

Identifying Conservation Areas on the Basis of Alternative Distribution Data Sets

JARED G. UNDERWOOD,* CATERINA D'AGROSA, AND LEAH R. GERBER

Ecology, Evolution and Environmental Science, School of Life Sciences, Arizona State University, College & University Drive, Tempe, AZ 85287-1501, U.S.A.

Abstract: *Distribution data on biodiversity features is a major component of conservation planning that are often inaccurate; thus, the true distribution of each feature is commonly over- or underrepresented. The selection of distribution data sets may therefore lead to variability in the spatial configuration and size of proposed reserve networks and uncertainty regarding the extent to which these networks actually contain the biodiversity features they were identified to protect. Our goals were to investigate the impact on reserve selection of choosing different distribution data sets and to propose novel methods to minimize uncertainty about target attainment within reserves. To do so, we used common prioritization methods (richness mapping, systematic reserve design, and a novel approach that integrates multiple types of distribution data) and three types of data on the distribution of mammals (predicted distribution models, occurrence records, and a novel combination of the two) to simulate the establishment of regional biodiversity reserves for the state of Arizona (U.S.A.). Using the results of these simulations, we explored variability in reserve placement and size as a function of the distribution data set. Spatial overlap of reserve networks identified with only predicted distribution data or only occurrence distribution data never exceeded 16%. In pairwise comparisons between reserves created with all three types of distribution data, overlap never achieved 50%. The reserve size required to meet conservation targets also varied with the type of distribution data used and the conservation goal; the largest reserve system was 10 times the smallest. Our results highlight the impact of employing different types of distribution data and identify novel tools for application to existing distribution data sets that can minimize uncertainty about target attainment.*

Keywords: biodiversity, conservation planning, distribution data, MARXAN, species richness, systematic reserve design

Identificación de Áreas de Conservación con Base en Conjuntos Alternativos de Datos de Distribución

Resumen: *Los datos de distribución de biodiversidad son un componente principal de la planificación de la conservación que a menudo son imprecisos; por lo tanto, la distribución real de cada característica comúnmente es sobre o subrepresentada. Por lo tanto, la selección de conjuntos de datos de distribución puede llevar a variabilidad en la configuración espacial y tamaño de las redes de reservas propuestas e incertidumbre en relación con la representatividad de los atributos de la biodiversidad que se identificaron para protección. Nos propusimos investigar el impacto de diferentes conjuntos de datos sobre la selección de reservas y proponer métodos novedosos para minimizar la incertidumbre sobre el logro de metas dentro de las reservas. Para ello, utilizamos métodos comunes de priorización (mapeo de riqueza, diseño sistemático de reservas, y un método novedoso que integra múltiples tipos de datos de distribución) y tres tipos de datos de la distribución de animales (modelos de distribución predictivos, registros de ocurrencia y una novedosa combinación de los dos) para simular el establecimiento de reservas regionales de biodiversidad en el estado de Arizona (E.U.A.). Con los resultados de estas simulaciones, exploramos la variabilidad de la ubicación y tamaño de reservas como una función del conjunto de datos de distribución. El traslape espacial de las*

*email jared.underwood@asu.edu

Paper submitted August 11, 2008; revised manuscript accepted May 1, 2009.

redes de reservas identificadas solo con los datos de distribución predictivos o solo datos de ocurrencia nunca excedieron 16%. En comparaciones pareadas entre reservas creadas con los tres tipos de datos de distribución, el traslape nunca alcanzó 50%. El tamaño de la reserva requerido para alcanzar metas de conservación también varió con el tipo de datos de distribución utilizados y la meta de conservación; el sistema de reservas más extenso fue 10 veces mayor que el menos extenso. Nuestros resultados resaltan el impacto de la selección de diferentes tipos de datos de distribución e identifican herramientas novedosas para la aplicación de datos de distribución existentes que puede minimizar la incertidumbre sobre el cumplimiento de metas.

Palabras Clave: biodiversidad, datos de distribución, diseño sistemático de reservas, MARXAN, planificación de la conservación, riqueza de especies

Introduction

Conservation planning efforts generally aim to identify priority areas for conservation action to ensure the long-term persistence of biodiversity (Margules & Pressey 2000). Two of the most frequently used methods to identify these areas are species richness mapping and systematic conservation planning (Fleishman et al. 2006; Pressey et al. 2007). Regardless of the prioritization method, the degree to which conservation planners are effective in representing their targets largely depends on the data they possess or gather (Flather et al. 1997; Grand et al. 2007). Ideally, these data would consist of accurate information on the true distribution of all aspects of biodiversity being considered. Nevertheless, such data are rarely available; thus, conservation-planning decisions are made with incomplete data and target attainment for the reserve system is uncertain (Ferrier 2002; Wilson et al. 2005; Rondinini et al. 2006).

