ANIMAL MODELING

Novel Approaches to Mapping Vertebrate Occurrence for the Northwest Gap Analysis Project

Jocelyn Aycrigg[1] and Gary Beauvais[2]

Introduction

The basic goal of the National Gap Analysis Program is to "keep common species common." Modeling and mapping the occurrence of vertebrate species within a state or region are used to achieve this goal. During previous state GAP projects, standardized methods of modeling and mapping vertebrates were used (for example, Merrill et al. 1996). However, since those projects have been completed new modeling approaches have been developed, particularly inductive modeling approaches that integrate relatively well with advanced geographic information systems (GIS) (for example, Carpenter et al. 1993; Phillips et al. 2006; see Elith et al. 2006 for review).

The Northwest Gap Analysis Project (NWGAP), which includes Idaho, Montana, Oregon, Washington, and Wyoming, began in September 2004. Our objectives for this project include creating consistent and current data products that are repeatable and standardized, and of high utility to resource managers. To meet these objectives we are proposing some novel approaches to modeling and mapping vertebrate occurrence. Similar to other state and regional projects we will create a species list, collect species occurrence data, assemble habitat and environmental data, and conduct species modeling. Our novel approaches address who we are collaborating with to collect species occurrence data, what we are aiming to map, and how we are conducting the modeling. Throughout our work we intend to involve local resource managers and biologists as much as possible to increase the accuracy and utility of our data products.

Who Will be Collecting Species Occurrence Data and Conducting Species Modeling?

We are working with the Natural Heritage Programs (NHP) in each of the 5 northwestern states to compile and organize all species occurrence data, provide local expertise and input into species modeling, and identify additional local and regional experts to involve in the project. The Northwestern NHPs are Montana Natural Heritage Program, Idaho Conservation Data Center, Washington Natural Heritage Program, Oregon Natural Heritage Information Program, and Wyoming Natural Diversity Database. The Wyoming Natural Diversity Database is acting as the central species modeling organization for NWGAP and is coordinating the activities of the other programs as well as producing final models and maps.

There are several advantages to collaborating with NHPs. First, each program contains a vast amount of existing expertise, occurrence data, and database and personnel infrastructure focused on describing and documenting the flora and fauna of their state, with special emphasis on species of conservation concern. Each program employs a lead zoologist who is recognized as an expert on the state fauna and environment, and who routinely collects, reviews, and summarizes vertebrate occurrence data. In effect, each NHP acts as a central clearinghouse for vertebrate occurrence data and status information for a given state.

Second, NHPs are responsible for managing species occurrence data only within their own state (although they regularly contact programs in neighboring states for information on species status, life history, and management).

[1] National Gap Analysis Program, University of Idaho, Moscow, Idaho.

[2]Wyoming Natural Diversity Database, University of Wyoming, Laramie, Wyoming.

By compiling regional occurrences for given species from five state-specific datasets, we can minimize the risk of obtaining duplicate records.

Third, each NHP uses a common database software, standards, and methods. This helps ensure that species occurrence records from different states have undergone similar quality control procedures, have similar content and format, and are easily integrated into a region-wide dataset. Species occurrences compiled for NWGAP will include a unique record identification, primary data source (for example, NHP zoologist, state biologist, etc.), genus and species, year of observation, month of observation, latitude/longitude of observation, and accuracy/precision of observation.

Fourth, the species occurrence data in each NHP are continually reviewed and quality-checked. Reviewing and quality-checking are necessary because individual species occurrences often are documented in many sources and NHPs commonly receive multiple submissions of the same observation. Without strict quality control individual species records could appear several times in the occurrence database.

This process minimizes the chance of duplicate records, which could bias subsequent estimates of a species distribution and (or) habitat use. Collaborating with NHPs for NWGAP ensures we will obtain a high quality, comprehensive compilation of occurrence records that requires minimal additional processing and screening.

Fifth, in addition to reviewing and quality checking the data, each NHP continually updates and maintains its occurrence database. Efficient access to updated data should make NWGAP products more dynamic. This will make re-running models for species with new occurrence data more feasible. By engaging NHPs as primary data sources, and employing more transparent and automated modeling methods, NWGAP cooperators should be able to update models and maps over short time frames.

What Will Be Modeled?

