

# Reserve Network Design in the Eastern Georgian Bay Region: Modelling the Eastern Massasauga Rattlesnake

Paul Zorn<sup>1</sup> and Justin Quirouette.<sup>2</sup>

<sup>1</sup>Parks Canada, Ontario Service Centre, Ecosystem Management Section, 111 Water Street East. Cornwall, Ontario K6H 6S3

<sup>2</sup>Georgian Bay Islands National Park, Box 28, Honey Harbour, Ontario P0E 1E0

## Abstract

*Georgian Bay Islands National Park is currently leading a multi-agency initiative to design a reserve network in the eastern Georgian Bay region. This initiative was derived from a multi-stakeholder workshop in 1994 that identified a vision for the area based on a network of core protected areas linked by conservation corridors (Geomatics Inc., 1994). Specifically, the purpose of this project is to design a system of cores and corridors that adequately represents the spatial requirements for species and is integrated with surrounding land use. The intent is to implement this reserve network in the regional municipality's official plan. Methods include a hierarchical approach that integrates geographical information systems (GIS), statistical modelling, conservation biology, behavioural ecology and landscape ecology into a comprehensive reserve network design that will be tested through subsequent monitoring and research plans. This paper focuses on the application of modelling methods to the eastern massasauga rattlesnake (*Sistrurus c. catenatus*) and compares the results of landscape-scale fragmentation analyses.*

## Introduction

Most published methods for designing reserve networks consisting of core areas and conservation corridors can be broadly categorized by two types: 1) landscape approaches based on concepts in landscape ecology and the analysis of remotely sensed data (e.g., Scott et al. 1993); and, 2) single-species approaches that focus on the habitat and dispersal requirements for specific rare species (e.g., Lamberson et al. 1994). Landscape approaches are generally intuitive, straightforward and have relatively smaller data demands than species approaches, however, their results are usually coarse and do not adequately represent the spatial requirements of species at a local scale. Single-species approaches take into account complex habitat associations and behavioural ecology. However, these approaches are too cumbersome to replicate for all species that would supposedly utilize a reserve network. In this study, two levels of analysis are employed beginning with coarse-level (landscape) methods and refined by fine-level (indicator species) methods. This hierarchical approach is designed to capitalize on the strengths, and minimize the constraints, of landscape and single-species studies. This multi-year project is currently in the species modelling phase where the habitat suitability of a selected suite of species is predicted.

## Methods

Spatial analyses were undertaken using Spans Explorer GIS 7.0 and Idrisi 2.0. Statistical modelling was developed using SPSS 7.5 and GS+. All spatial data

were carefully screened for data quality, consistency in scale and positional accuracy. Exported data for statistical analysis were screened for outliers, departures from normality and violation of statistical assumptions.

### ***Landscape Fragmentation Analysis***

Initial identification of potential core areas in the reserve network design were undertaken through habitat fragmentation analyses. Landsat TM imagery (1994, maximum likelihood classification) was reclassified to depict forest habitat classes. One hundred metre buffer areas were then calculated around road data from the National Topographic Data Base (scale = 1:50,000) and overlaid on top of the forest habitat layer and contiguous, unfragmented forest patches were identified (Figure 1).

### ***Logistic Regression Modelling***

Methods for the prediction of habitat utilization by indicator species throughout the study area were first tested on the *S. catenatus*. Logistic regression was the selected approach to develop the species probability models because dependent variables are dichotomous and logistic regression is tolerant of departures from normality among predictor variables (Norusis/SPSS Inc. 1997). Data used in the model to predict the habitat utilization of *S. catenatus* includes: observational data from the Element Occurrence database managed by the Natural Heritage Information Centre, Peterborough, Ontario; land cover (e.g., road density); habitat attributes (e.g., habitat patch type, size, perimeter); and, biophysical (e.g., elevation, aspect) data (Figure 2).

In all, 18 independent variables were used in conjunction with presence/absence data (dependent variable) for logistic regression modelling.

