

ISSUES AND CHALLENGES OF INCORPORATING FUZZY SETS IN ECOLOGICAL MODELING

VINCENT B. ROBINSON

*University of Toronto, 3359 Mississauga Road North,
Mississauga, ON L5L 1C6 Canada*

Abstract. An information-based framework is presented for spatially explicit GIS-based ecological modeling. Within this framework some of the important issues and challenges of incorporating fuzzy sets in spatially explicit population models (SEPM) are discussed. Examples of current work are used to illustrate the main issues and challenges facing the incorporation of fuzzy sets in ecological modeling. Among the challenges to be discussed are fuzzy-based techniques for data acquisition, model control/evaluation, heterogeneous representations of spatial data, parameterization of models, and hypothesis testing. There is special attention given to the issue of habitat modeling and presence/absence problem. Many scientific issues facing the incorporation of ecological models will be raised such as hypothesis testing, and relationship between statistical analysis and fuzzy techniques. Software availability is discussed as challenge for past and future. use of fuzzy techniques directed at handling uncertainty in GIS-based SEPM.

Keywords: fuzzy sets, spatially explicit population model, geographic information system, habitat model, ecological geographic variables

1. Introduction

Changes in the landscape of either natural or human cause can significantly influence the distribution of species over space and through time. Spatially explicit population models (SEPMs) have developed to the level of simulating how animals may disperse across a landscape, hence affecting the metapopulation dynamics of a species (Lurz et al., 2001). These models explicitly link landscape information with species specific ecological knowledge. The spatially explicit nature of SEPMs makes them especially attractive for use in developing policies, strategies, and tactics for a wide range of environmental protection and security issues. Indeed, they are

valued in environmental protection/conservation efforts precisely because of their utility in contrasting future management scenarios (Liu et al., 1995).

It is the inherent fuzziness of biological organization that has prompted some to apply fuzzy techniques in ecological studies (Schaefer and Wilson, 2002). It is also accepted now that fuzzy techniques can be used to model, or manage, uncertainties inherent in GIS databases and processing (Robinson, 2003). Precisely because of their spatially explicit nature, these models have been increasingly used in conjunction with geographic information systems (GIS). The modeling of bird or mammal metapopulation dynamics generally requires data representing the landscape and the population dynamics of particular species of interest. Thus, this chapter first presents an information-based framework for spatially explicit GIS-based ecological modeling. Within this framework an overview of how and where fuzzy sets may be used to discuss strategies for incorporating fuzzy information processing in SEPM efforts. Among the challenges discussed are fuzzy-based techniques for data acquisition, model control, heterogeneous representations of spatial data, parameterization of models, and hypothesis testing. The relationship between fuzzy techniques and statistical methods in ecological modeling is briefly addressed. The challenge of adopting fuzzy-based techniques in the ecological modeling field(s) will also be discussed. Many scientific issues facing the incorporation of ecological models will be raised along with those of cost and software availability.

In its presentation of the issues and challenges of incorporating fuzzy techniques in SEPM, this chapter will tend to focus upon bird and small mammal species. Birds have been used for millennia as indicators of environmental health or degradation. In addition, they may carry pathogens such as avian flu. Small mammals are among the most easily studied mammals, especially in relation to landscape attributes. Like birds, they also can carry pathogens such as rabies and plague. Thus, these two broad categories of biological entities are of particular relevance to the issue of environmental protection and security due to their responses to landscape level changes and their importance as potential vectors of human pathogens.

2. Information Theoretic Framework

Lima and Zollner (1996) discuss information-based approaches to modeling the movement and dispersal of animals. These approaches can be considered to represent a continuum that at one extreme is characterized by theoretical studies assuming animals disperse in random directions for random distances settling in the nearest detectable habitat patch(s). Such studies rarely use actual landscapes, hence are of limited value for management, or protection purposes. At the other extreme are approaches

that attribute considerable cognitive abilities to animals that include the use of spatial memory and learning. These models often make use of techniques from the field of artificial intelligence to model. These also tend to be poorly matched with real world landscape data and often make use of theoretical landscapes. In between these two extremes are those approaches that are intermediate in spatial scale where animals are typically given knowledge only about their nearby landscape and have no information about greater landscape. The model is often based on animals moving in the direction of the greatest detectable resource abundance or dispersing in direction of best detectable living site.

These approaches can be loosely categorized as being population- or individual-based models. As depicted in Figure 1, population-based models typically make use of a maximum dispersal distance to determine which habitat patches are potential destinations for individuals that need to be relocated from a patch with a surplus population to a patch of habitat that can accommodate a greater population density. Hence, the relationship is actually between the spatial entities of patches. The allocations are usually performed using random draws. Often there is little regard for obstacles to movement found in an intervening landscape matrix. In contrast, individual-based models base the relocation of individuals on distribution of habitat,

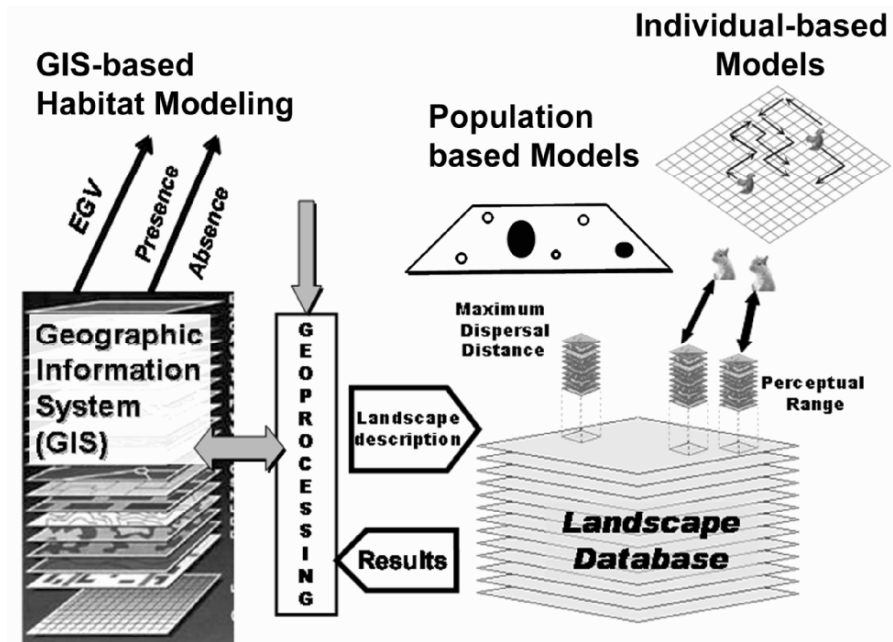


Figure 1. Conceptual overview of the relationship between GIS-based landscape modeling, population-based models of movement, and individual-based models of movement