Occurrence records and predicted distribution models are often the data used to map biodiversity (Ferrier 2002; Wilson et al. 2005; Rondinini et al. 2006). Species occurrence records are documented observations of a species in a given location. Three major problems with species occurrence data are that sampling effort is not equal in all areas (spatial bias), effort is often biased toward certain species (taxonomic bias), and contemporary observations are rare for many species (temporal bias) (Smith et al. 2000; Wilson et al. 2005; Rondinini et al. 2006). These sampling issues can lead to underestimates of the true species distribution. In other words, an error of omission occurs because a species may incorrectly appear absent from a portion of its true distribution because the area has not been sampled. In addition, because many available occurrence records for the United States were collected in the early 20th century, the current distribution of a species may not be accurately reflected by these data (Rondinini et al. 2006). To overcome this temporal bias, new observation data would need to be collected. Nevertheless, limited funds often preclude the collection of new species occurrence data and thereby restrict understanding of current distributions (Smith et al. 2000; Fleishman et al. 2001).

Because of the uncertainty (spatial, temporal, and taxonomical biases) associated with species occurrence data, predicted distribution models based on parameters deemed necessary for species survival have been developed (Ferrier 2002; Rondinini et al. 2006; Rodriguez et al. 2007). Nevertheless, both the quality and availability of environmental data and the modeling techniques used can result in uncertainty, and can over- or underestimate a species' distribution (e.g., Loiselle et al. 2003; Vaughn & Ormerod 2005; Rodriguez et al. 2007). This uncertainty can lead to errors of commission. A commission error occurs when areas identified for protection fail to contain the conservation targets they were meant to protect (Rondinini et al. 2006). Despite the challenges associated with both types of distribution data, they continue to be used in the conservation planning process (Funk & Richardson 2002; Rondinini et al. 2006; Rodriguez et al. 2007).

The quantity and quality of data used in the planning process are known to significantly alter the outcome (e.g., Flather et al. 1997; Loiselle et al. 2003; Elith et al. 2006; Grand et al. 2007); however, the extent to which alternative distribution data sets identify similar areas in regional or local conservation planning efforts is not well understood. We did not attempt to identify a set of reserves for implementation; rather, we highlight the spatial disparity between areas identified for conservation when different types of distribution data are used. We also developed novel techniques to reduce uncertainty in meeting reserve-system conservation targets that take spatial constraints into account.

Methods

Data

We used data on 71 mammal species that inhabit terrestrial, aquatic, arboreal, and fossorial habitats in Arizona (U.S.A.). This suite of species was chosen because it represents a range of ecological niches, rarity, and threats (e.g., widespread vs. localized, threatened vs. secure, generalist vs. specialist) commonly considered when selecting a group of species on which to base reserve design

(Pearson 1994). Mammals are frequently used in conservation planning as a surrogate group (e.g., Yahnke et al. 1998; Ceballos & Ehrlich 2006; Forbosh et al. 2007).

Predicted distribution models for each species were developed as part of the Southwest Regional Gap Analysis Project. This modeling process used a Boolean approach and resulted in habitat being predicted where all the required attributes (land-cover type [e.g., vegetation community], elevation, slope, aspect, distance to perennial water, landform, soils, and hydrologic units) occurred together. For each species, information was obtained from the literature and species experts to identify the appropriate range of each attribute (Boykin et al. 2007). Thresholding for the models was performed as they were developed. If a location contained all attributes required by the species, then the species was assumed to occur at that location. This technique creates no probability of occupancy data for the locations identified. The level of accuracy associated with these data was assessed as part of Boykin et al. (2007) and was within the acceptable range for data sets of predicted distributions. We made no modifications to the data set.

Occurrence data consisted of 26,766 historic point locality records (1900–1999) compiled from various sources (e.g., state natural heritage programs, biological atlases, primary literature, and collections) by Bender et al. (2005). The level of accuracy associated with these occurrence data is variable; however, data points that could not be mapped with <5-km² accuracy were not included in the data set (Bender et al. 2005). We projected all data to Albers equal-area projection.

As is commonly the case, the occurrence records and predicted distribution models probably misrepresent the true distributions of our chosen species. Our objective, however, was to explore the consequences of using various types of distribution data, not to identify a set of reserves for implementation. Thus what was important was that all general data patterns for common surrogate species were represented, not the accuracy of the data. For example, in our data sets, rare or specialist species generally have fewer occurrence records and are predicted to have a smaller geographical distribution, and more common or more wide-ranging species normally have a greater number of occurrence records and predicted distributions that are more spatially extensive. Most importantly, these data are representative of those commonly available to conservation planners (Rondinini et al. 2006).

Dispersal-Limited Data

In an effort to balance the errors of commission and omission that may occur when assigning a species' distribution, we developed a hybrid of occurrence and predicted-distribution data sets (dispersal limited). To create a

dispersal-limited distribution, we included the location of all known occurrences and the predicted distribution surrounding each occurrence record within the species' dispersal capability. We used an average dispersal distance for each species. We based these averages on dispersal data from NatureServe (2008) species databases. Various dispersal distances, however, could be incorporated into this method depending on the risk aversion of the planner to committing an error of commission or omission. To minimize errors of commission, one could use the minimum known dispersal distance and to minimize errors of omission, one could use the maximum known dispersal distance.

The advantages of our dispersal-limited distribution data stem from the assumption that occurrence records are less likely to have errors of commission than predicted distribution data. If this is true, a dispersal-limited distribution controls errors of commission associated with predicted distributions by incorporating only the predicted habitat within the dispersal distance of an occurrence record. Additionally, the incorporation of surrounding predicted habitat limits errors of omission common to occurrence records by increasing a species' range to nearby locations with similar habitat.