### ***Variables Sampling***

The effects of spatial autocorrelation can highly reduce the precision of regression models through the violation of assumptions regarding independence among variables (W. Li et al. 1997). Spatial autocorrelation for this analysis was directly measured using Moran's *I* coefficient. Patterns in Moran's *I* were compared at various lag distances for all independent variables. Following these comparisons, the optimal lag distance where Moran's *I* approached zero for all variables was identified. Based upon the selected lag distance variables were sampled using a random systematic sampling method and exported as a "point in polygon" table (N=856) for statistical analysis.

### ***Principle Components Analysis***

Multi-collinearity can also reduce the precision of regression models through violation of assumption of interdependence among independent variables. This study solved for multi-collinearity by transposing the raw data into orthogonal components using principle components analysis with a varimax rotation. Since varimax is an orthogonal rotation the principle components represents the original data in a form that is simplified (17 variables expressed by 12 components) and possesses no multi-collinearity. Twelve principal components were extracted based on the loadings from the rotated solution (Table 1). These 12 components were then saved as separate variables that were then used as independent variables in the logistic regression modelling process.

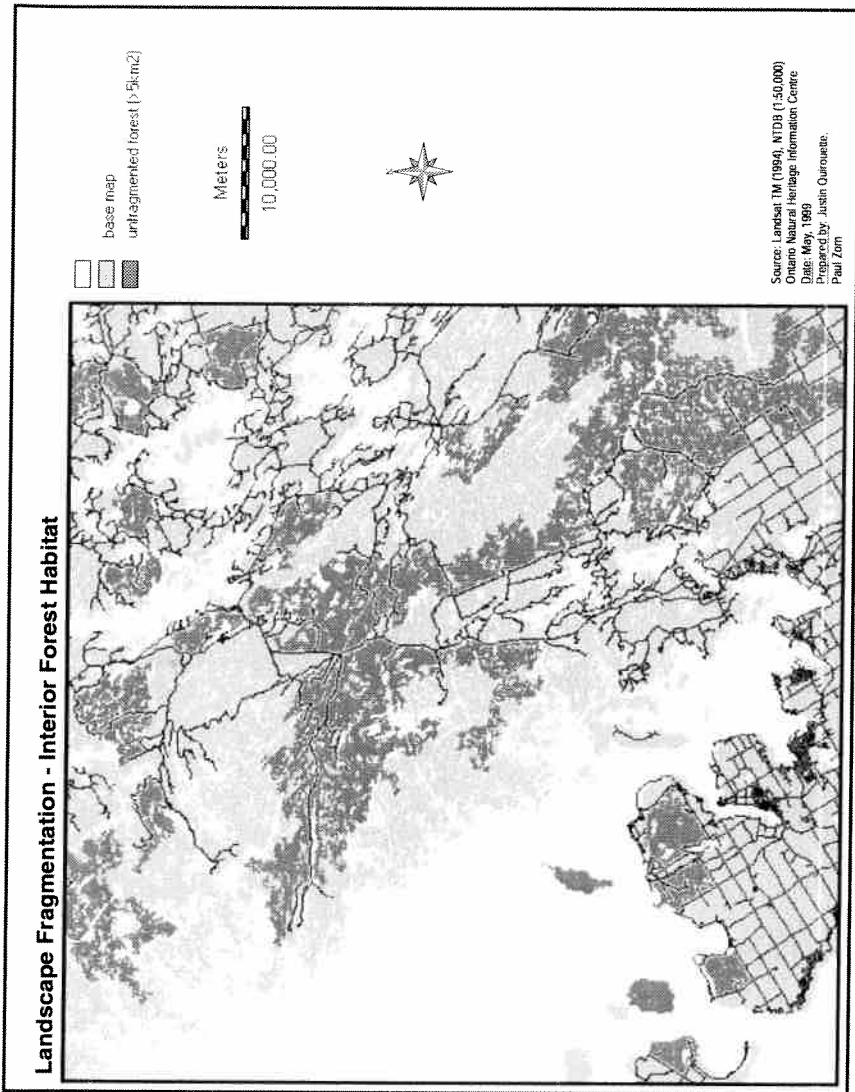


Figure 1: Interior forest habitat map generated through landscape fragmentation analysis using geographical information system.



Figure 2: Example of an independent variable map that was used as an input for logistic regression modelling.