We intend to map *range*, *distribution*, and *habitat quality* separately for each terrestrial vertebrate in the Northwest region (Figure 1). By providing this information we intend to make regional species models useful and applicable to local land and resource managers.

We consider a species' *range* as the total area occupied; similarly, the spatial limits within which a species can be found (Morrison and Hall 2002). Range maps, then, are typically coarse-grained depictions of a species' spatial patterning defined almost entirely by geographic space, with little consideration of underlying environmental features. We will produce range maps based on aggregations of all pre-defined map units known to be occupied by a given species. Map units for NWGAP range maps will be 10-digit Hydrologic Unit Codes (HUCs).

A species' *distribution* is defined as a finer-scale depiction of a species spatial patterning, which relies on identification of multivariate environments suitable for occupation by a given species; similarly, the "spread" or "scatter" of a species within its range (Morrison and Hall 2002). A species' distribution is a spatial subset of its range.

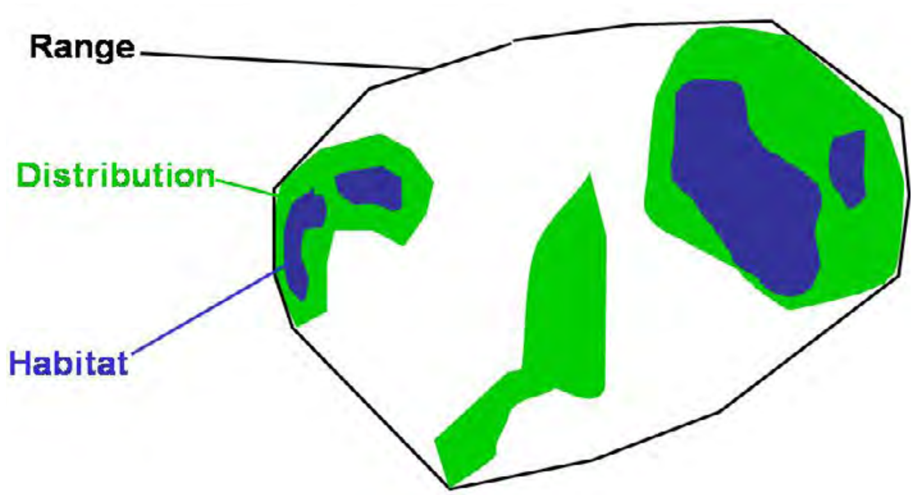

Figure 1. A hypothetical species' range, distribution, and habitat. Habitat is a spatial subset of distribution and distribution is a spatial subset of range.

In contrast to range maps, where map units are pre-determined blocks of geographic space that are occupied by individuals, distribution maps show the intersection of multiple environmental gradients that are potentially occupied by individuals (Beauvais and Master 2005). Distribution maps not only are finer in grain than range maps, but also depict more distinction between suitable and unsuitable environments.

A species' *habitat* is the combination of resources and conditions that promote occupancy, survival, and reproduction (Morrison and Hall 2002; Beauvais and Master 2005). Just as a species' distribution can be seen as a spatial subset of its range, a species' habitat is a spatial subset of its distribution. Habitat maps are inherently difficult to produce, especially for large areas, because the conditions that lead to positive rates of survival and reproduction vary at fine spatial and temporal scales, and often are not represented in large-extent digital maps. However, *habitat quality* maps, whereby different portions of a species' distribution are ordinally ranked (that is, high, medium, low) based on their predicted long term contribution to reproduction and survival can be produced by modifying distribution maps with additional information (Beauvais and Master 2005). We intend to produce habitat quality maps for terrestrial vertebrates in the Northwest by compiling expert opinion on environmental factors that correlate with rates of survival and reproduction, and modifying distribution maps accordingly.

How Will Modeling Be Done?

Previous state and regional Gap projects have used a deductive, or "expert systems", modeling approach in which information about habitat associations are synthesized from experts and literature reviews (Figure 2). Types of habitat association data include land cover, elevation, soil, proximity to water, and climatic gradients. A predicted species distribution based on these habitat associations is made, and then intersected with and sometimes modified by species occurrence data. Species experts then review the draft distribution map, propose edits, and the map is revised to a final stage.