Assigning Presence

We divided the state of Arizona into 25-km² planning units (PUs; 12,092) in ArcGIS (version 9, ESRI, Redlands, California). Each PU represented a single potential reserve location. Although assigning presence on a smaller resolution is possible, the selection of the 25-km² planning unit was recommended as appropriate by land managers in the planning area on the basis of acquisition and management requirements. Smaller parcels of land were seen as inefficient from managerial and ecological perspectives. Using each type of distribution data, we assigned presence or absence for every species to all PUs. For point localities (occurrence data), if the record was within a PU, the species was considered present. For dispersal-limited and predicted distribution data, if any portion of a species' predicted range was in the PU, then the species was considered present.

Unsurprisingly, there was a large discrepancy in the number of PUs containing each type of distribution data: 100% of the PUs contained some predicted distribution data, whereas only 33% contained occurrence distribution data, and only 51% contained dispersal-limited distribution data. These differences illustrate the disparity in spatial distribution of data inherent to the data sets.

Richness Mapping

Species richness has often served as a basis for identifying areas of conservation priority. For each type of data, we based identification of reserves that would protect species richness (i.e., richness reserves) on the

relative number of species within each PU (Orme et al. 2005; Grenyer et al. 2006). We defined richness reserves as the richest 5% of planning units for each type of distribution data (Reyers et al. 2000; Orme et al. 2005; Shriner et al. 2006). When identifying richness reserves for each type of data, we excluded PUs with no data.

Systematic Reserve Design

We used MARXAN (version 2.0.2), a reserve selection algorithm, to systematically design reserve networks (Ball & Possingham 2000). Within MARXAN we selected the commonly used simulated annealing process that generates the near-minimum set of PUs required to achieve conservation goals, also referred to as the “best” reserve network (Ball & Possingham 2000). Simulated annealing allows for multiple near-optimum solutions that can then be compared on the basis of financial, political, social, or other constraints. The relative conservation importance of each PU can be assessed on the basis of its selection frequency (irreplaceability; Warman et al. 2004; Shriner et al. 2006). Selection-frequency scores allowed us to identify and compare the location of PUs with high selection frequency across the different types of distribution data as an alternative to comparing the locations of the final, best reserve network.

To illustrate the influence of distribution data on reserve placement, we first analyzed the scenario in which the reserve network represented all species at least once at a minimal cost (Shriner et al. 2006; Grand et al. 2007). Cost in these scenarios is not reflective of economic or ecological value of each PU, but is simply the area required for the reserve network.

We ran the model under the assumption that presence in PUs was assigned on the basis of a single type of distribution data, either predicted, dispersal limited, or occurrence. When parameterizing the selection algorithm, we did not limit the area of the potential reserve network. We set the cost of PUs to one because we considered the potential value of each PU as equal. We set the species-penalty factor (penalty assigned to a reserve for not meeting conservation targets) at 10,000 to ensure that the reserve system met specified conservation targets for all species (Ball & Possingham 2000). We ran 1000 iterations for each scenario.

To better understand the potential impacts of reserve connectivity and aggregation, we ran a second set of scenarios with the same parameters as mentioned previously, but used the boundary length modifier tool in MARXAN to force aggregation and connectivity of the reserve systems (Ball & Possingham 2000). In our third set of scenarios, we assessed the impact of setting a relative goal (10% of total occurrences as dictated by each type of distribution data) as opposed to an absolute goal (10 locations for each species regardless of total possible).

We used the output from MARXAN for three analyses. First, we compared the locations for each of the best reserve networks. Second, we compared the locations of the top 10% of PUs that had a selection frequency greater than zero as identified for each type of distribution data (hereafter referred to as highly selected areas, not an inclusive reserve network). Third, we used the selection frequency results to assess the variance in reserve placement within a single type of distribution data.

Adaptive Method

To provide conservation planners with a tool to minimize the risk of committing commission errors while still considering spatial constraints, we developed a novel method that incorporates all three types of distribution data (predicted, occurrence, and dispersal limited), hereafter referred to as the adaptive method. The use of multiple types of distribution data can reduce commission error because occurrence and dispersal-limited distribution data are used to meet most targets. Nevertheless, some degree of spatial flexibility is retained because predicted distribution data is used to meet remaining targets.

To illustrate this technique, we set an absolute conservation target of 10 locations per species (Rodrigues & Gaston 2001; Pressey et al. 2003). We required that 50% (5) of conservation targets be met with distribution as identified by occurrence records, 30% (3) as identified by the dispersal-limited data, and 20% (2) as identified by the predicted distribution data. We used all three types of distribution data to assign presence to PUs. For each species, however, we allowed presence in a PU to be based on only one type of distribution data. When presence in a PU could be assigned with more than one type of distribution data, we used the data that supplied the most restricted range for that species. In all cases, the most restrictive type was occurrence data followed by dispersal-limited data, and the most widespread was predicted distribution data. This method allowed PUs to be selected for our reserve network on the basis of only one type of distribution data. We compared the results of the adaptive method with results from the standard procedure in which only one type of distribution data was used (either occurrence or predicted distribution data) as the basis for identifying 10 locations for each species.