Variable	Rotated Component Matrix											
	1	2	3	4	5	6	7	8	9	10	11	12
Aspect	-.062	.045	-.030	.008	-.027	-.073	.009	.977	-.014	-.028	.109	-.077
Dense Coniferous Area	-.014	-.014	.996	-.012	-.011	.002	.000	-.014	-.016	.010	.001	-.009
Dense Coniferous Perimeter	-.018	-.018	.996	-.014	-.014	-.003	.004	-.019	-.009	.018	.001	-.011
Dense Deciduous Area	-.031	-.030	-.013	-.023	.990	.005	-.006	-.012	.029	.043	.015	.025
Dense Deciduous Perimeter	-.030	-.030	-.013	-.023	.990	.001	.001	-.017	.027	.035	.009	.017
Elevation	.094	.009	-.020	-.018	.040	.143	-.053	-.083	.123	.097	.044	.965
Illumination	-.038	.018	.004	.004	-.004	.033	.984	.009	-.030	-.049	.038	-.048
Incidence Angle	.072	-.009	.001	.005	.024	.239	.044	.119	.021	.042	.953	.043
Mixed, Mainly Coniferous Area	-.026	-.028	-.011	.994	-.021	-.006	.002	.003	-.032	-.015	.003	-.002
Mixed, Mainly Coniferous Perimeter	-.031	-.034	-.014	.993	-.025	.000	.002	.007	-.036	-.018	.002	-.017
Mixed, Mainly Deciduous Area	.993	-.036	-.016	-.029	-.030	.044	-.022	-.037	-.007	.041	.037	.047
Mixed, Mainly Deciduous Perimeter	.993	-.038	-.016	-.030	-.032	.041	-.023	-.033	-.006	.044	.038	.051
Road Density	-.015	-.005	-.024	-.064	.051	.063	-.032	-.015	.980	.109	.021	.117
Slope	.035	-.021	-.013	-.003	-.009	.929	.207	.080	.023	.000	.063	.051
Sparse Mixed Forest Area	-.036	.997	-.016	-.030	-.030	-.019	.010	.022	-.001	.010	-.004	.008
Sparse Mixed Forest Perimeter	-.037	.997	-.017	-.032	-.031	-.020	.011	.027	-.004	.005	-.005	.000
Water Buffer	.078	.015	.028	-.030	.073	.055	-.052	-.030	.110	.978	.041	.094

Table 1. Rotated component matrix of independent variables for logistic regression model. Extraction method = principle component analysis. Rotation method = varimax.

### Logistic Regression

Twenty percent of the dataset from the "point in polygon" table was randomly selected and set aside for post-model verification. The remaining 80% were used as inputs into a logistic regression model predicting habitat utilization for *S. catenatus* using a forward stepwise enter method. Post model diagnostics were used to determine the statistical validity of the output model. These diagnostics included the Hosmer and Lemeshow Goodness-of-Fit Test and tests for normality among standardized residuals. Once confident of the validity of the model,  $\beta$  values for each variable entered into the model were used to create a probability surface in GIS using map algebra functions based on the following logistic regression equations:

$$\text{Equation 1: } Z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_i X_i$$

$$\text{Equation 2: } \text{Prob} = 1/(1+e^{-Z})$$

### Probability Surface Cutoff Values

The logistic regression model created a smooth probability surface that predicts the habitat utilization of *S. catenatus* (Figure 3). This surface was reclassified to categorize areas of high probability in order to answer the question: at what probability value will an area be categorized as being suitable for *S. catenatus* (e.g., 60%, 75%, 90%)? This probability cutoff value was determined by overlaying the 20% of the point data not used in the development of the model on to the probability surface. Using GIS, the ratio of presence points to absence points lying within and out of the habitat suitability polygons created by varying cutoff values was calculated. The optimal cutoff value was determined by the probability value that resulted in the highest presence to absence point ratio residing within the habitat suitability polygons.

## Results

The logistic regression model for *S. catenatus* was significant ( $X^2 = 21.3115$ ,  $P = .0064$ ) with a correct classification between presence and absence performance of 76.13%. The standardized residuals did not significantly differ from normal. The optimal cutoff value for the resultant probability surface with the highest presence to absence point ratio was .725 (Figure 4).

