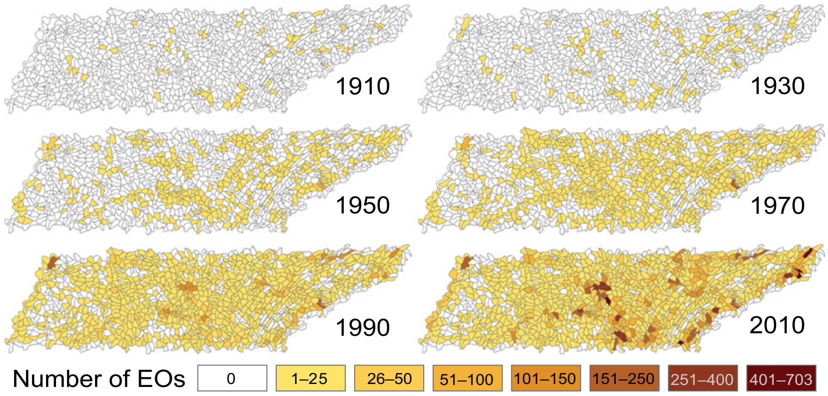

Fig. 2. Number of element occurrences (EOs), including repeats, accumulated in each watershed since 1900 at six snapshots in time. Darker colours have a higher density of EOs. For example, the darkest watershed in 1970 has 155 EOs which it accumulated between 1900 and 1970. Of the 1152 watersheds, 925 have at least one EO by 2010, 695 have ≤25 and 56 have ≥100.

to 1910. The ranking in the second iteration (1911) was based on EOs from 1900 to 1911 and so on.

3. Rank Watersheds: Repeat Step 1 with the updated data set from Step 2.

4. Calculate V_t: Calculate the difference between the rankings between the current year and previous year as defined above (1 − Spearman's ρ). This gives the magnitude of change in priorities for the current year. In the first iteration, we get V_{1910} by comparing the rankings from 1910 and 1909. In Fig. S2 (Supporting information), we summarize the result of delaying updates of records and priorities.

5. Repeat: Repeat Steps 2–4 through the year 2010, adding the most recent records for each iteration such that

$$V_{1910} = 1 - \rho\,(\text{ranking in 1910, ranking in 1909})$$
$$V_{1911} = 1 - \rho\,(\text{ranking in 1911, ranking in 1910})$$
$$\vdots$$
$$V_{2010} = 1 - \rho\,(\text{ranking in 2010, ranking in 2009})$$

STATISTICAL ANALYSES

To test if additional data have a significant effect on priorities, we tested whether the 5% confidence limit (5% CL) about the median of V contained zero. To calculate the 5% CL, we used bias-corrected accelerated bootstrapping from 10 000 samples of