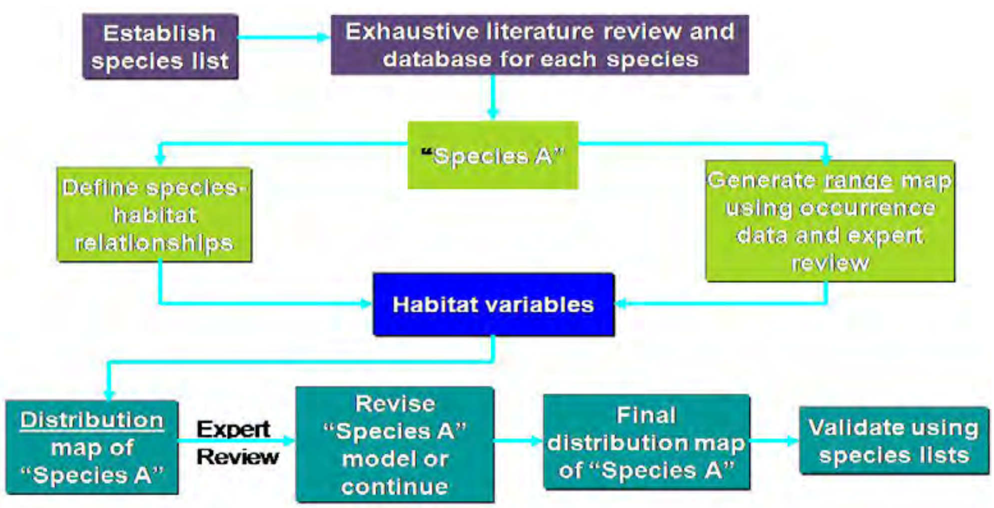

Figure 2. Steps involved with deductive modeling. These steps were taken to create species distribution maps for state-based GAP projects

Deductive models generally work well for well-known species that have clear, well-defined relationships with major environmental variables. They also are often the only option for modeling the distribution of very poorly-studied species for which there are few mapped occurrence records. However, deductive models typically over predict distributions when habitat associations or mapped land cover types are too general, when target species use environments that are difficult to define with satellite imagery, or when a species' distribution is strongly driven by interactions among environmental variables. If well mapped, land cover types are adequate surrogates for some complicated environmental patterns (Pressey 2004), but this is not true in all cases (Brooks et al. 2004).

In contrast to deductive modeling, inductive modeling is an approach whereby the multivariate environments at points of known occurrence are statistically summarized, and then extrapolated across the study area (Figure 3). Inductive modeling approaches (for example, Carpenter et al. 1993; Phillips et al. 2006; see Elith et al. 2006 for review) have gained in popularity for several reasons: resulting predictions are precise and repeatable; calculation methods are explicit and transparent; ease of use has increased as computer (particularly desktop GIS) technology has advanced; their primary inputs—georeferenced records of species occurrence and digital layers of environmental features—are increasingly available; and most inductive models can be easily re-run and improved with time and additional data (Elith et al. 2006).

Several algorithms have been developed to produce inductive models of species distributions, such as BioClim (Nix 1986), DOMAIN (Carpenter et al. 1993), GARP (Stockwell and Peters 1999), and MaxEnt (Phillips et al. 2004; Phillips et al. 2006). Recent reviews and comparisons of algorithms indicate that opportunistically-collected species occurrence data, if processed carefully, are appropriate for producing accurate distribution models, and that some algorithms (such as MaxEnt) produce generally more accurate and robust models than others (Elith et al. 2006; Hernandez et al. 2006; Phillips et al. 2006).

Most inductive modeling algorithms work best with abundant and well-distributed points of known species' presence, and suspected species' absence. The predictive power of an algorithm decreases as data quantity and quality decrease. Existing data may not meet the standards required by some algorithms, and additional sampling to boost data quantity and quality is often prohibitively expensive and time-consuming (Brooks et al. 2004).