## Discussion

Areas with a .725 or greater probability of *S. catenatus* habitat utilization (Figure 4) are considered as potential core areas for integration into the reserve network design within the Georgian Bay Islands National Park ecosystem. Similar models will be conducted for a variety of species that represent various taxonomic groups, habitat types, body size, and dispersal behaviour. Each of these species probability models are to be field tested through incorporation into the park's research and monitoring plans.

Figure 5 displays the degree to which results from forest fragmentation analyses overlap with predicted habitat for *S. catenatus*. Since *S. catenatus* is not an interior forest species the lack of significant overlap is not surprising, however, the dissimilarity does support the need for an integration of landscape and species scale analyses when conducting reserve network design initiatives. Without the explicit

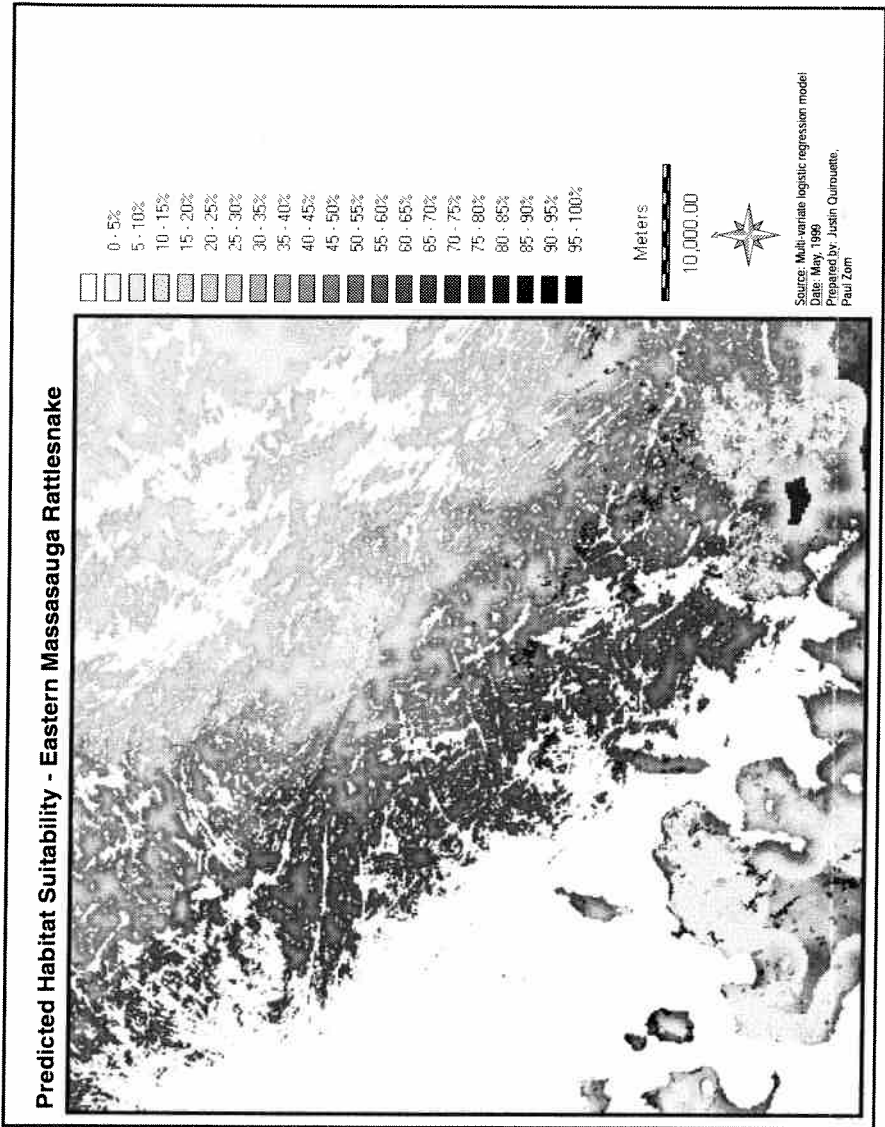


Figure 3: Probability surface predicting habitat suitability for *S. catenas* through multivariate logistic regression modelling.