and other landscape features, that are within the perceptual range of an individual (or social group). Hence the spatial relationship is more directly represented as being between an animal and landscape features.

Figure 1 also depicts in a conceptual manner some of the commonalities found in most all information-based modeling approaches. Information-based approaches use some version of perceptual range. In the case of population level models it is equivalent to the use of maximum dispersal distance. However complex or simplistic, they use rules of movement that control the movement of animal entities. They rely on a standard of plausibility (Lima and Zollner, 1996). The major reason for relying on a standard of plausibility is both the lack of spatially explicit behavior information on species' movements and difficulty of gathering such information (Zollner and Lima, 2005).

The process of going from ecogeographic variables (EGV) to arrive at a spatially explicit habitat model remains an important phase in a GIS-based approach. In addition, the behavioral and population model parameters contribute to the interpretation and execution of the rules of movement in concert with landscape data. At each step in this process there are questions of uncertainty to be addressed. Robinson (2002) discusses in detail the role fuzzy information processing may play at each step in this process. He notes the inherent fuzziness of the geographic information that forms the basis for the information used by the ecological model. In addition, he shows how fuzziness is inherent in the basic parameters of the models of dispersal/movement.

3. Incorporating Fuzzy Sets in GIS-Based Landscape Representations

In the context of the framework summarized in Figure 1, there are two major topics to consider when preparing to use a GIS database to support a SEPM exercise. First, there is the representation of the landscape without specific reference to habitat. Second, of crucial importance is the modeling of habitat.

3.1. LANDSCAPE INFORMATION

There are a number of sources of fuzziness inherent in the construction of a GIS database that may, or may not, subsequently be used for SEPM. Many GIS databases that are subsequently used for landscape level SEPM purposes were not originally assembled for that specific purpose. Figure 2 summarizes some the main input streams to a GIS database by the type of inputs. One of the major input streams is from sensor data. Land cover layers are typically used to reclassify the landscape into habitat versus

nonhabitat patches; and remote sensing inputs are a common input upon which those classifications are made. The fuzzy classification of and representation of remote sensing data is a well-studied problem, yet its potential use and/or effect on subsequent use in landscape level studies is less well studied (Arnot et al., 2004). Techniques such as fuzzy c-means (or k-means) clustering or related techniques can be used to produce a classification of land cover that preserves the inherent fuzziness that characterizes such representations (Comber et al., 2005; Robinson 1988).

Another major input stream is from expert opinion (Figure 2). In the context of these types of ecological models, a good example of such data would be bird count data collected by experienced bird surveyors. As noted later this is inherently fuzzy. Expert opinion on land cover and other environmental variables may also be characterized as linguistic variables (Bordogna et al., 2006; Zhu et al., 2001). Sketch maps of some form are not uncommon. Even in the case of bird counts there is a kind of sketch map technique whereby the observer uses a bull's-eye form to approximate the location of a bird in terms of distance and direction. Other sketch maps depicting routes have been shown to be amenable to analysis using F-histograms and linguistic descriptions (Skubic et al., 2004).

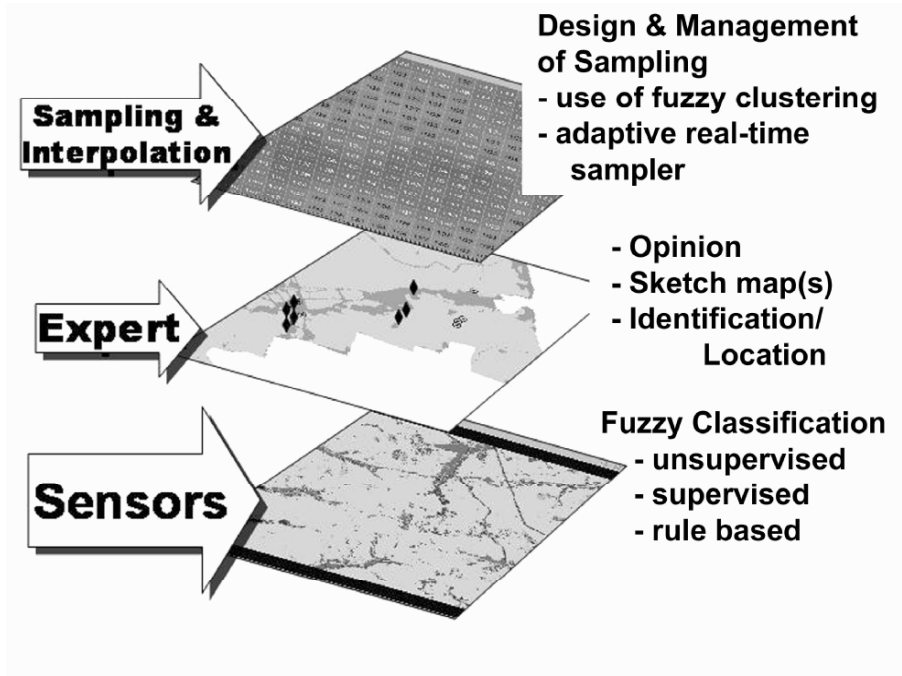


Figure 2. Major input streams used to create GIS database for landscape level studies and some associated fuzzy techniques