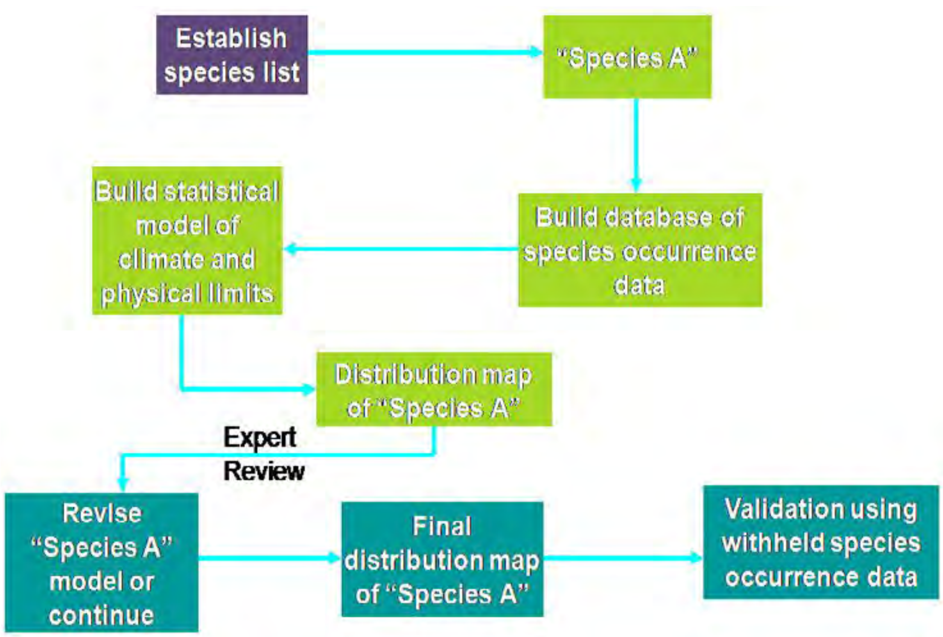

Figure 3. Steps involved with inductive modeling. These steps will be followed to create inductive species occurrence models for NWGAP.

Clearly, there are strengths and limitations to both deductive and inductive modeling approaches. For NWGAP we anticipate using purely deductive modeling for a few widespread species, species with coarse and well-established environmental relationships, and species whose available occurrence data are wholly inappropriate to inductive approaches. However, for most species we intend to combine the strengths of both modeling approaches to produce robust distribution models (Figure 4). This approach is designed to use the strengths of one approach to compensate for the weaknesses in the other (see discussion in Brooks et al. 2004).

Beauvais et al. (2003) describe the proposed combined modeling approach we intend to use for NWGAP (Figure 5). They collected species occurrence records for swift fox (*Vulpes velox*) in the five states of the U.S. Forest Service Rocky Mountain Region. These data were used as input to the DOMAIN (Carpenter et al. 1993) modeling algorithm to obtain a distribution map based on physical and climatic parameters (Figure 5). This map then was intersected with a deductive model of landcover types deemed suitable for occupation by swift fox. The intersection of the two maps encompassed 92 percent of the swift fox points in an independent validation data set, and was presented as the final predictive distribution map for the species in this region.