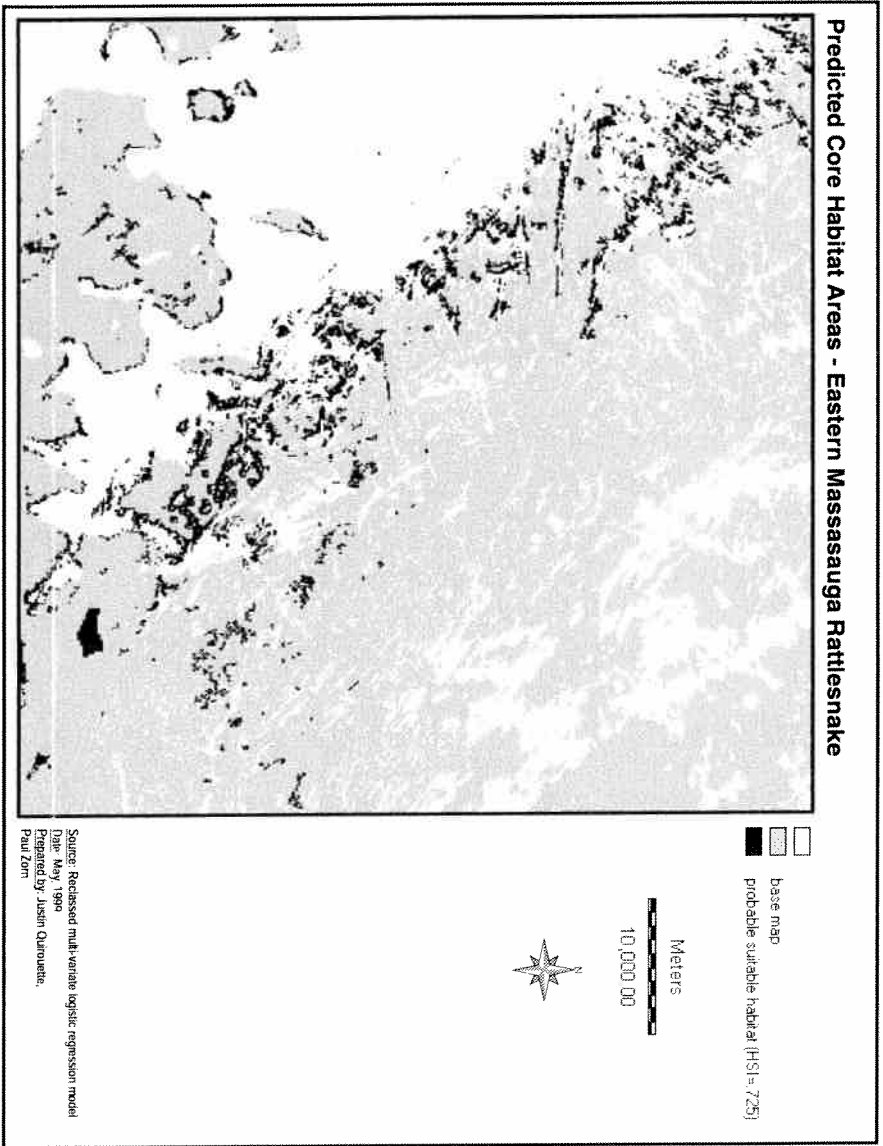


Figure 4: Predicted core habitat area for *S. catenatus* determined by reclassing probability surface according to a .725 cutoff value.