The third major input stream is associated many conventional methodologies used to create GIS-database layers from measurements on the ground based on some spatially explicit sampling scheme then interpolated to fill in the space to create a complete layer. One of the more common layers created in this fashion is the digital elevation model (DEM). Others include phenomena like soils, vegetation, wetlands, etc. In this context fuzzy clustering has been used to construct a sampling strategy from remotely sensed data or other reconnaissance information then interpolation methods used to generate a nonfuzzy representation such as soil characteristics (Odeh et al., 1990). It is less common for explicitly fuzzy techniques to be used for spatial interpolation (Dragicevic, 2005), especially in connection with efforts focused on SEPM. In contrast to this use of fuzzy sets, Graniero and Robinson (2003) have shown how a real-time adaptive sampler using fuzzy logic can increase the accuracy of such data collection while simultaneously lowering the cost.

3.2. HABITAT MODELING

Commonly, GIS-based models of animal movement require spatially explicit representation of the distribution of habitat as well as nonhabitat landscape elements. Uncertainty in the construction of a habitat model can have many different sources. It is also clear that in many cases defining a crisp boundary between habitat and nonhabitat is not ecologically meaningful, nor realistic. For example, Dettmers and Bart (1999) constructed a Boolean logic rule-based model of worm-eating warblers in included the condition that location have a slope between 35–44. However, does this mean that a location with a slope of 44.5 or 35.1 should be considered not habitat? Typically theories of the niche provide the theoretical underpinning to support the definition of suitability of a location as habitat for a particular species. The niche concept is usually described with reference to similarity of an optimal, or best, state. Thus, most habitat models rely on this logic. This has led some to suggest that whenever similarity-based concepts are used to construct a habitat model, that the model construction and testing should be based on using fuzzy logic (Hill and Binford, 2002).

3.2.1. *Fuzziness of Habitat Suitability Index (HSI)*

One common methodology is based on the Habitat Suitability Index (HSI). Burgman et al. (2001) point out several possible sources of uncertainty in the construction of a typical HSI. They emphasize the reliance, due to limited data, on expert opinion. They proceed to demonstrate how to use fuzzy numbers to encapsulate the uncertainty of the relationship between the value of an EGV (e.g., distance to forest) and level of suitability (i.e.,

ranges from 0 to 1). Reliability bounds for the best estimates of suitability were shown to vary significantly as a function of the characteristics of individual patches of landscape. Their example using the Florida scrub-jay illustrates how dependent some avian habitat modeling exercises are on expert opinion.

Hill and Binford (2002) also discuss the inherent fuzziness of the HSI. They make special reference to the reliance on expert opinion. Although they do not develop the concept of linguistic variables (Zadeh, 1978), they do demonstrate how a HSI model can be constructed using ambiguous categories using the theory of fuzzy sets. However, their demonstration of its application is not a GIS-based modeling approach. This lack of a GIS-based methodology is typical of HSI development. Hence, HSI results can often be of limited value to landscape level modeling efforts.

3.2.2. *The presence/absence problem*

Although expert opinion is often the basis for HSI, there is considerable effort directed at obtaining presence (and/or absence) data. Avian studies often rely upon detection of individuals through sight or sound (Austen et al., 2001; Jenkins et al., 2003). Mammalian studies often use trapping, photography, or tracks to determine presence (Carroll et al., 1999; Oehler and Litvaitis, 1996; Orrock et al., 2000). In both cases it is often tacitly assumed, for modeling purposes, that lack of detection is tantamount to absence. Hence the prevalence of methods assuming presence/absence data (Fielding and Bell, 1997).

In their study of why sparrow distributions do not match the spatially explicit predictions of habitat models, Jenkins et al. (2003) noted there were Commission errors where some habitat does not contain birds. These are areas where prior events have depleted species numbers, but because the birds do not disperse long distances, they do not reoccupy it quickly. This is consistent with theories of metapopulation dynamics.

Omission errors where some birds remain in unsuitable locations (i.e., nonhabitat or habitat of low quality) that were formerly suitable locations. This situation may result from the species having high site fidelity (Jenkins et al., 2003).

This has led to competing, nonfuzzy statistical methodologies for predicting the suitability of landscape elements as habitat (Brotons et al., 2004). Some of these methods include the use of pseudoabsence approaches, researchers are cautioned to be mindful of their study design as well as biases inherent in the presence data (Pearce and Boyce, 2006). Although this group of techniques makes use of expert opinion mostly as a general guide as to what EGV to include, recent results suggest that any

more detailed inclusion of expert opinion may actually decrease the predictive power of the model (Pearce et al., 2001; Seoane et al., 2005). Such results call into question the uncritical inclusion of expert opinion.

In summary, an estimate of only the amount of habitat or only the bird population may present an overly optimistic view of a species plight. Hence, the fuzziness inherent in determination of what constitutes habitat through presence/nonpresence data can have significant implication for environmental protection and conservation. Note that habitat in this context is represented as a crisp, nonfuzzy concept when, in fact, it more accurately is represented as a fuzzy concept.