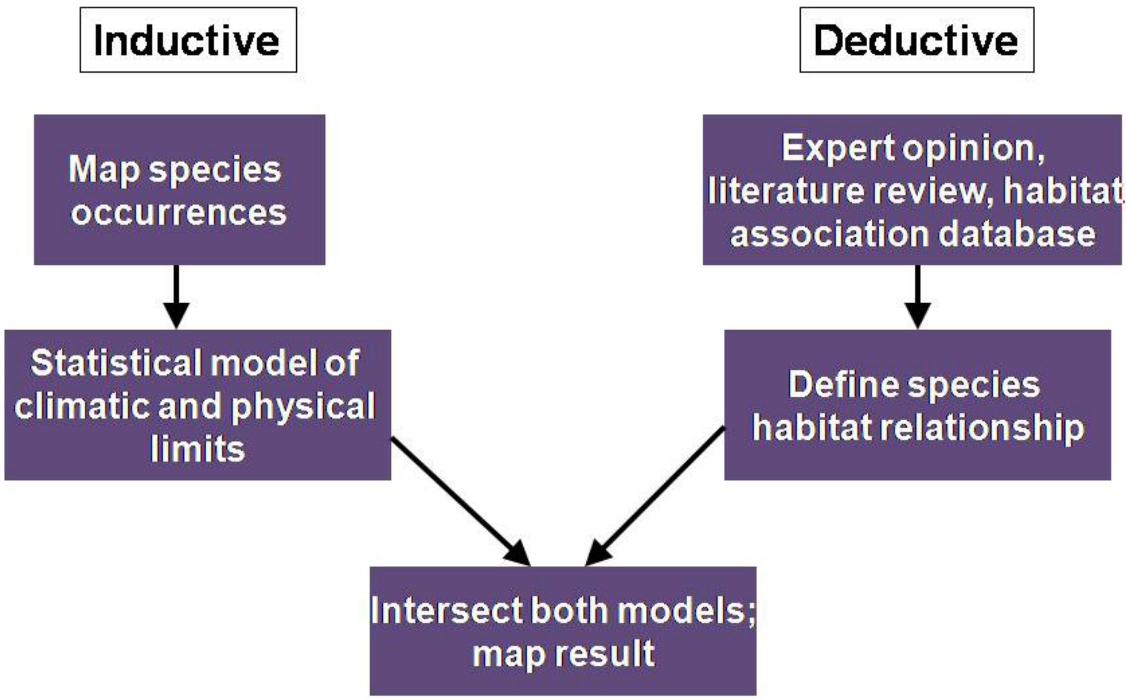

Figure 4. Our proposed method of combining the strengths of deductive and inductive modeling to construct better species occurrence models. This diagram describes the approach that will be used for creating species' distributions for NWGAP

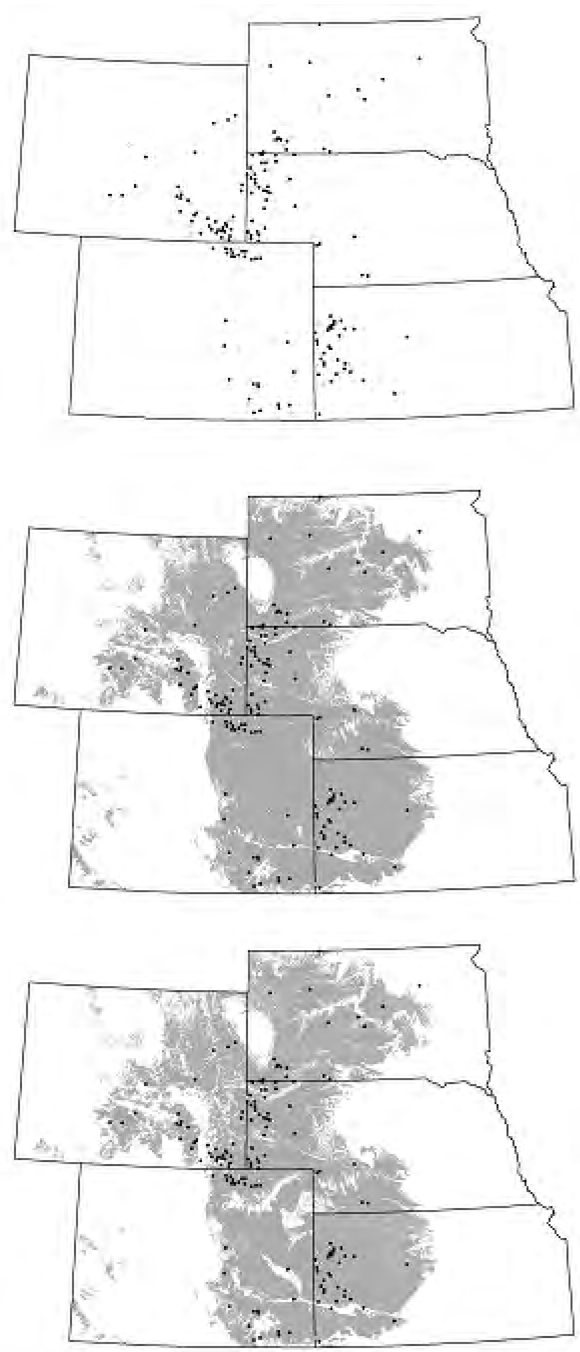

Figure 5. An example of combining inductive with deductive modeling using species occurrence and habitat association data for the swift fox (Vulpes velox; Beauvais et al. 2003). The top map shows the species occurrence data available in the five states of U.S. Forest Service (USFS) Region 2. Solid circles (75 percent of total) were used in the model and open circles (25 percent of total) were withheld and used for independent validation of the model. The middle map shows the results of the inductive model using the DOMAIN modeling algorithm with the gray areas indicating areas to be included in the model. The bottom map is the intersection of the two maps above resulting in a predicted distribution map for swift fox, which include 92 percent of the independent validation points (that is, open circles). Gray areas show (1) habitats associated with the species, as selected by Gap Analysis teams within each state, and (2) are within the suitable inductive model for the species. Note the areas that have been removed from the middle map to refine the predictive distribution in the bottom map.