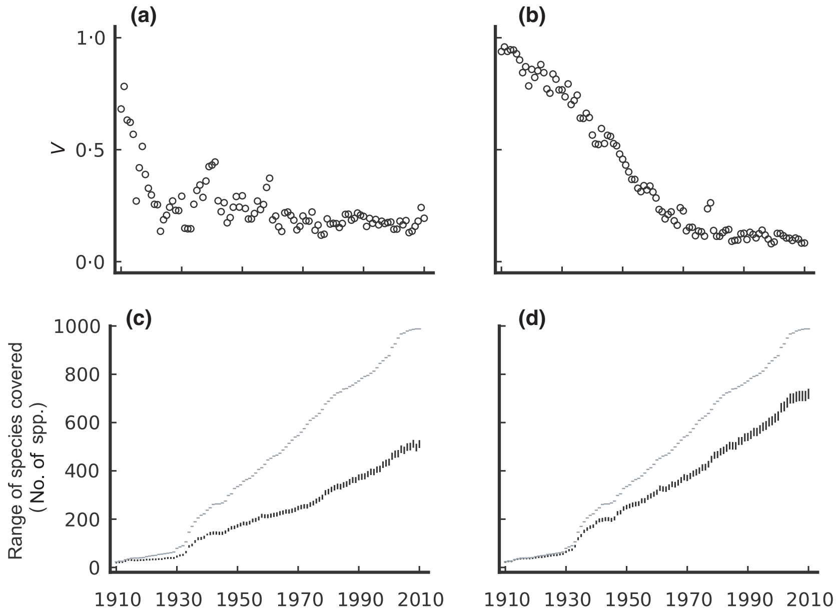

Fig. 3. Representative results of ranking watersheds by their frequency in near-optimal solutions to the maximal coverage problem (MCP; Cabeza & Moilanen 2001). Solutions to the MCP choose a fixed subset of all watersheds that maximally cover known species. Conservation priorities are sensitive to additional data, but that sensitivity declines with a smaller watershed budget and over time and eventually levels out. Panels (a) and (c) correspond to rankings created by choosing 20 watersheds to solve the MCP; panels (b) and (d) correspond to a 70-watershed solution. (a), (b) effect of 1 year of additional element occurrences, from 1910 to 2010. (c), (d) range of species covered by solutions to the MCP. Bottoms and tops of vertical, lower black bars are minimum and maximum number of species covered by solutions to the MCP. Horizontal, higher grey bars are the number of species with at least one occurrence by that year.

Table 1. Effect of additional data on priorities overall and when regressed against time. Conservation priorities are sensitive to additional data, but that sensitivity declines with a smaller watershed budget and over time and may level out in the future. Sensitivity of the ranking method to the number of sites used in the prioritization was tested. Subsequently, (column 2) the overall magnitude of change brought on by 1 year of additional element occurrences (EOs) was tested, as well as (columns 3–8) the trend of changing priorities over time. Finally, the regression models were used to predict the change in priorities from additional EOs in (column 9) 2010 and (column 10) 2030 if current collection conditions hold. Sensitivity tests without [$slope_2$] and [break] fields use a one-segment regression

Test	Med(V)	Intercept	$Slope_1$	$Slope_2$	Break	R^2	ΔAICc	V_{2010}	V_{2030}
1-watershed	0	–	–	–	–	–	–	–	–
5-watersheds	**0·14**	−1·0	0·6	–	–	0·03	2	**0·17**	**0·18**
10-watersheds	**0·18**	0·2	0·6	–	–	−0·01	1	**0·19**	**0·19**
20-watersheds	**0·20**	84·6	**−43·9**	−1·3	1920	0·73	83	**0·16**	**0·13**
30-watersheds	**0·19**	40·4	**−20·7**	−1·5	1937	0·88	118	**0·13**	0·10
40-watersheds	**0·22**	37·9	**−19·4**	−1·3	1944	0·94	165	**0·14**	**0·12**
50-watersheds	**0·21**	35·0	**−17·8**	−0·9	1954	0·97	199	**0·14**	**0·12**
60-watersheds	**0·20**	29·8	**−15·1**	−0·1	1965	0·98	203	**0·14**	**0·14**
70-watersheds	**0·31**	27·6	**−13·9**	−1·7	1971	0·98	177	**0·09**	0·06
80-watersheds	**0·37**	24·9	**−12·5**	−2·0	1977	0·99	173	**0·08**	0·04
90-watersheds	**0·40**	23·8	**−12·0**	−2·3	1978	0·99	146	**0·09**	0·05

Table columns are [Med(V)] = median of V across all years; [intercept] = model intercept; [$Slope_{1(2)}$] = slope of the first (second) segment in $V \cdot year^{-1} \cdot 10^{-3}$; [Break] = breakpoint (year) of the two segments; [R^2] = adjusted R^2; [ΔAICc] = ΔAICc of the piecewise model not chosen.; [V_{2010}] = predicted value of V in 2010; and [V_{2030}] = predicted value of V in 2030.
Values in bold are significant at the 95% confidence level.
[Correction added after first online publication on 31 Oct 2014: regression statistics amended].

the same size as the original sample (usually 101 data points, one from each year within 1910–2010). Bootstrapping was performed in MATLAB r2012b (MathWorks, Inc., Natick, MA, USA).