We refer to the adaptive method as adaptive because planners can adapt how conservation targets are met on the basis of the importance they assign to spatial constraints or risk of commission error. We used a 5:3:2 (occurrence: dispersal limited: predicted) requirement to meet targets to illustrate the method, but if spatial constraints are the most important consideration, more targets could be met with predicted distribution data (2:3:5). Nevertheless, if risk of commission errors is the most

important consideration, more targets could be met with occurrence records or dispersal-limited distribution data (6:4:0). Additionally, if planners do not want to create or use dispersal-limited distribution data, the adaptive method can still be used with only occurrence and predicted distribution data given the same considerations (e.g., 5:5, 7:3, 3:7).

Analyses

We assessed the impact of alternative types of distribution data to conservation planning outcomes by examining the total number of PUs in each reserve network and spatial congruence between sets of reserves created with alternative types of distribution data or conservation goals. Given the effects of spatial autocorrelation, we calculated congruence with a modified Jaccard similarity coefficient (van Jaarsveld et al. 1998). We modified the standard formula to include two new terms: boundaries and potential boundaries. We defined *boundaries* as the number of occasions when a PU selected for the smallest set of reserves did not overlap a PU in the larger reserve network but bordered one or more PUs from the larger network. Each boundary was given a value of one-half. We defined *potential boundaries* as one-half the number of nonoverlapping PUs for the smallest set of reserves. If a PU was selected by both scenarios, then it received a value of 1 in our analysis. Therefore, our calculation for the Jaccard coefficient became: (number PUs with shared + number of boundaries $[\frac{1}{2}]$) / (number of additional PUs selected for scenario A + number of additional PUs selected for scenario B + number of potential boundaries). We performed pairwise comparisons on the following sets of best reserves and highly selected areas: occurrence versus predicted; dispersal limited versus predicted; dispersal limited versus occurrence; and results of our adaptive method with each type of distribution data.

To assess the variance in reserve placement within a single type of distribution data, we used the selection frequency scores. This allowed us to compare the spatial overlap across 1000 reserve networks selected with the same type of distribution data. To assess the variance, we determined the number of PUs selected 90% and 75% of the time and calculated what percentage of the corresponding best reserve network that number represented. Low variance would occur if a large percentage of the PUs in the corresponding best reserve network were selected more than 75% of the time.

Finally, we investigated whether PU selection for reserve systems depended on the type of distribution data. We used the Wilcoxon signed-ranks test to assess each pairwise comparison of reserve systems to determine whether PU selection, or frequency of selection, depended on the type of distribution data.

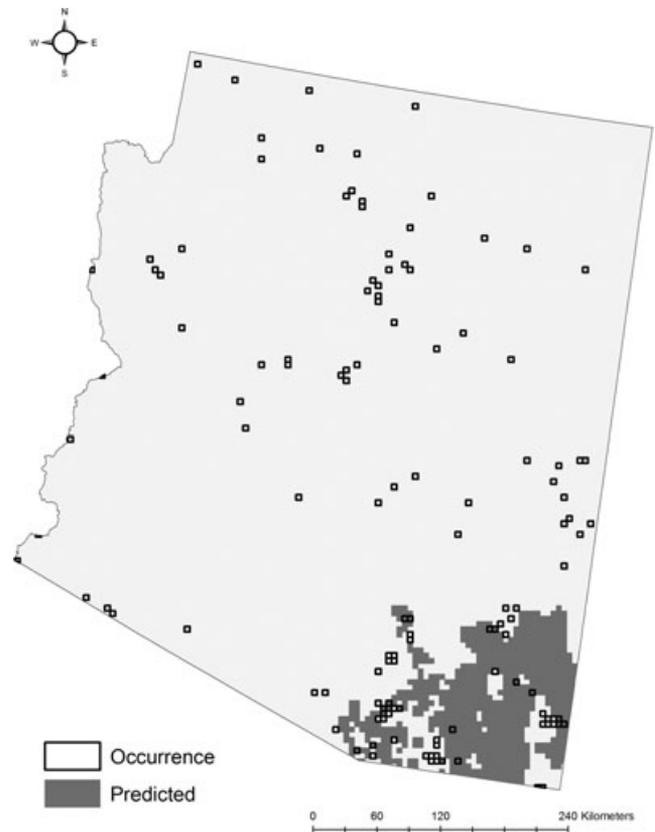


Figure 1. Visual representation of overlap in species-richness reserves in Arizona (U.S.A.) created with occurrence or predicted distribution data. Richness reserves were identified by selecting the richest 5% of planning units for each type of distribution data.

Results

Richness Reserves

The location of richness reserves depended on the type of distribution data available. Overlap between richness reserves created with predicted versus occurrence distribution data was only 3%. Richness reserves identified on the basis of only predicted data were clustered in the southeastern corner of the state and were more uniformly distributed when we used only occurrence data (Fig. 1).