Summary

Our species modeling effort for NWGAP is collaborating with NHPs in the five Northwestern states. This collaboration will effectively and efficiently collect and process species occurrence data and provide access to vital biological expertise. The result will be a high quality and comprehensive compilation of species occurrence data. Furthermore, we are striving to create multiple maps (range, distribution and habitat quality), which represent multiple scales, for each terrestrial species occurring in the Northwest. Our intent is to make regional models more useful to local natural resource managers. Finally, in an effort to improve the accuracy of the final map for each species, we are combining the strengths of inductive and deductive modeling.

We have finalized the species list and assembled the species occurrence data in the Northwest. In 2007, we began modeling and within 3 years will have final species models and maps for use in the Northwest gap analysis as well as other conservation efforts.

References Cited

Beauvais, G., and L. Master. 2005. Proposed procedure for modeling range, distribution, and habitat quality of vertebrates for Northwest GAP. Proposal submitted to National Gap Analysis Program.

Beauvais, G.P., R. Thurston, and D. Keinath. 2003. Predictive range maps for 15 species of management concern in the Rocky Mountain Region of the USDA Forest Service. Report prepared for the U.S. Geological Survey-National Gap Analysis Program by the Wyoming Natural Diversity Database-University of Wyoming, Laramie, Wyoming, USA.

Brooks, T., G.A.B. Da Fonseca, and A.S.L. Rodrigues. 2004. Species, data, and conservation planning. *Conservation Biology* 18:1682-1688.

Carpenter, G., A.N. Gillison, and J. Winter. 1993. DOMAIN: A flexible modeling procedure for mapping potential distributions of plants and animals. *Biodiversity and Conservation* 2:667-680.

Elith, J., C.H. Graham, R.P. Anderson, M. Dudík, S. Ferrier, A. Guisan, R.J. Hijmans, F. Huettmann, J.R. Leathwick, A. Lehmann, J. Li, L.G. Lohmann, B.A. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J.M. Overton, A.T. Peterson, S.J. Phillips, K. Richardson, R. Scachetti-Pereira, R.E. Schapire, J. Soberón, S. Williams, M.S. Wisz, and N. E. Zimmermann. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29:129-151.

Hernandez, P., C.H. Graham, L.L. Master, and D.L. Albert. 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography* 29:773-785.

Merrill, E.H., T.W. Kohley, M.E. Herdendorf, W.A. Reiners, K.L. Driese, R.W. Marrs, and S.H. Anderson. 1996. *The Wyoming Gap Analysis Project Final Report*. Wyoming Cooperative Fish and Wildlife Research Unit, University of Wyoming, Laramie.

Morrison, M.L., and L.S. Hall. 2002. Standard terminology: Toward a common language to advance ecological understanding and application. Pages 43-52 in *Predicting Species Occurrences: Issues of accuracy and scale*. Editors: J.M. Scott, P.J. Heglund, and M.L. Morrison, et al., Island Press.

Nix, H.A. 1986. A biogeographic analysis of Australian elapid snakes. Pages 4-15 in *Atlas of Australian elapid snakes*. Series 8 Editor: R. Longmore. Commonwealth Scientific, Industrial, and Research Organization (CSIRO) Publishing, Collingwood, Victoria, Australia.

Phillips, S.J., R.P. Anderson, and R.E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modeling* 190:231–259.

Phillips, S.J., J. Dudík, and R.E. Schapire. 2004. A maximum entropy approach to species distribution modeling. *Proceedings of the 21st International Conference on Machine Learning*, Banff, Canada, 2004.

Pressey, R.L. 2004. Conservation planning and biodiversity; assembling the best data for the job. *Conservation Biology* 18:1677-1681.

Stockwell, D.R.B. and D.P. Peters. 1999. The GARP modeling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science* 13:143-158.