We examined trends in priority changes over time or data set size using ordinary least squares regression. In this data set, time and $\log_{10}$(data set size) are highly collinear ($R^2 = 0·99$), indicating they cannot be included in the same regression (Quinn & Keough 2002). Therefore, we regressed priority changes against time and data set size separately. Due to the similarity of results when regressing against time or data set size, we focus on regressions in time in the main text (but see Table S1, Supporting information for data set size results). We expected changes in priorities to decrease as we accumulated data. To confirm this, we tested that the 95% confidence intervals of the slopes of the above regressions are negative.

For some of the watershed budgets, we observed that the data could be clearly separated into two distinct segments (e.g. Fig. 3a,b and Fig. S1, Supporting information). Thus, we created piecewise regressions using the *segmented* package in R v2.12.1 to further examine the results (Muggeo 2008). The piecewise regression optimization is sensitive to initial guesses of breakpoints, so we visually estimated the breakpoints and then used the piecewise model with the convergence closest to our estimations (Muggeo 2008). For each sensitivity test, we tested piecewise models with one or two segments and compared AICc scores to determine which offered the more parsimonious fit to the data. AICc model comparison explicitly considers the trade-off between model fit using maximum likelihood and parsimony through the number of parameters (Crawley 2007).

We also wanted to know if changes in priorities might be sustained over the near future. This was primarily determined by interpolating the predicted value of V and its 5% significance in 2010 using the one or two-segment model chosen by AICc competition. We also extrapolated to 2030 to assess a more distant level of change if current conditions hold.

Results

We found that spatial conservation priorities are generally sensitive to additional data and the number of watersheds in which conservation can take place (Table 1: column 2). We also found that the sensitivity of priorities decreased over time (Table 1: columns 4 & 5; Fig. 3a,b), but that the trend levelled out after some time and the point at which this occurred depended on the watershed budget (Table 1: columns 4, 5 & 6; Fig. 3a,b). Finally, we found that additional data may continue to affect priorities in the future if current collection conditions hold (Table 1: columns 9 & 10).

Taking two exemplar budgets from those shown in Table 1 for illustrative purposes (20 watersheds and 70 watersheds), Fig. 3a,b illustrates the magnitude of change in priorities over time and Fig. 3c,d shows the number of species covered over time indicating the performance of the conservation planning process (see also Fig. S3, Supporting information). In Fig. 3c,d, the lower set of vertical, black bars shows the numbers of species covered by solutions to the MCP over time. Compare this to the higher, horizontal grey bars, which show how many species were known to occur in Tennessee in that year. The height of the lower bars relative to the higher bars shows how well the prioritization performed in each year. A comparison between Figs 3c and 3d shows that a watershed budget of 70 led to greater coverage of known species than using a lower budget of 20 watersheds. We also assess the performance of prioritizations relative to the full data set in Fig. S3 (Supporting information).

In the first part of our analysis, we wanted to know if, in general, 1 year of additional EOs affects conservation priorities. Provided more than one watershed is being considered for conservation action, we found that 1 year of additional EOs significantly changed the ranking of watersheds (Table 1: second column). Values in the second column of Table 1 reveal the magnitude of overall changes in priorities brought on by additional EOs, with larger values indicating more change per year of additional EOs. Note that the overall magnitude of change brought on by additional EOs increases with the watershed budget (moving down second column in Table 1).

Next, we assessed how changes in priorities changed over time. Changes in priorities over time were pronounced (Fig. 3a,b, Table 1, Fig. S4, Supporting information maps the spatial changes in priorities through time). The exemplar budgets shown in Fig. 3a,b reveal how large changes in priorities near the beginning of the time period are followed by a steep decline earlier and steadier decline later, but differ in how steep that initial decline is and when it occurs. These differences are quantified in Table 1. The fourth and fifth columns in Table 1 show the main results of our analysis of trends in changing priorities. In eight sensitivity tests, the degree to which priorities changed with 1 year of additional EOs decreased over time (Table 1: column 4). In seven of those eight tests, this result was true in both the first and second time periods (Table 1: columns 4 & 5).