Systematically Designed Reserve Networks

Best reserve networks and highly selected areas varied greatly in size and location, depending on the type of distribution data used. In scenarios with an absolute conservation goal, the use of occurrence data led to a reserve configuration that required the greatest total area to meet conservation goals (Table 1). When dispersal-limited data were used, an intermediate size for reserve

Table 1. Size of reserve networks (planning units) for hypothetical regional biodiversity reserves in Arizona (U.S.A.).

Scenario*	No. of occurrences protected per species	No. of planning units
Best reserves		
predicted	1	8
dispersal limited	1	14
occurrence	1	21
predicted	10	87
dispersal limited	10	90
adaptive	10	151
occurrence	10	218
predicted	10%	1185
dispersal limited	10%	368
occurrence	10%	112
Best reserves w/aggregation		
predicted	1	848
dispersal limited	1	1,999
occurrence	1	2,226
Highly selected areas		
predicted	1	1
dispersal limited	1	3
occurrence	1	3
predicted	10	9
dispersal limited	10	10
adaptive	10	35
occurrence	10	53
predicted	10%	1270
dispersal limited	10%	543
occurrence	10%	254
Richness reserves		
predicted	NA	763
occurrence	NA	126

*Best reserve networks, identified by MARXAN as requiring the least area; highly selected areas, top 10% of areas identified by the reserve-selection algorithm as necessary for the best reserve network; best reserves with aggregation, planning units selected for the best reserve network were required to be aggregated; richness reserves, planning units within the top 5% of species richness.

networks was discernible. At both absolute target levels, the smallest reserve networks were those created with predicted distribution data (Table 1). Spatial overlap between best reserve networks for absolute targets ranged from 0% to 47% (Table 2). When the conservation target was set at one location for each species, the greatest spatial overlap was observed between reserves created with occurrence and dispersal-limited data (Fig. 2; 13%), and no spatial overlap occurred between reserves created with predicted distribution data and other types of distribution data. Spatial overlap was higher between all reserve comparisons when the absolute target level was increased to 10 locations per species (5–47%). The highest degree of spatial overlap occurred between reserves created with occurrence data and our adaptive method (47%).

When we incorporated principles of aggregation and connectivity, we found a pattern of spatial overlap similar

Table 2. Pairwise comparisons of spatial overlap (Jaccard's coefficient) in reserve networks when different types of distribution data are used to identify best reserves or highly selected areas for species richness in Arizona (U.S.A.).

Comparison*	No. of occurrences protected per species	Jaccard's coefficient (%)
Highly selected areas		
predicted, occurrence	1	0
predicted, dispersal limited	1	0
dispersal limited, occurrence	1	14
predicted, occurrence	10	6
predicted, dispersal limited	10	6
dispersal limited, occurrence	10	9
adaptive, predicted	10	6
adaptive, occurrence	10	31
adaptive, dispersal limited	10	15
predicted, occurrence	10%	8
predicted, dispersal limited	10%	14
dispersal limited, occurrence	10%	28
Best reserves		
predicted, occurrence	1	0
predicted, dispersal limited	1	0
dispersal limited, occurrence	1	13
predicted, occurrence	10	6
predicted, dispersal limited	10	25
dispersal limited, occurrence	10	4
adaptive, predicted	10	35
adaptive, occurrence	10	47
adaptive, dispersal limited	10	40
predicted, occurrence	10%	3
predicted, dispersal limited	10%	8
dispersal limited, occurrence	10%	12
Best reserves w/aggregation		
predicted, occurrence	1	9
predicted, dispersal limited	1	15
dispersal limited, occurrence	1	23

*Best reserve networks, identified by MARXAN as requiring the least area; highly selected areas, top 10% of areas identified by the reserve-selection algorithm as necessary for the best reserve network; best reserves with aggregation, planning units selected for the best reserve network were required to be aggregated.

to the results of absolute conservation goals (9–23%). Nevertheless, the area requirements for such scenario were much larger than for absolute targets.

With a relative target (10%), the overlap in the best reserve systems was minimal (3–12%; Table 2). In a reversal from absolute targets, however, the use of predicted distribution data required the largest reserve system and the use of occurrence data the smallest (Table 1).

A comparison of the highly selected areas (see Methods) showed a lack of spatial overlap for all approaches when the target level was set at one (Table 2; 0–14%). Greater spatial overlap was apparent when the target was increased (10 locations, 10%; Table 2). Spatial overlap at this elevated conservation level was least apparent for reserves created with predicted and occurrence data (6% and 8%, respectively).