3.2.3. Fuzzy habitat modeling with presence data

Like many fields, fuzzy c-means (or, k-means) was one of the first methods used in ecological analysis. In this case it is used commonly to develop classifications of ecological communities (Equihua, 1990). It has since been used to characterize the gradations in landscape-related characteristics within areas where species such as bats (Medellin et al., 2000) and squirrels (Bridges, 2003) are present. In these cases, it has been shown to represent more faithfully the internal heterogeneity of the habitat, thus avoiding the

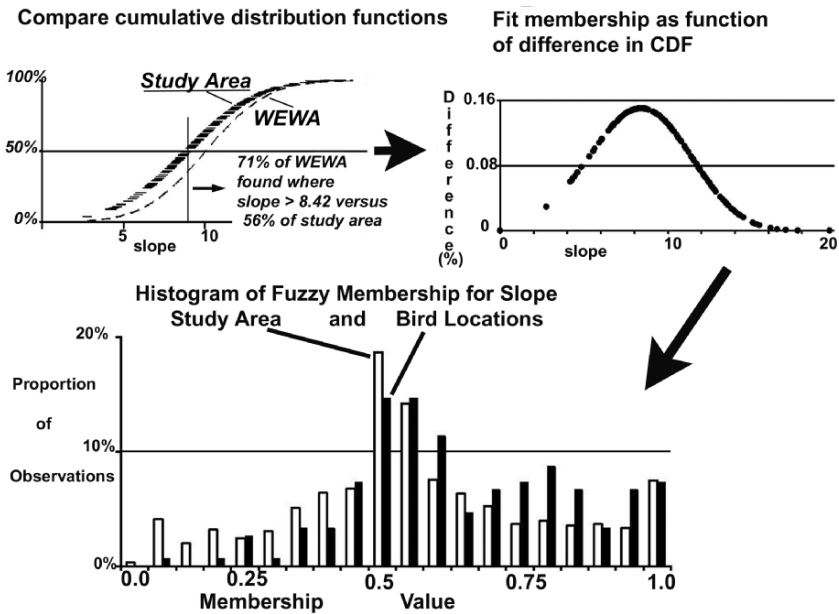


Figure 3. Example of an experiment inferring habitat rules by fuzzification of CDF-based methodology using presence data

oversimplification that occurs with crisp classifications. However, this approach has not been used to actually infer habitat from presence-only data.

One approach to inferring a fuzzy habitat model from presence-only data is currently under exploration (Hodgson, in progress). Figure 3 summarizes a portion of the process whereby cumulative distribution functions (CDFs) between presence data and the surrounding study area are compared to determine if there is any difference in the two. If there is a difference it is used in conjunction with membership function generation to determine if the variable can be used as an important EGV describing habitat for the species. In concept this is a fuzzification of the nonfuzzy approach described by Dettmers and Bart (1999). The end result would be a GIS raster layer in which each cell is coded according to the degree to which it is a member of the habitat set.

4. Fuzziness and Model Parameters

Attention has largely focused on the uncertainty concerning landscape level GIS information as input to support SEPM while few studies have investigated the effects of uncertain model parameters. One such study, using a crisp classification of habitat/nonhabitat landscapes found that errors in dispersal-mortality parameters resulted in greatest prediction errors, followed by mobility errors, and lastly landscape errors (Ruckelshaus et al., 1997). Results such as those in Ruckelshaus et al. (1997) suggest that focusing solely on the fuzziness of the landscape representations may not be as important to the reliability of model predictions as focusing on addressing the uncertainty in model parameters. Their results suggest that uncertainty surrounding dispersal parameters is a significant problem that should receive attention from those constructing SEPMs. Since crucial parameters in the dispersal/movement models concern decisions of where to move to and when to stop moving, it can be considered analogous to a control problem. Indeed Robinson (2002) has drawn parallels between control of mobile agents, robotic, and otherwise, to the problem of modeling individual animals in support of SEPM. Similarly, fuzzy cellular automata have been used to model insect infestation spread in a forested region of British Columbia (Bone et al., 2006).

In one GIS-based modeling effort, a fuzzy membership function was used to model the likelihood that a grid cell can be reached by a squirrel starting from a source patch. Results using the fuzzy approach were more consistent with field data (DeGenst et al., 2001). This suggests the potential utility of fuzzy sets in the parameterization of spatially explicit models at the population level model (see Figure 1).

Another approach is the modeling of animals at the individual level where vital parameters controlling dispersal/movement, such as perceptual range, that cannot be precisely derived from field and/or experimental work. For example, Mech and Zollner (2002) report the perceptual range of several common sciuridae as ranges, such as for the gray squirrel the perceptual range is between 300 m and 400 m. However it is not biologically meaningful to say 400.1 m is completely and utterly not perceived while a landscape feature at 400 m is perceived. Typically perceptual ranges are modeled as if they occur on isotropic surface rather than the anisotropic surfaces that exist in reality. It has been shown that such context dependent perceptual ranges can have an effect on the movement behavior of individuals (Olden et al., 2004).

Robinson and Graniero (2005) describe how an object-oriented probe mechanism can be used to incorporate fuzziness in an effective manner at the level of model parameters such as the perceptual range. Essentially the individual animals are simulated as a population of objects where each class has a different decision-making model of whether to move, where to move, and when to stop moving. However, each object does not directly draw its information from the GIS-based database, but rather through a probe mechanism (Graniero and Robinson, in press). Robinson and Graniero (2005) model squirrel dispersal using a fuzzy set representation of perceptual range. It is modeled as a function of distance on an isotropic surface. In