We also explored in more detail how trends in changing priorities differed between the first and second time periods in each sensitivity test. We did this by comparing the values in the fourth and fifth columns of Table 1 and illustrate this in Fig. 3a,b. There was support in most cases for distinguishing the two time periods evident in each sensitivity test, as can be seen in Fig. 3a,b and tested by AICc competition in column eight in Table 1. When using a watershed budget larger than 10, there was a larger negative trend in changing priorities in the earlier time period than the later (Fig. 3a,b, Table 1: columns 4 & 5). Last, as we increased the watershed budget, the changing effect of additional EOs on priorities decreased more gradually in the first time period (moving down fourth column in Table 1).

Finally, we tested if the continuing change in priorities apparent in Fig. 3a,b might continue into the future if data collection conditions persist. The results of this analysis can be seen in the last two columns of Table 1. Bold values in the last two columns of Table 1 reveal that under several sensitivity tests, we expect additional EOs to continue to affect priorities if collection conditions persist. For instance, when we used a watershed budget of 60, we predicted an effect of 1 year of EOs in both 2010 and 2030. Contrarily, in four cases, additional EOs affected priorities in 2010 (Table 1, column nine) but are not predicted to do so in 2030 (Table 1, last column). The even split in the results of this part of our analysis, along with the fact that we performed extrapolation, makes it unclear how common the continuing effect of additional EOs will be in the future for our study system.

In Appendix S1 (Supporting information), we describe results for other sensitivity tests including our choice of conservation objective and controls for data reliability and sampling bias.

Discussion

How do additional observations of species change spatial conservation priorities? The importance of this question should be evident by the growing spatial prioritization literature (Kukkala & Moilanen 2013) as well as the ongoing use of opportunistically collected EOs for prioritization. We addressed this question by examining how the ranking of watersheds in the U.S. state of Tennessee changed as species observations were recorded over the past century. We ranked watersheds by their complementary contribution to conserving species richness and assessed how our results were affected by the number of watersheds considered for conservation action. Our methods and results can give insight into state-level prioritizations for watershed actions that focus on across-watershed complementary richness, to areas early in data collection and to those with a long history of species observations.

Perhaps our most important finding is that when additional data have a significant effect on priorities in 2010, additional data are also likely (54% of cases) to have a significant effect in the future if data collection conditions hold. Because our complementarity analysis uses species identities to determine priorities, small additions of infrequently occurring species have a larger effect on priorities than when prioritizing by local species richness only (e.g. Appendix S1, Fig. S1, Supporting information), and we expect this to persist in the future as long as data collection conditions continue.

Another key take-home message from our analysis is that additional observations tended to have a decreasing effect on priorities as we amassed data. While this phenomenon is well documented in studies focusing on the accuracy of prioritizations (Grantham *et al.* 2008), we offer a novel data context and explanation for why this occurs. The causal mechanism is due to the complementarity goal we used: additional observations increased the evenness of assemblages across watersheds, lowering the probability that one more observation in a watershed made that watershed necessary for a high-richness solution.

Our results have multiple competing consequences for conservation. In cases similar to ours, new observations may continue to significantly determine current spatial priorities and thus should be collected for that purpose. At the same time, current prioritizations will not necessarily match those for next year, so we do not expect additional observations to determine, on their own, long-term priorities. Our results also suggest that initial observations in data-poor regions will have the greatest effect on determining priorities. Subsequent observations then serve to refine those priorities. Finally, we found that our results were sensitive to conservation goals and to controls for data reliability and sampling bias (Appendix S1, Supporting information). Specifically, when ranking by local richness, the degree of decrease in priority changes over time was smaller as was the overall magnitude. Similarly, when ranking by local richness, future priorities are only expected to be affected by additional EOs when using a low tolerance for data reliability. The weaker effect of additional EOs on priorities when ranking by local richness is because the local richness objective ignores what is unique about different watersheds and thus misses influential variation in the data. Taking our results in aggregate, we suggest that the most effort in species observation be put forth early when a conservation group enters a new area. However, species observations should not cease since new observations will help refine priorities and update them as conditions change. In locations other than Tennessee, we expect similar patterns. The time in the collection record at which the qualitative shift from large changes in priorities with a rapid decrease in changes to smaller, but persistent changes in priorities – characterized in Fig. 3a,b – will depend on the conservation context. Managers in other locations could therefore use an analysis such as ours to determine the balance of data collection and conservation actions over time.