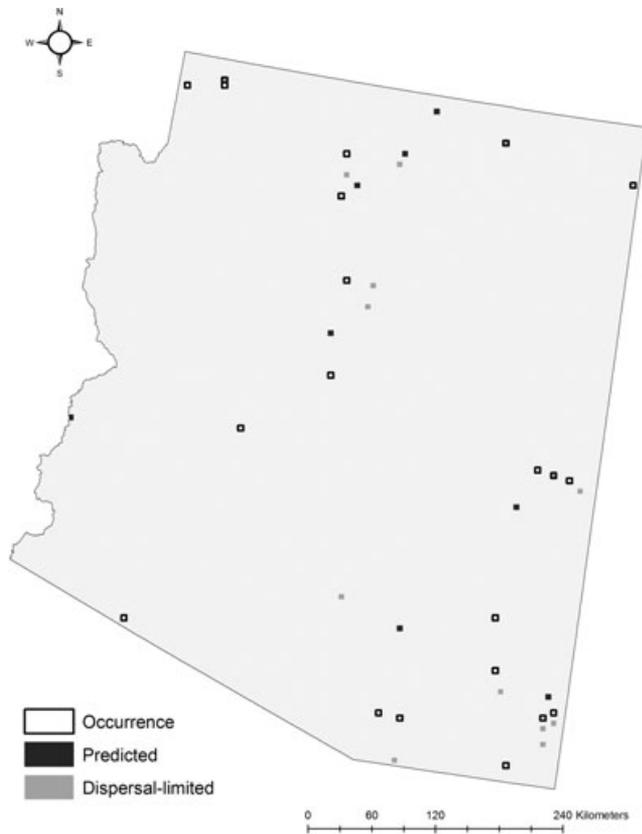


Figure 2. Visual representation of overlap in “best” reserve networks in Arizona (U.S.A.). Reserves identified with occurrence, predicted, or dispersal-limited distribution data. “Best” reserve networks were those identified by the reserve-selection algorithm MARXAN as the near optimal set of areas that protects one location for each species and minimizes area requirements.

Reserve Selection Variance

The overlap between reserve systems identified with a single type of distribution data was also low. For predicted distribution data, when the target was one location for each species, or 10% of a species' locations, no PU was selected more than 30% of the time. When the target was 10 locations for each species, nine PUs (10% of the corresponding best reserve size) were selected more than 90% of the time, and no other PUs were selected more than 20% of the time. For occurrence data, when the target was one location for each species, three PUs were selected more than half the time. When the target was 10% of the species' locations, two PUs were selected more than half the time. When the target was 10 locations for each species, 54 PUs (25% of the corresponding best reserve size) were selected 90% of the time and 75 PUs (34% of the corresponding best reserve size) were selected 75% of the time. For dispersal-limited data, when the target was one location for each species, or 10% of

a species' locations, no PU was selected more than 40% of the time. When the target was 10 locations for each species, 10 PUs (11% of the corresponding best reserve size) were selected 90% of the time and 12 PUs (13% of the corresponding best reserve size) were selected 75% of the time.

When we assessed which PUs were selected for inclusion in reserves, regardless of the conservation goal or which two types of distribution data were compared, the results were significantly different for each type of data ($p < 0.001$). When we analyzed the frequency with which PUs were selected for inclusion in each scenario, we also found significant data-dependent differences for all comparisons ($p < 0.001$).

Discussion

Data Coverage and Conservation Goals

We believe the patterns of spatial overlap between the reserve networks we documented are the result of interplay between two factors: conservation goals and the spatial coverage inherent to each type of distribution data. When the spatial coverage of distribution data was limited, flexibility in selecting PUs was reduced. Planning-unit selection can also be constrained as the conservation target for each species is increased. The impacts of data coverage and conservation target, as well as interactions between the two, were apparent in our analysis of reserve selection variance for each type of distribution data. Reserves identified with our occurrence data set (most limited coverage) showed the least variance in the PUs selected, whereas those from the predicted distribution data (broadest coverage) had the greatest variance. Across all types of distribution data, variance in which PUs were selected decreased as the absolute conservation target increased. Interaction between these two factors constrained the scenario that paired occurrence data (limited spatial coverage) with the highest absolute conservation target (10 locations) so that it had the least variance in which PUs were selected.

Although our results should be considered in the context of alternative types of distribution data, the underlying driver for the differences we observed appears to be the inherent discrepancy in spatial coverage within each data set. Our results show that the coverage patterns inherent to each type of distribution data set (predicted distribution = broad coverage, occurrence data = limited coverage, dispersal-limited distribution data = intermediate coverage) can result in disparity of reserve location.

We developed our dispersal-limited distribution data and adaptive method in response to the inherent differences in spatial coverage of each type of data. These methods present techniques for use with absolute conservation targets that maximize spatial coverage of the

species distribution data set yet minimize uncertainty associated with the planning process. Both the adaptive method and the use of dispersal-limited data identified reserve systems of intermediary size relative to reserve systems developed with only occurrence or only predicted distribution data.

Our novel methods could also be used when occurrence data are generally available but scarce for certain species. For example, in our case study, several species lacked a sufficient number of occurrence records to allow for the selection of 10 separate locations. Although reserve selection on the basis of only occurrence data would minimize uncertainty that the species were included in the reserve network, it would result in one or more species not being adequately protected at the desired level. To remedy this problem and minimize uncertainty, our adaptive method could be used because it allows for occurrence data to be incorporated yet the process is not solely dependent on this type of data. Another approach would be to reasonably expand data coverage of the occurrence data set by using our dispersal-limited technique. Nevertheless, if a situation arose in which occurrence data were available for every PU, there would be no need for predicted distribution data, dispersal-limited distribution data, or our adaptive method.