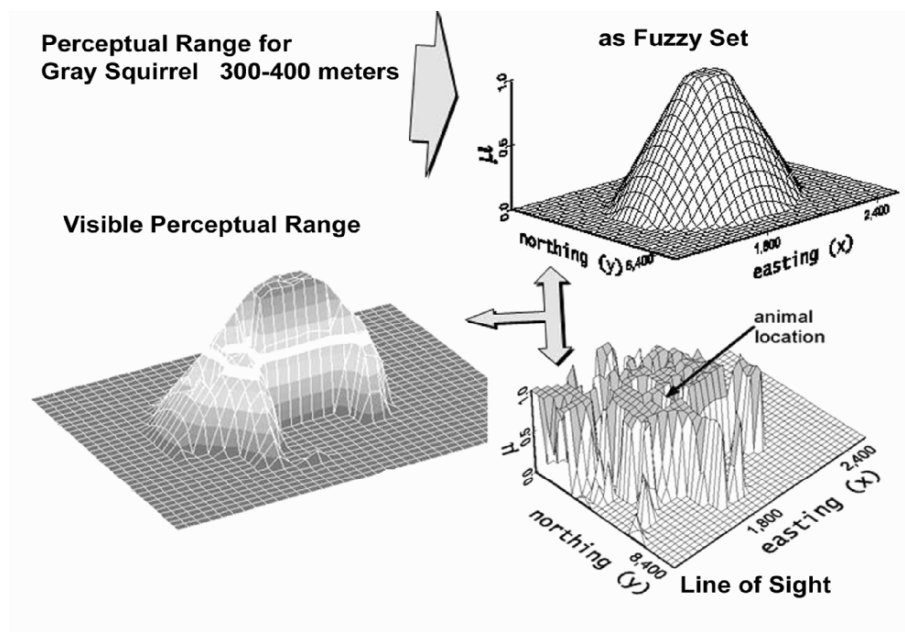


Figure 4. Derivation of context sensitive perceptual range as a fuzzy set

addition, a line of sight set is calculated to represent the degree to which landscape elements are visible to a squirrel (i.e., animal object) up in a tree to arrive at a fuzzy set of visible landscape. The combination of perceptual range and visible landscape provided a set of landscape elements who were members to some degree in the visible perceptual range represented as a fuzzy set (see Figure 4). This joint fuzzy representation of perceptual range and visibility formed an individual object's view of the landscape database derived from the GIS layers. The landscape information was then used in a fuzzy-based decision-making model of that had two submodels. The movement submodel concerned the locational choice for the next move. Once the individual object had been moved to that location, the residence submodel was used to determine whether or not it would stay in the location. In other words, did that location meet the requirements for establishing a home range that would sustain an individual? Details of these submodels can be found in Robinson and Graniero (2005).

Since fuzzy set theory is particularly rich in set connectives, four decision models were used to simulate dispersal behavior. Three were based on fuzzy set theory using compensatory, noncompensatory, and Yager style connectives. For comparison, a fourth model was a crisp (nonfuzzy) model. Hence there were four populations (i.e., classes) of animal objects with each population using a different decision model. Figure 5 illustrates how the different decision models yield differing spatially explicit behavior. Thus, this approach illustrated the ability of a fuzzy-based approach to represent variations in decision making as might occur within populations. Since success at finding a home range location is critical to support modeling of metapopulation dynamics and species spread, this result illustrates how fuzziness incorporated in a spatially explicit ecological model may help explain variations in rates of recolonization or dispersal.

Since squirrels tend to set up home ranges in areas not contested by conspecifics, simulations were for landscapes with different levels of conspecific coverage. The base line situation was the case of social fencing where there were no "holes" in the conspecific landscape. Variations in the base line were implemented by randomly allowing for 20%, 40%, and 60% percent of the landscape to be composed of unoccupied "holes." A conspecific "hole" was therefore a spatially explicit goal for a disperser. Results of these simulations showed that variations in landscape perception, and consequent use of the information in the decision submodels, did have an effect on the simulated animal's ability to find a suitable location for establishing a home range. Furthermore, in all situations the simulated animals using the fuzzy-based submodels were more successful at finding a home range than were those using a submodel based on crisp logic (Robinson and Graniero, 2005).

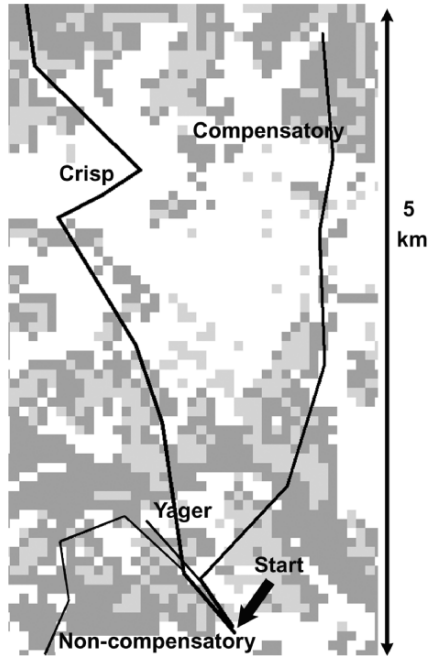


Figure 5. An example of the effect of differing decision models on spatial behavior of an individual given the same starting location and surrounding landscape. Note that the gray regions are deciduous forest while the white regions tend to be nonforest

The example shown in Figure 5 illustrates that a range of behaviors within a single species can be modeled. This is more realistic than assuming all animal agents have a single, identical decision submodels. In addition, the simulations yielded plausible results that do not rely upon stochasticity, thus providing a more satisfying framework for understanding how movement behavior may be influenced by landscape and ecological factors as well as the decision submodels. However, it remains to be determined whether or not such results would actually have significant impact on the ultimate demographic dynamics of a SEPM.

5. Hypothesis Testing

Although fuzzy techniques have been used to classify landscape elements, characterize model parameters, and characterize within habitat/community variations, they have rarely been used to directly test hypotheses. When discussing fuzzy techniques in conjunction with statistical analysis two primary issues arise:

Use of fuzzy techniques versus statistical analysis: In other words, can fuzzy techniques be used to appropriately accept or reject a hypothesis?

Use of fuzzy techniques along with statistical techniques: In other words, can the two approaches be used in a complementary fashion rather than as competing paradigms?

In the end, there is some question as to whether or not fuzzy techniques should, or even can, be evaluated against nonfuzzy techniques. One of the problems with constructing such an evaluation is the specification of a metric of evaluation that is meaningful to both.

5.1. FUZZY TECHNIQUES VERSUS STATISTICAL ANALYSIS

In habitat modeling it is common to see logistic regression used as if the assumption of presence/absence has been met. This, and other approaches, relies upon probability theory via the use of inferential statistics. As an alternative to this approach fuzzy-based techniques of evaluating habitat using presence data are beginning to emerge (Hill and Binford, 2002; Hodgson, in progress). For example, Hill and Binford (2002) present a fuzzy-based tests that allows habitat models to be accepted or rejected according to the degree to which they under/overestimate habitat potential. They suggest that their approach provides a means of not only using approximate reasoning, but also a means of providing defensible conclusions. In the context of ecological modeling culture, this issue remains a significant challenge for those researchers seeking to incorporate fuzzy sets in their modeling efforts.