Measures of performance of conservation plans enable decision makers to assess whether more information is needed before acting (Polasky & Solow 2001), what management actions to take (Walters & Hilborn 1978), and to rank methods for creating new plans (Grantham *et al.* 2010). While we did focus on differences between priorities over time, the number of species covered by of choices of watersheds for conservation drove the prioritization process (Fig. 3c,d, Fig. S3, Supporting information). As such, our analysis is most similar to passive adaptive management (Walters & Hilborn 1978; Williams 2011), in that the decisions we make are refined as we gain information, but differs in that we do not assume that choices in one year affect those in the next year.

We necessarily made several decisions that may have affected our results. First, priorities were updated annually, which may partially explain small priority changes overall. Longer update periods increase the median change in priorities (Fig. S2, Supporting information). Secondly, we chose to use raw point occurrences rather than modelled data to prioritize watersheds. As a result, conservation actions focusing on the highest priority watersheds in any one year would be under-representative of potentially important areas (Rondinini *et al.* 2006). Aims to create comprehensive conservation plans should, when possible, use a mix of raw point occurrences and modelled data (Rondinini *et al.* 2006). That being said, we anticipate many of the effects we find will carry over to cases where practitioners are combining the two data types. Thirdly, our ranking method used two pieces of information to come up with relative rankings of watersheds: the spatial locations and species identities of EOs. This was done intentionally, so we could directly relate changes in species observations to changes in priorities. Additional information would increase the direct applicability of our results to Tennessee. For instance, the costs and patterns of land use change and management over time would have made apparent in priority setting the trade-off of these factors with species coverage. We could have also chosen to base our analyses on the unranked irreplaceabilities or local richnesses of watersheds rather than transforming to ranks first. Our assumption was that all decision variables required for a ranking of watersheds were encompassed in their irreplaceabilities and thus keeping additional information was unnecessary. In reality, conservation actions will rely on relationships among players, detailed site histories, short-term opportunity and other such information which is rarely recorded over such long time-spans. Finally, our statistical choices affect the inferences that can be drawn from our analyses. For example, the change in priorities in 1 year was calculated on two watershed rankings whose data sets overlapped substantially, and the overlap grew over time. Therefore, each ranking was not independent of the earlier rankings. As such, particular significance levels should be interpreted cautiously.

Prioritization is a common and necessary part of conservation planning. Here, we have provided insight into how regular updating of priorities is affected by additional data. Unlike previous studies, our study used a conservation-relevant data set that spans 111 years, which enabled us to explore long-term trends others could not. Additionally, we used simple prioritizations that did not account for data weaknesses. As such, our results may reflect practice more closely than other studies. Our results suggest that conservation planners can expect additional observations to alter priorities when conservation goals are complementarity-based.

Acknowledgements

We thank C. Dumoulin, R. Fovargue, G. Iacona, E. Larson, N. Sutton, and S. Ward for helpful feedback and J. Wisby and S. Mathes for help acquiring and processing species distribution data. Species distribution data were provided by Tennessee Natural Heritage Program. This work was completed while AWM was funded by a National Science Foundation fellowship: Scalable Computing and Leading Edge Innovative Technologies and a separate assistantship by the Department of Ecology & Evolutionary Biology at the University of Tennessee, Knoxville.

References

Ahrends, A., Rahbek, C., Bulling, M.T., Burgess, N.D., Platts, P.J., Lovett, J.C. *et al.* (2011) Conservation and the botanist effect. *Biological Conservation*, **144**, 131–140.

Ando, A. (1998) Species distributions, land values, and efficient conservation. *Science*, **279**, 2126–2128.

Araujo, M.B., Thuiller, W., Williams, P.H. & Reginster, I. (2005) Downscaling European species atlas distributions to a finer resolution: implications for conservation planning. *Global Ecology and Biogeography*, **14**, 17–30.