Predicted distribution data will remain important in areas where little or no occurrence data exist or if the PUs are very large and therefore the probability of occupation increases (Wilson et al. 2005; Rondinini et al. 2006; Shriner et al. 2006). Predicted distributions are also likely to become an increasingly important component of conservation planning as climate change and other dynamic processes are incorporated into planning efforts.

Uncertainty in Conservation Planning

Uncertainty in conservation planning can arise in a variety of ways. Our use of dispersal-limited distribution data and the novel adaptive application of systematic conservation planning illustrates one method to deal with the inherent uncertainty associated with data on species distribution and the process of conservation planning. The techniques we used are most useful if the available predicted distribution data overrepresents the true range of the species.

Other methods may also deal with problems of uncertainty in conservation planning. For example, information-gap decision theory seeks to identify strategies that will be the least likely to fail because of uncertainty. Although a useful tool, the methodology can be complex. To incorporate info-gap theory, one must develop a performance requirement, a mathematical-process model that measures performance as a result of management, and a model describing uncertainty in the conservation planning process (McDonald-Madden et al. 2008).

Additional methods to deal with uncertainty in species distribution data include collecting more-systematic presence and absence data, incorporating probability of occurrence data directly, and setting more-stringent thresholds when developing modeled distributions (see Wilson et al. 2005; Elith et al. 2006). Although these methods decrease the level of uncertainty associated with conservation planning, they often require additional time, money, or outside expertise that can make these alternatives prohibitive in terms of cost or feasibility. Our methods are intended to be an easy way to modify existing data so that uncertainty is reduced when the time, skills, or resources are not available to attempt a more complex solution. For example, the predicted distribution data we used were developed by the U.S. Geological Survey, and one of the desired outcomes was a data set that provided species distribution data for state and local conservation planning efforts (Boykin et al. 2007). A cursory analysis of these data showed that the models of predicted distribution overrepresented the expected distributions for many species in Arizona. With our methods we were able to reduce the errors of commission in our planning effort.

Although we did not perform quantitative tests that proved our methods reduced uncertainty, our underlying assumptions about the risk of omission versus commission errors associated with each type of distribution data are generally accepted in the literature (Rondinini et al. 2006). For example, the results of one recent study in which the accuracy of numerous species predictive modeling methods were analyzed, showed that even the most effective methods are generally only slightly better than chance at correctly predicting the presence of a species (Elith et al. 2006). This supports our efforts to further refine even the most accurate predicted distribution data, including data with a probability of occurrence, by restricting predicted distribution data to areas around known locations (dispersal-limited approach) or limiting its use in the planning process (adaptive method). Ultimately, it is rarely possible to quantify the exact uncertainty and error associated with distribution data (Wilson et al. 2005; Rondinini et al. 2006; Rodriguez et al. 2007). The true benefit of our novel approaches is, therefore, nearly impossible to quantify. Nevertheless, quantification of the benefits would be an important future research step for determining the advantage of these theoretical methods. We recognize that our analyses were conducted at a regional scale and that the advantages of our adaptive method might change with scale or increased availability of occurrence data.

Although data availability commonly drives the type of distribution data used in designing reserves, our results show that the choice of distribution data can greatly affect the final product. There was little spatial overlap between reserve networks created with the most commonly available types of species distribution data, and within a single type of distribution data, variance in the

reserve placement occurred. This lack of spatial congruence within and between different types of distribution data has real implications when arguing for the placement of reserves in a particular location. Although the spatial configuration of reserves is noteworthy, more important is that the targets for which the reserves were created actually reside within its boundaries. Our adaptive method and dispersal-limited approach address this problem and can minimize uncertainty that conservation targets are actually represented in the reserve network while still taking spatial constraints into account.

Acknowledgments

This study was supported by grants from the Arizona Game and Fish Department Heritage Fund, the National Science Foundation (award 0347960) and the Arizona Nongame Wildlife Checkoff. We gratefully acknowledge the contributions of J. Francis, Nongame Senior GIS Analyst, members of the Gerber lab, and PhD committee members for J. Underwood (D. Pearson, E. Wentz, A. Smith) for their insightful comments and guidance.