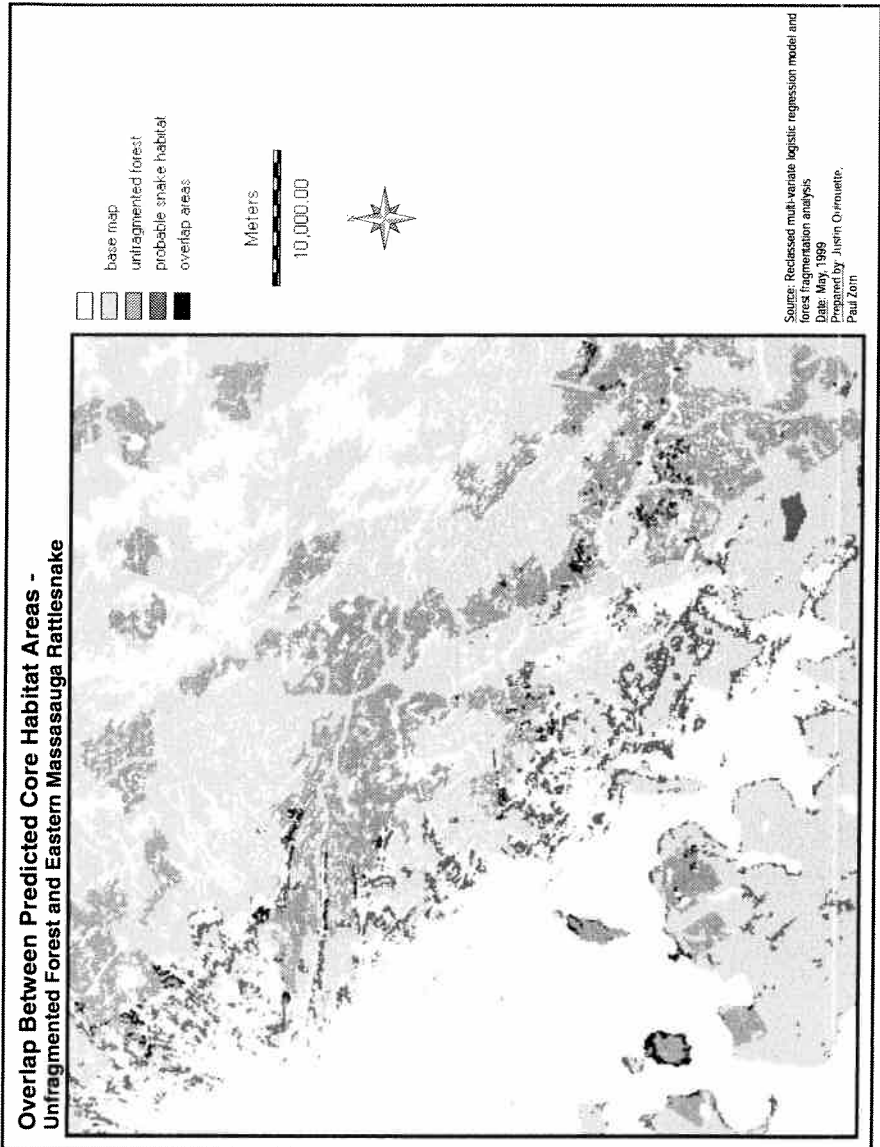


Figure 5: Spatial difference between core habitat areas identified through forest fragmentation and species modelling for the eastern massasauga rattlesnake. The degree of overlap suggests the need for analysis at both landscape and species scales.

incorporation of species habitat models, landscape fragmentation approaches may be insufficient to adequately identify habitat requirements for certain species.

Georgian Bay Islands National Park is continuing with multiple species logistic regression modelling and will be undertaking ground truthing beginning the 2000 field season. This work will form the basis for observational and telemetry studies where behavioural and demographic data will be collected and incorporated into the reserve network design process (e.g., dispersal patterns for determining the effectiveness of corridor design). Through iterative research and monitoring the reserve network initiative will not be a discreet project but the basis of the Georgian Bay Islands National Park's long term ecosystem conservation program.

For more information on this, or any, component of Georgian Bay Islands National Park's ecosystem management program please visit: [www.terra-plex.com/GBINP](http://www.terra-plex.com/GBINP).

## References

- Geomatics Inc. 1994. *Georgian Bay Greater Ecosystem Bioregional Study*. Campbellville, ON.
- Lamberson, R.H., B. R. Noon, C. Voss and K.S. McKelvey. 1994. Reserve design for terrestrial species: the effects of patch size and spacing on the viability of the Northern Spotted Owl. *Conservation Biology*. 8(1): 185-195.
- Li, W. et al. 1997. A regression model for the spatial distribution of red-crown crane in Yancheng Biosphere Reserve, China. *Ecological Modelling*. 103: 115-121.
- Norusis/SPSS Inc. 1997. *SPSS Professional Statistics 7.5*. Chicago, IL.
- Scott, J.M. et al. 1993. Gap analysis: a geographic approach to protection of biological diversity. *Wildlife Monographs*. 123.