Arponen, A., Lehtomäki, J., Leppänen, J., Tomppo, E. & Moilanen, A. (2012) Effects of connectivity and spatial resolution of analyses on conservation prioritization across large extents. *Conservation Biology*, **26**, 294–304.

Bottrill, M.C., Joseph, L.N., Carwardine, J., Bode, M., Cook, C., Game, E.T. *et al.* (2008) Is conservation triage just smart decision making? *Trends in Ecology & Evolution*, **23**, 649–654.

Brooks, T.M., Mittermeier, R.A., da Fonseca, G.A.B., Gerlach, J., Hoffmann, M., Lamoreux, J.F., Mittermeier, C.G., Pilgrim, J.D. & Rodrigues, A.S.L. (2006) Global biodiversity conservation priorities. *Science*, **313**, 58–61.

Cabeza, M. & Moilanen, A. (2001) Design of reserve networks and the persistence of biodiversity. *Trends in Ecology & Evolution*, **16**, 242–248.

Carwardine, J., O'Connor, T., Legge, S., Mackey, B., Possingham, H.P. & Martin, T.G. (2012) Prioritizing threat management for biodiversity conservation. *Conservation Letters*, **5**, 196–204.

Coll, M., Piroddi, C., Albouy, C., Ben Rais Lasram, F., Cheung, W.W.L., Christensen, V. *et al.* (2012) The Mediterranean Sea under siege: spatial overlap between marine biodiversity, cumulative threats and marine reserves. *Global Ecology and Biogeography*, **21**, 465–480.

Crawley, M.J. (2007) *The R Book*. John Wiley & Sons, Ltd, Chichester, England.

Csuti, B., Polasky, S., Williams, P.H., Pressey, R.L., Camm, J.D., Kershaw, M. *et al.* (1997) A comparison of reserve selection algorithms using data on terrestrial vertebrates in Oregon. *Biological Conservation*, **80**, 83–97.

De Ornellas, P., Milner-Gulland, E.J. & Nicholson, E. (2011) The impact of data realities on conservation planning. *Biological Conservation*, **144**, 1980–1988.

Diamond, J.M. (1975) The island dilemma: lessons of modern biogeographic studies for the design of natural reserves. *Biological Conservation*, **7**, 129–146.

Didier, K.A., Glennon, M.J., Novaro, A., Sanderson, E.W., Strindberg, S., Walker, S. & Di Martino, S. (2009) The Landscape Species Approach: spatially-explicit conservation planning applied in the Adirondacks, USA, and San Guillermo-Laguna Brava, Argentina, landscapes. *Oryx*, **43**, 476–487.

Dobson, A.P., Rodriguez, J.P., Roberts, W.M. & Wilcove, D.S. (1997) Geographic distribution of endangered species in the United States. *Science*, **275**, 550–553.

Felinks, B., Pardini, R., Dixo, M., Follner, K., Metzger, J.P. & Henle, K. (2011) Effects of species turnover on reserve site selection in a fragmented landscape. *Biodiversity and Conservation*, **20**, 1057–1072.

Freitag, S. & van Jaarsveld, A.S. (1998) Sensitivity of selection procedures for priority conservation areas to survey extent, survey intensity and taxonomic knowledge. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **265**, 1475–1482.

Gaston, K.J. & Rodrigues, A.S.L. (2003) Reserve selection in regions with poor biological data. *Conservation Biology*, **17**, 188–195.

Gladstone, W. & Davis, J. (2003) Reduced survey intensity and its consequences for marine reserve selection. *Biodiversity & Conservation*, **12**, 1525–1536.

Grantham, H.S., Moilanen, A., Wilson, K.A., Pressey, R.L., Rebelo, T.G. & Possingham, H.P. (2008) Diminishing return on investment for biodiversity data in conservation planning. *Conservation Letters*, **1**, 190–198.

Grantham, H.S., Wilson, K.A., Moilanen, A., Rebelo, T. & Possingham, H.P. (2009) Delaying conservation actions for improved knowledge: how long should we wait? *Ecology Letters*, **12**, 293–301.