Fuzzy clustering has been used to investigate the description of vegetation communities (Equihua, 1990) and spatial structure of animal populations (Feoli and Zerihun, 2000; Schaefer and Wilson, 2002). Schaefer and Wilson (2002) showed how fuzzy clustering can depict the ambiguous structure of caribou and walleye population's spatial structure. The advantage over the hard clustering technique was that it was able to convey the degree of uncertainty of belonging to a population cluster of both unambiguous as well as problematic individuals. Using both crisp and fuzzy clustering techniques Nicholls and Tudorancea (2001) demonstrated how fuzzy cluster analysis whereby species can be ranked according to their contribution to the structure of the dendrogram. Until this demonstration there had been few good methods for accomplishing this quantitatively (Nicholls and Tudorancea, 2001).

5.2. FUZZY TECHNIQUES WITH STATISTICAL TECHNIQUES

As suggested by Smithson (2005), it seems reasonable to consider that the joint, complementary use of fuzzy and statistical techniques as a powerful way of testing hypotheses. Equihua (1990) used analysis of variance to compare the results of fuzzy clustering with a crisp approach. In an analysis of bat habitat in a disturbed tropical forest environment, Medellín et al. (2000) used generalized linear models to regress bat community variables against a fuzzy membership score derived from fuzzy clustering. However, the statistical analysis of fuzzy membership patterns remains relatively unknown in GIS-based ecological modeling even though there are now techniques out there that will explicitly test for spatially explicit patterns (Hagen-Zanker et al., 2005).

Rather than using statistical analysis simply as an analysis of the outcome of a fuzzy-based exercise, it has been demonstrated in other fields that it can be used to choose fuzzy membership functions and form fuzzy rules. Focusing on the problem of uncertain parameter estimation in a hydroecological model, Mackay et al. (2003) illustrate the use of a two-stage methodology, where in the first stage Monte Carlo simulations are run. Then each simulation result is evaluated by regressing simulated evaporative fraction from a regional hydroecological model and surface temperature. For each regression, the coefficient of determination (R^2) is used as a fuzzy measure of the goodness of fit for its respective simulation result. Hence the fuzzy set is composed of the set of R^2 measures for all simulations, to which a fuzzy information-theoretic tool is applied to form a restricted set in which only good simulations are retained. A restricted set is used as an ensemble solution in the second stage of parameter estimation. This illustrates the potential for a combined statistical-fuzzy approach to address the issue of uncertainty in model parameters. However, it has yet to be developed in a convincing manner in the context of SEPM.

6. Software

The cost of developing and implementing fuzzy computational tools may have been, in the past, an obstacle to incorporating fuzzy sets in GIS-based ecological modeling. To some extent, it remains an issue since it is still relatively rare for a commercially available GIS software system to support fuzzy information processing. A notable exception is the IDRISI GIS software package (Eastman, 1999). However, there are add-on software works, not commercially available, that are coupled with commercially available GIS products such as ESRI's products (Benedikt et al., 2002; Yanar and Akyurek, 2006). However, to date these tools have not yet had

substantial presence the ecological modeling literature. Other tools, outside GIS software packages, have been developed.

Although the code for the fuzzy c-means algorithm was published in 1984 (Bezdek et al., 1984), fuzzy clustering remains a relatively infrequently used technique in ecological modeling. Shortly after Bezdek et al.'s (1984) paper appeared, applications in the classification of land cover using remote sensing data did appear (Fisher and Pathirana, 1994; Robinson and Thongs, 1986). Even though it is now a technique well known to remote sensing specialists, the additional information afforded by a fuzzy classification of land cover is generally not exploited by GIS-based ecological modeling efforts. The availability of software such as the FuzME (Minasny et al., 2002) package have supported the use of fuzzy clustering in ecological modeling (Bridges, 2003; Schaefer and Wilson, 2002). In addition to FuzME, the fuzzy c-means and FANNY techniques reported in Nicholls and Tudorancea (2001) are now freely available as part of the R statistical project (Dimitriadou et al., 2005; Maechler, 2005). The experience of fuzzy clustering illustrates that the cost and availability of software to support fuzzy techniques in GIS-based ecological modeling is a very real issue. The challenge is to not only make these techniques available but also to illustrate their use, utility, and availability to the ecological modeling community at large.

7. Concluding Comments

Fuzzy clustering techniques have been used sporadically in habitat studies that form the basis for inferring habitat patches from a GIS database. The results of using fuzzy clustering techniques suggest that they preserve, and represent, more faithfully the spatial variability of habitat than do the more traditional crisp techniques. This is of particular importance when incorporating rare community types in habitat analysis since crisp techniques tend to overlook, or eliminate, their contribution to habitat. Thus, one of the main issues and challenges of incorporating fuzzy sets in ecological modeling is the degree to which the results are ecologically meaningful, supportive, or not, of a given hypothesis, and ecologically plausible.

Perhaps more importantly is the modeling process itself. The degree to which the model and its parameterization have plausible ecological interpretations is a challenge. Like most applications in the ecological literature, this chapter's discussion has exclusively considered type 1 fuzzy sets. Since uncertainty is the fundamental reason that models incorporate fuzzy sets, it is important to consider the fact that the fuzzy membership functions are themselves subject to uncertainty. Such uncertainty is modeled using type 2 fuzzy sets (Fisher et al., in press). It would be useful

for the purposes of constructing fuzzy ecological models to have an appreciation of whether or not, the added computational complexity and burden of type 2 fuzzy sets would result in any significant improvement of the models and their performance.