Literature Cited

- Ball, I. R., and H. P. Possingham. 2000. MARXAN (V1.8.2): Marine reserve design using spatially explicit annealing, a manual. University of Queensland, Brisbane, Australia.
- Bender, D. J., C. H. Flather, K. R. Wilson, and G. C. Reese. 2005. Regional data to support biodiversity assessments: terrestrial vertebrate and butterfly data from the Southwest. General technical report RMRS-GTR-152. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. Fort Collins, Colorado.
- Boykin, K. G., et al. 2007. Predicted animal habitat distributions and species richness. Pages 679–750 in J. S. Prior-Magee, K. G. Boykin, D. F. Bradford, W. G. Kepner, J. H. Lowry, D. L. Schrupp, K. A. Thomas, and B. C. Thompson, editors. Southwest regional gap analysis final report. U.S. Geological Survey, Gap Analysis Program, Moscow, Idaho.
- Ceballos, G., and P. R. Ehrlich. 2006. Global mammal distributions, biodiversity hotspots, and conservation. *Proceedings of the National Academy of Sciences of the United States of America* **103**:19374–19379.
- Elith, J., et al. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29**:129–51.
- Ferrier, S. 2002. Mapping spatial pattern in biodiversity for regional conservation planning: where to from here? *Systematic Biology* **51**:331–363.
- Flather, C. H., K. R. Wilson, D. J. Dean, and W. C. McComb. 1997. Identifying gaps in conservation networks: of indicators and uncertainty in geographic-based analyses. *Ecological Applications* **7**:531–542.
- Fleishman, E., R. Mac Nally, J. P. Fay, and D. D. Murphy. 2001. Modeling and predicting species occurrence using broad-scale environmental variables: an example with butterflies of the Great Basin. *Conservation Biology* **15**:1674–1685.
- Fleishman, E., R. F. Noss, and B. R. Noon. 2006. Utility and limitations of species richness metrics for conservation planning. *Ecological Indicators* **6**:543–553.
- Forbeseh, P. F., M. Eno-Nku, and T. C. H. Sunderland. 2007. Priority setting for conservation in south-west Cameroon based on large mammal surveys. *Oryx* **41**:255–262.
- Funk, V. A., and K. S. Richardson. 2002. Systematic data in biodiversity studies: use it or lose it. *Systematic Biology* **51**:303–316.
- Grand, J., M. P. Cummings, T. G. Rebelo, T. H. Ricketts, and M. C. Neel. 2007. Biased data reduce efficiency and effectiveness of conservation reserve networks. *Ecology Letters* **10**:364–374.
- Grenyer, R., et al. 2006. Global distribution and conservation of rare and threatened vertebrates. *Nature* **444**:93–96.
- Loiselle, B. A., C. A. Howell, C. H. Graham, J. M. Goerck, T. Brooks, K. G. Smith, and P. H. Williams. 2003. Avoiding pitfalls of using species distribution models in conservation planning. *Conservation Biology* **17**:1591–1600.
- Margules, C. R., and R. L. Pressey. 2000. Systematic conservation planning. *Nature* **405**:243–253.
- McDonald-Madden, E., P. W. J. Baxter, and H. P. Possingham. 2008. Making robust decisions for conservation with restricted money and knowledge. *Journal of Applied Ecology* **45**:1630–1638.
- NatureServe. 2008. NatureServe explorer: an online encyclopedia of life. Version 7.0. NatureServe, Arlington, Virginia. Available from <http://www.natureserve.org/explorer> (accessed June 2008).
- Orme, C. D. L., et al. 2005. Global hotspots of species richness are not congruent with endemism or threat. *Nature* **436**:1016–1019.
- Pearson, D. L. 1994. Selecting indicator taxa for the quantitative assessment of biodiversity. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **345**:75–79.
- Pressey, R. L., R. M. Cowling, and M. Rouget. 2003. Formulating conservation targets for biodiversity pattern and process in the Cape Floristic Region, South Africa. *Biological Conservation* **112**:99–127.
- Pressey, R. L., M. Cabeza, M. E. Watts, R. M. Cowling, and K. A. Wilson. 2007. Conservation planning in a changing world. *Trends in Ecology & Evolution* **22**:583–592.
- Reyers, B., A. S. van Jaarsveld, and M. Kruger. 2000. Complementarity as a biodiversity indicator strategy. *Proceedings of the Royal Society of London Series B-Biological Sciences* **267**:505–513.
- Rondinini, C., K. A. Wilson, L. Boitani, H. Grantham, and H. P. Possingham. 2006. Tradeoffs of different types of species occurrence data for use in systematic conservation planning. *Ecology Letters* **9**:1136–1145.
- Rodrigues, A. S. L., and K. J. Gaston. 2001. How large do reserve networks need to be? *Ecology Letters* **4**:602–609.
- Rodriguez, J. P., L. Brotons, J. Bustamante, and J. Seoane. 2007. The application of predictive modelling of species distribution to biodiversity conservation. *Diversity and Distributions* **13**:243–251.
- Shriner, S. A., K. R. Wilson, and C. H. Flather. 2006. Reserve networks based on richness hotspots and representation vary with scale. *Ecological Applications* **16**:1660–1673.
- Smith, A. T., et al. 2000. Databases tailored for biodiversity conservation. *Science* **290**:2073–2073.
- van Jaarsveld, A. S., et al. 1998. Biodiversity assessment and conservation strategies. *Science* **279**:2106–2108.
- Vaughan, I. P., and S. J. Ormerod. 2005. The continuing challenges of testing species distribution models. *Journal of Applied Ecology* **42**:720–730.
- Warman, L. D., A. R. E. Sinclair, G. G. E. Scudder, B. Klinkenberg, and R. L. Pressey. 2004. Sensitivity of systematic reserve selection to decisions about scale, biological data, and targets: case study from southern British Columbia. *Conservation Biology* **18**:655–666.
- Wilson, K. A., M. I. Westphal, H. P. Possingham, and J. Elith. 2005. Sensitivity of conservation planning to different approaches to using predicted species distribution data. *Biological Conservation* **122**:99–112.
- Yahnke, C. J., I. G. de Fox, and F. Colman. 1998. Mammalian species richness in Paraguay: the effectiveness of national parks in preserving biodiversity. *Biological Conservation* **84**:263–268.