Grantham, H.S., Pressey, R.L., Wells, J.A. & Beattie, A.J. (2010) Effectiveness of biodiversity surrogates for conservation planning: different measures of effectiveness generate a kaleidoscope of variation. *PLoS One*, **5**, e11430.

Groves, C.R., Game, E.T., Anderson, M.G., Cross, M., Enquist, C., Ferdaña, Z. *et al.* (2012) Incorporating climate change into systematic conservation planning. *Biodiversity and Conservation*, **21**, 1651–1671.

Henson, A., Williams, D., Dupain, J., Gichohi, H. & Muruthi, P. (2009) The Heartland conservation process: enhancing biodiversity conservation and livelihoods through landscape-scale conservation planning in Africa. *Oryx*, **43**, 508–519.

Joseph, L.N., Maloney, R.F. & Possingham, H.P. (2009) Optimal allocation of resources among threatened species: a project prioritization protocol. *Conservation Biology*, **23**, 328–338.

Kukkala, A.S. & Moilanen, A. (2013) Core concepts of spatial prioritisation in systematic conservation planning. *Biological Reviews*, **88**, 443–464.

Magurran, A.E., Baillie, S.R., Buckland, S.T., Dick, J.M., Elston, D.A., Scott, E.M., Smith, R.I., Somerfield, P.J. & Watt, A.D. (2010) Long-term datasets in biodiversity research and monitoring: assessing change in ecological communities through time. *Trends in Ecology & Evolution*, **25**, 574–582.

Margules, C.R. & Pressey, R.L. (2000) Systematic conservation planning. *Nature*, **405**, 243–253.

Martin, L.J., Blossey, B. & Ellis, E. (2012) Mapping where ecologists work: biases in the global distribution of terrestrial ecological observations. *Frontiers in Ecology and the Environment*, **10**, 195–201.

Mateo, R.G., de la Estrella, M., Felicísimo, Á.M., Muñoz, J. & Guisan, A. (2013) A new spin on a compositionalist predictive modelling framework for conservation planning: a tropical case study in Ecuador. *Biological Conservation*, **160**, 150–161.

Metcalfe, K., Delavenne, J., Garcia, C., Foveau, A., Dauvin, J.C., Coggan, R., Vaz, S., Harrop, S.R. & Smith, R.J. (2013) Impacts of data quality on the setting of conservation planning targets using the species–area relationship. *Diversity and Distributions*, **19**, 1–13.

Muggeo, V.M.R. (2008) segmented: an R package to fit regression models with broken-line relationships. *R News*, **8**, 20–25.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B. & Kent, J. (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853–858.

Olden, J.D., Jackson, D.A. & Peres-Neto, P.R. (2002) Predictive models of fish species distributions: a note on proper validation and chance predictions. *Transactions of the American Fisheries Society*, **131**, 329–336.

Polasky, S. & Solow, A.R. (2001) The value of information in reserve site selection. *Biodiversity and Conservation*, **10**, 1051–1058.

Pressey, R.L., Johnson, I.R. & Wilson, P.D. (1994) Shades of irreplaceability: towards a measure of the contribution of sites to a reservation goal. *Biodiversity and Conservation*, **3**, 242–262.

Pressey, R.L., Cabeza, M., Watts, M.E., Cowling, R.M. & Wilson, K.A. (2007) Conservation planning in a changing world. *Trends in Ecology & Evolution*, **22**, 583–592.

Pryce, B., Iachetti, P., Wilhere, G., Ciruna, K., Floberg, J., Crawford, R. *et al.* (2006) *Okanagan Ecoregional Assessment, Volume 1 – Report*. The Nature Conservancy, Victoria, BC.

Quinn, G.P. & Keough, M.J. (2002) *Experimental Design and Analysis for Biologists*. Cambridge University Press, New York.

Rainho, A. & Palmeirim, J.M. (2013) Prioritizing conservation areas around multispecies bat colonies using spatial modeling. *Animal Conservation*, **16**, 438–448.

Redford, K.H., Coppolillo, P., Sanderson, E.W., Da Fonseca, G.A.B., Dinerstein, E., Groves, C. *et al.* (2003) Mapping the conservation landscape. *Conservation Biology*, **17**, 116–131.