This chapter has but scratched the surface in presenting a variety of issues and challenges regarding the incorporation of fuzzy sets in GIS-based ecological modeling. The topics are an intersection of research and applications in geographic information science, behavioral ecology, landscape ecology, remote sensing, fuzzy systems, and database theory.

Acknowledgments

The author gratefully acknowledges the partial support of Discovery Grant RGPIN 44611-02 from the Natural Sciences and Engineering Research Council (NSERC) of Canada. The partial support of NATO to support the author's participation in the workshop held in Kyiv, Ukraine is most gratefully acknowledged.

References

- Anot, C., Fisher, P. F., Wadsworth, R., and Wellens, J., 2004, Landscape metrics with ecotones: pattern under uncertainty, *Landscape Ecology* **19**:181–195.
- Austen, M. J., Francis, C. M., Burke, D. M., and Bradstreet, M. S. W., 2001, Landscape context and fragmentation effects on forest birds in Southern Ontario, *The Condor* **103**:701–714.
- Benedikt, J., Reinberg, S., and Riedl, L., 2002, A GIS application to enhance cell-based information modeling, *Information Sciences* **142**:151–160.
- Bezdek, J. C., Ehrlich, R., and Full, W., 1984, FCM: the fuzzy c-means clustering algorithm, *Computers and Geosciences* **10**:191–203.
- Bone, C., Dragicevic, S., and Roberts, A., 2006, A fuzzy-constrained cellular automata model of forest insect infestations, *Ecological Modelling* **192**:107–125.
- Bordogna, G., Chiesa, S., and Geneletti, D., 2006, Linguistic modelling of imperfect spatial information as a basis for simplifying spatial analysis, *Information Sciences* **176**:366–389.
- Bridges, L. M., 2003, *Spatial Scale and Environmental Structure: Habitat Selection of Adult Grey Squirrels (Sciurus carolinensis) in Central Ontario*, Master's thesis, Trent University.
- Brotons, L., Thuiller, W., Araujo, M. B., and Hirzel, A. H., 2004, Presence-absence versus presence-only modelling methods for predicting bird habitat suitability, *Ecography* **27**:437–448.
- Burgman, M. A., Breininger, B. W., Duncan, B. W., and Ferson, S., 2001, Setting reliability bound on habitat suitability indices, *Ecological Applications* **11**:70–78.
- Carroll, C., Zielinski, W. J., and Noss, R. F., 1999, Using presence-absence data to build and test spatial habitat models for the Fisher in the Klamath Region, U.S.A., *Conservation Biology* **13**:1344–1359.

- Comber, A., Fisher, P., and Wadsworth, R., 2005, What is land cover? *Environment and Planning B: Planning and Design* **32**:199–209.
- DeGenst, A., Canters, F., and Gulink, H., 2001, Uncertainty modeling in buffer operations applied to connectivity analysis, *Transactions in GIS* **5**:305–326.
- Dettmers, R. and Bart, J., 1999, A GIS modeling method applied to predicting forest songbird habitat, *Ecological Applications* **9**:152–163.
- Dimitriadou, E., Hornik, K., Leisch, F., Meyer, D., and Weingessel, A., 2005, *The e1071 Package*, R Project for Statistical Computing, <http://cran.r-project.org/>.
- Dragicevic, S., 2005, Multi-dimensional interpolations with fuzzy sets, In: *Fuzzy Modeling with Spatial Information for Geographic Problems*, Petry, F. E., Robinson, V. B., and Cobb, M. A., eds., Springer, Berlin, pp. 143–158.
- Eastman, J. R., 1999, *Idrisi32 Guide to GIS and Image Processing*, Clark Labs, Worcester, MA.
- Equihua, M., 1990, Fuzzy clustering of ecological data, *Journal of Ecology* **78**:519–534.
- Feoli, E. and Zerihun, W., 2000, Fuzzy set analysis of the Ethiopian rift valley vegetation in relation to anthropogenic influences, *Plant Ecology* **147**:219–225.
- Fielding, A. H. and Bell, J. F., 1997, A review of methods for assessment of prediction errors in conservation presence/absence models, *Environmental Conservation* **24**:38–49.
- Fisher, P. F. and Pathirana, S., 1994, The evaluation of fuzzy membership of land cover classes in the suburban zone, *Remote Sensing of Environment* **34**:121–132.
- Fisher, P., Cheng, T., and Wood, J., in press, Higher order vagueness in geographical information: empirical geographical population of type-n fuzzy sets, *Geoinformatica* **29**:106–128.
- Graniero, P. A. and Robinson, V. B., 2003, A real-time adaptive sampling method for field mapping in patchy, heterogeneous environments, *Transactions in GIS* **7**:31–54.
- Graniero, P. A. and Robinson, V. B., in press, A probe mechanism to couple spatially explicit agents and landscape models in an integrated modelling framework, *International Journal of Geographical Information Science*
- Hagen-Zanker, A., Straatman, B., and Uljee, I., 2005, Further developments of a fuzzy set map comparison approach, *International Journal of Geographical Information Science* **19**:769–785.
- Hill, K. E. and Binford, M. W., 2002, The role of category definition in habitat models: practical and logical limitations of using Boolean, indexed, probabilistic, and fuzzy categories, In: *Predicting Species Occurrences: Issues of Accuracy and Scale*, Scott, J. M., Heglund, P. J., Morrison, M. L., Hafer, J. B., Raphael, M. G., Wall, W. A., and Samson, F. B., eds., Island Press, Washington, DC, pp. 97–106.
- Hodgson, P., in progress, *An Exploratory Study in the Use of Fuzzy Set to Model the Habitat of Worm-eating Warblers in the Land Between the Lakes National Recreation Area (tentative title)*, Master's thesis, University of Toronto.
- Jenkins, C. N., Powell, R. D., Bass, O. L., Jr., and Pimm, S. L., 2003, Why sparrow distributions do not match model predictions, *Animal Conservation* **6**:39–46.
- Lima, S. L. and Zollner, P. A., 1996, Towards a behavioral ecology of ecological landscapes, *Trends in Ecology and Evolution* **11**:131–135.
- Liu, J., Dunning, J. B., Jr., and Pulliam, H. R., 1995, Potential effects of a forest management plan on Bachman's sparrows (*Aimophila aestivalis*): linking a spatially explicit model with GIS, *Conservation Biology* **9**:62–75.
- Lurz, P. W. W., Rushton, S. P., Wauters, L. A., Bertolino, S., Currado, I., Massoglio, P., and Shirley, M. D. F., 2001, Predicting grey squirrel expansion in North Italy: a spatially explicit modelling approach, *Landscape Ecology* **16**:407–420.

