Jeffrey Swanson  
NRS 509 Project

Statistical Applications of GIS

The use of a GIS to support and provide analysis for statisticians proved to be a very extensive topic, so this paper will be focused on the use of GIS in habitat selection estimation studies. As an aspiring environmental statistician the amount of easily obtainable geographically referenced data available through a GIS is an exciting proposition. As computer systems become more powerful and databases become larger and more complex the future looks very bright for these types of studies.

The majority of GIS usage in habitat selection and usage studies is for adding additional categorical variables to the model. In a study of gray partridges these included merely the habitat type in which they were found. (Alldredge et al. 1998) But in a study of winter habitat selection of pronghorn the GIS provided not only the habitat type where they were seen, but elevation data, how far the animal was from water, and their distance to the nearest road. (Alldredge et al. 1998) A 1998 study by W.P. Erickson et al. seemed to take advantage of many of GIS’ capabilities. GIS was used to find a suitable sampling area for moose habitat in burned and unburned areas near a major river in Alaska, then was used to map the sampling route taken in the airplane to create equal flight times over each habitat type, to provide categorical variables (mostly land cover type), and finally to map the probability of use function over the region. Finally, in a study of goshawk habitat selection in southern Alaska, GIS was used to plot positions obtained through radio telemetry to determine land cover and elevation of known hawk nesting locations. (Pendleton et al. 1998) While most of the categorical variable support could be done by brute force (getting out there to the sampling points and looking around and taking measurements), GIS has made this part of the exercise much easier and significantly cheaper (especially for areas that are remote and would be extremely expensive in time and money to visit). This allows for relocation of resources to other areas to create better studies. A study by Otis et al. (1998) brings to the table something that wouldn’t be feasible without a GIS. In many habitat selection studies it is impossible to replicate results to provide stronger statistical evidence because it is impossible to find a similar sampling unit to the one used in the study. Through the use of models in GIS and with various data layers it is now possible to survey large areas and find areas that are nearly identical to use as replication in a study. While each of these studies used a different statistical method of arrive at their conclusions each used GIS as a very integral part of the data analysis.

While the popularization and refinement of GIS has made possible many studies that would not have been feasible without it, it comes with some caveats that have been ignored until lately. A paper by Cressie et al (2003) brings to light the importance of quantifying the location error evident in all geographic data. Whether the error comes from collection the data (through improper surveying, or GPS error), error due to pixel size (and even error from the pixels being arranged in a regular grid), or overlaying data sets that have similar but not identical scales, it needs to be recognized. These errors are largely ignored in the estimation of error in a model or a surface simply
because they are not easily quantifiable and would significantly increase the time and knowledge needed to undertake such analyses.

The use of a modern GIS can greatly reduce the cost of many studies and is an excellent source of supporting data when dealing with geographically referenced data. As databases grow in size and computer power increases it will become extremely feasible to undertake extremely complicated statistical studies of habitat usage.

This paper is concerned with the value of knowing (and compensating for) location error in geographic data. Error in the provided location is often overlooked by geostatisticians. It provides examples of location error propagation in several types of situations. Including surface estimation from continuous data, error in predictions based on grid sampled data, error associated with raster based data, and error as a result of using scales of data that are comparable but not identical. Then a method for reducing and taking into account these various types of error is discussed, and an example using artificially created data is shown. The method involves a Monte-Carlo process to estimate the error surface. Finally, the method is used to refine a kriging estimation of the process.


This paper follows the history of wildlife population assessment methods from mark-recapture methods used in the early 20th century to present day methods which take advantage of a GIS. They also discuss the future. Currently a GIS makes setting up experiments easy (by easily integrating a topo map into a sampling plan) but also facilitates real-time satellite tracking. The integration of a GIS into population assessment and management operations will become even more important in the future with the availability of increasingly higher quality geographic data.


This paper gave a summary of a study on moose habitat selection in the Innoko NWR in rural Alaska. Moose locations were sampled and then buffer areas were created around each location to determine the land coverage using GIS. The land coverage data was obtained from Lansat TM data and fell into 22 categories in the sampling area. These data were used to estimate the probability that a moose would choose an area for habitat. These estimates were then mapped back onto the sampling areas based on the model that was created.


This paper addresses statistical methods for determining habitat selection in animal species. It discusses a few different types of studies which have recently been completed, including a study of moose in burned and unburned areas, a study of gray partridges, and a study of the winter habitat selection of pronghorn. The pronghorn study utilized GIS support to add elevational, slope, and diversity of vegetation to the location data received from satellite collars. These were then applied to determine
areas of feasible winter habitat using polytomous logistic regression. It then covers the methods applied to each situation and concludes that most resource selection methods require use of log-linear models, which are much more complex than the methods previously accepted.


This paper highlights the importance of GIS usage in planning and executing studies of habitat selection. In the past, it was unfeasible to replicate these studies because it is difficult to locate exact replications of habitat types so it was therefore impossible to make inferences based on any statistical proof. But with a GIS it becomes trivial through the use of land cover data and known habitat areas, to select similarly sized, shaped, and bordered areas to replicate the sampling. Traditional statistical methods can then be applied and inferences about habitat usage can be made. These methods are then discussed.


This paper covers the use of compositional analysis of habitat with GIS support to determine usage by Goshawks in Alaska. There is a need for conservation in these areas because of the timber harvest that is currently removing habitat. The hawks were tagged with radio locators and the locational data was collected by radio-telemetry. Habitat types were determined with GIS data based on land cover data and elevation data. Analysis was completed using a randomization technique (used because of the small sample sizes and non-normal distribution of hawks) to determine that there was no selection of habitat among the hawks, they merely used what was available to them.