Rondinini, C., Wilson, K.A., Boitani, L., Grantham, H. & Possingham, H.P. (2006) Tradeoffs of different types of species occurrence data for use in systematic conservation planning. *Ecology letters*, **9**, 1136–1145.

Schloss, C.A., Lawler, J.J., Larson, E.R., Papendick, H.L., Case, M.J., Evans, D.M. *et al.* (2011) Systematic conservation planning in the face of climate change: bet-hedging on the Columbia Plateau. *PLoS One*, **6**, e28788.

Simaika, J.P., Samways, M.J., Kipping, J., Suhling, F., Dijkstra, K.D.B., Clausnitzer, V., Boudot, J.P. & Domisch, S. (2013) Continental-scale conservation prioritization of African dragonflies. *Biological Conservation*, **157**, 245–254.

Smith, R.K., Freeman, P.L., Higgins, J.V., Wheaton, K.S., FitzHugh, T.W., Ernstrom, K.J., Das, A.A. (2002) *Priority Areas for Freshwater Conservation Action: A Biodiversity Assessment of the Southeastern United States*. The Nature Conservancy.

Stein, B.A. (2002) *States of the Union: Ranking America's Biodiversity*. NatureServe, Arlington, VA.

Stein, B.A., Kutner, L.S. & Adams, J.S. (eds.) (2000) *Precious Heritage: The Status of Biodiversity in the United States*. Oxford University Press, New York.

Tennessee Department of Environment and Conservation. (2014) *Year 2012 303(d) List: Final Version*. Nashville, TN.

Tennessee Wildlife Resources Agency. (2005) *Tennessee's Comprehensive Wildlife Conservation Strategy. TWRA Wildife Technical Report 05-08*. Nashville, TN.

The Nature Conservancy. (2006) *Standards for Ecoregional Assessments and Biodiversity Visions*. Arlington, VA.

Vane-Wright, R.I., Humphries, C.J. & Williams, P.H. (1991) What to protect?—Systematics and the agony of choice. *Biological Conservation*, **55**, 235–254.

Walters, C.J. & Hilborn, R. (1978) Ecological optimization and adaptive management. *Annual Review of Ecology and Systematics*, **9**, 157–188.

Watson, J.E.M., Grantham, H.S., Wilson, K.A. & Possingham, H.P. (2011) Systematic conservation planning: past, present and future. *Conservation Biogeography* (eds R.J. Ladle & R.J. Whittaker), pp. 136–160. John Wiley & Sons, Ltd., Chichester, UK.

Williams, B.K. (2011) Passive and active adaptive management: approaches and an example. *Journal of Environmental Management*, **92**, 1371–1378.

Wilson, C.R. (2011) *Documenting and Protecting Biodiversity on Land Trust Projects: An Introduction and Practical Guide* (ed. K. Mueller). Land Trust Alliance, Washington, DC.

Wilson, K.A., Westphal, M.I., Possingham, H.P. & Elith, J. (2005) Sensitivity of conservation planning to different approaches to using predicted species distribution data. *Biological Conservation*, **122**, 99–112.

Wisz, M.S., Hijmans, R.J., Li, J., Peterson, A.T., Graham, C.H., Guisan, A. & Group, N.P.S.D.W. (2008) Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, **14**, 763–773.

Zafra-Calvo, N., Cerro, R., Fuller, T., Lobo, J.M., Rodríguez, M.Á. & Sarkar, S. (2010) Prioritizing areas for conservation and vegetation restoration in post-agricultural landscapes: a Biosphere Reserve plan for Bioko, Equatorial Guinea. *Biological Conservation*, **143**, 787–794.

Received 27 February 2014; accepted 29 August 2014
Handling Editor: Brendan Wintle

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Appendix S1. Additional sensitivity test methods and results.

Fig. S1. Plots of V against time for all sensitivity tests.

Fig. S2. Changes in priorities as the time between priority updates increases.

Fig. S3. Performance of priorities at covering known species over time.

Fig. S4. Spatial distribution of watershed richnesses and irreplaceabilities over time.

Table S1. Median and regression results for V vs. dataset size.