- Mackay, D. S., Samanta, S., Ahl, D. E., Ewers, B. E., Gower, S. T., and Burrows, S. N., 2003, Automated parameterization of land surface process models using fuzzy logic, *Transactions in GIS* 7:139–153.
- Maechler, M., 2005, *The cluster Package*, R Project for Statistical Computing, <http://cran.r-project.org/>.
- Mech, S. G. and Zollner, P. A., 2002, Using body size to predict perceptual range, *Oikos* 98:47–52.
- Medellin, R. A., Equihua, M., and Amin, M. A., 2000, Bat diversity and abundance as indicators of disturbance in neotropical rainforests, *Conservation Biology* 14:1666–1675.
- Minasny, B. and McBratney, A. B., 2002, FuzME version 3.5, Australian Centre for Precision Agriculture, The University of Sydney, Australia, December 12, 2003.
- Nicholls, K. H. and Tudorancea, C., 2001, Application of fuzzy cluster analysis to Lake Simcoe crustacean zooplankton community structure, *Canadian Journal of Fisheries and Aquatic Sciences* 58:231–240.
- Odeh, I. O. A., McBratney, A. B., and Chittleborough, D. J., 1990, Design of optimal sample spacings for mapping soil using fuzzy k-means and regionalized variable theory, *Geoderma* 47:93–112.
- Oehler, J. D. and Litvaitis, J. A., 1996, The role of spatial scale in understanding responses of medium-sized carnivores, *Canadian Journal of Zoology* 74:2070–2079.
- Olden, J. D., Schooley, R. L., Monroe, J. B., and Poff, N. L., 2004, Context-dependent perceptual ranges and their relevance to animal movements in landscapes, *Journal of Animal Ecology* 73:1190–1194.
- Orrock, J. L., Pagels, J. F., McShea, W. J., and Harper, E. K., 2000, Predicting presence and abundance of a small mammal species: the effect of scale and resolution, *Ecological Applications* 10:1356–1366.
- Pearce, J. L., Cherry, K., Drielsma, M., Ferrier, S., and Whish, G., 2001, Incorporating expert opinion and fine-scale vegetation mapping into statistical models of faunal distribution, *Journal of Applied Ecology* 38:412–424.
- Pearce, J. L. and Boyce, M. S., 2006, Modelling distribution and abundance with presence-only data, *Journal of Applied Ecology* 43:405–412.
- Robinson, V. B., 1988, Some implications of fuzzy set theory applied to geographic databases, *Computers, Environment, and Urban Systems* 12:89–97.
- Robinson, V. B., 2002, Using fuzzy spatial relations to control movement behavior of mobile objects in spatially explicit ecological models, In: *Applying Soft Computing in Defining Spatial Relations*, Matsakis, P. and Sztandera, L. M., eds., Physica-Verlag, Heidelberg, pp. 158–178.
- Robinson, V. B., 2003, A perspective on the fundamentals of fuzzy sets and their use in geographic information systems, *Transactions in GIS* 7:3–30.
- Robinson, V. B. and Graniero, P. A., 2005, Spatially explicit individual-based ecological modeling with mobile fuzzy agents, In: *Fuzzy Modeling with Spatial Information for Geographic Problems*, Petry, F. E., Robinson, V. B., and Cobb, M. A., ed., Springer, Heidelberg, pp. 299–334.
- Robinson, V. B. and Thongs, D., 1986, Fuzzy set theory applied to the mixed pixel problem of multispectral landcover databases, In: *Geographic Information Systems in Government*, Opitz, B. K., ed., A. Deepak Publishing, Hampton, VA, pp. 871–885.
- Ruckelshaus, M., Hartway, C., and Kareiva, P., 1997, Assessing the data requirements of spatially explicit dispersal models, *Conservation Biology* 11:1298–1306.
- Schaefer, J. A. and Wilson, C. C., 2002, A fuzzy structure of populations, *Canadian Journal of Zoology* 2235–2241.

- Seoane, J., Bustamante, J., and Diaz-Delgado, R., 2005, Effect of expert opinion on the predictive ability of environmental models of bird distribution, *Conservation Biology* **19**:512–522.
- Skubic, M., Blisard, S., Bailey, C., Adams, J. A., and Matsakis, P., 2004, Qualitative analysis of sketched route maps: translating a sketch into linguistic descriptions, *IEEE Transactions on Systems, Man, and Cybernetics* **34**:1275–1282.
- Smithson, M., 2005, Fuzzy set inclusion: linking fuzzy set methods with mainstream techniques, *Sociological Methods and Research* **33**:431–461.
- Yanar, T. A. and Akyurek, Z., 2006, The enhancement of the cell-based GIS analyses with fuzzy processing capabilities, *Information Sciences* **176**:1067–1085.
- Zadeh, L. A., 1978, Fuzzy sets as a basis for a theory of possibility, *Fuzzy Sets and Systems* **1**:3–28.
- Zhu, A. X., Hudson, B., Burt, J., Lubich, K., and Simonson, D., 2001, Soil mapping using GIS, expert knowledge, and fuzzy logic, *Soil Science Society of America Journal* **65**:1463–1472.
- Zollner, P. A. and Lima, S. L., 2005, Behavioral tradeoffs when dispersing across a patchy landscape, *Oikos* **108**:219–230.