



This is to certify that the

dissertation entitled

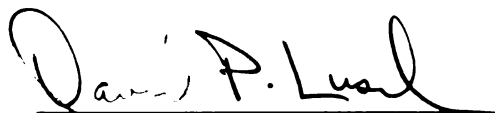
SPATIAL MODELING OF SEASONAL HOMES
IN THE UPPER GREAT LAKES STATES

presented by

Bradley Aric Shellito

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Geography



Major professor

Date June 22, 2001

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE
SEP 18 2002		
NOV 19 2002		
JUL 31 2004		
06 04 04		
APR 03 2007		
12 27 06		

**SPATIAL MODELING OF SEASONAL HOMES IN THE UPPER GREAT LAKES
STATES**

By

Bradley Aric Shellito

A DISSERTATION

**Submitted to
Michigan State University
In partial fulfillment of the requirements
For the degree of**

DOCTOR OF PHILOSOPHY

Department of Geography

2001

ABSTRACT

Seasonal home location is influenced by a variety of factors, including many related to land use. Water bodies, natural areas, agriculture, and distance from cities are some of the qualities that affect the location and distribution of seasonal homes. This research examines the 1990 seasonal home distribution for the Upper Great Lakes States of Michigan, Wisconsin, and Minnesota at a minor civil division (MCD) level.

Geographic Information Systems are used to construct independent variables over the three state level. Modeling techniques, including the application of logit regression and artificial neural networks, are used to examine this distribution. These models are used to determine the principal predictors of seasonal home distribution in the region. The research also serves as a test of the applicability of combining these modeling approaches with GIS and recreational tourism related theory. Conclusions are drawn about the principal predictors of seasonal home distribution as well as a comparison of modeling approaches. Finally, future directions and expansions of this research are discussed.

For my parents, David and Elizabeth

ACKNOWLEDGEMENTS

I would like to extend my deepest thanks to the professors who have helped guide me through this process. Grateful appreciation and thanks goes out to Dr. Bryan Pijanowski for all of his help with the research and data involved with this project and for all of his assistance and guidance over the last few years. Very special thanks also goes to Dr. David Lusch for all of his help, direction and assistance, along with serving as committee chair. Thanks also go out to Dr. Jon Burley, Dr. David Campbell, and Dr. Daniel Stynes for invaluable assistance, advice, and patience in completing this research.

I would also like to thank Dr. Daniel Brown for all of his help and guidance during my first few years at MSU and the formative stages of this dissertation. Also, thanks go to Dr. Michael Chubb for his help as well. Special thanks go to Mr. James Biles for his endless patience as well as suggestions and advice regarding the statistics and modeling approaches used in this project. Thanks for statistical help must also go out to Dr. Bruce Pigozzi and Dr. Ashton Shortridge. Thanks also to Mr. Gaurav Manik and Mr. Snehal Pithadia for their assistance with the neural networks used here.

Thanks also go out the people and agencies whose contributions to the data sources made this project possible: Dr. Mike Vasievich of the US Forest Service and the Great Lakes Ecological Assessment Project; Dr. Stuart Gage, Director of the Computational Ecology and Visualization Laboratory; the Michigan and Minnesota branches of RESAC; the Michigan and Wisconsin Departments of Natural Resources; the MDEQ; and the NASA LULCC project.

Finally, I'd like to say thank you to Dr. Ron Shaklee, for starting me down this road in the first place.

TABLE OF CONTENTS

LIST OF TABLES	vii
LIST OF FIGURES	viii
CHAPTER 1	
INTRODUCTION	1
Research Questions and Objectives	1
Definitions	2
Format of the Dissertation	4
CHAPTER 2	
LITERATURE REVIEW	6
Introduction	6
Theory in Geography	7
Land Use and Landscape Ecology	10
Seasonal Homes and Land Use	11
Seasonal Homes and Public Lands	16
Seasonal Home Development	18
Seasonal Homes' Link to Recreation and Tourism	21
Tourism Development and Land Use	22
Modeling Recreation and Tourism Issues	23
Geographic Information Science	25
GIS and Recreational Tourism	26
GIS, Seasonal Homes, and Real Estate	27
Conclusion	27
CHAPTER 3	
STUDY AREA	29
The Upper Great Lakes States and Seasonal Homes	29
Spatial Distribution of Seasonal Homes	32
Conclusion	35
CHAPTER 4	
METHODS	37
Introduction	37
Modeling Framework	37
Unit of Analysis	39
Dependent Variable	40
Spatial Autocorrelation	40
Independent Variables	41
Development of Independent Variables	43
Correlation of Dependent and Independent Variables	51
Principal Components Analysis	51
Logistic Regression and the Logit Model	53
The Artificial Neural Network Model	54

Conclusion	57
CHAPTER 5	
MODEL RESULTS	58
Introduction	58
Spatial Autocorrelation	58
PCA Results	60
Logit Model Results	63
Neural Network Model Results	64
Examination of Independent Variables	64
Logit Model	65
Neural Network Model	65
The Principal Predictors	67
Spatial Distribution of Seasonal Homes	67
Model Comparison	69
The Spatial Error Signal and Residual Analysis	71
Logit Estimates vs. Neural Network Estimates	77
Conclusion	77
CHAPTER 6	
DISCUSSION	79
Introduction	79
Dependent Variable	79
Independent Variables	82
Bias in the System	84
Methods and Model Building	89
The Science Under the System	91
Implications for Land Use and Potential Land Use Change	93
Conclusion	98
CHAPTER 7	
FUTURE DIRECTIONS	99
A Baseline Approach	99
Regionalization	101
Temporal Modeling	102
Adapting for Land Use Change	105
Other Applications of the Framework	108
Methodological Exploration	109
Spatial Statistics Approaches	111
Conclusion	112
BIBLIOGRAPHY	113
APPENDIX A: Independent Variable Maps	126

LIST OF TABLES

Table 3.1: Metro MCDs and their numbers of seasonal homes, 1990. . . .	34
Table 3.2: Types of MCDs and the average % seasonal they contain. . . .	35
Table 4.1: Bivariate Correlations.	52
Table 5.1: Initial Eigenvalues.	60
Table 5.2: PCA Rotated Component Matrix.	61
Table 5.3: Initial beta values and solved values.	63
Table 5.4: Reduced variable runs of neural network.	66
Table 5.5: Comparison of model rankings of principal predictors.	67
Table 5.6: Mean Absolute Deviation of the modeling approaches.	71
Table 7.1: US Housing and Seasonal Homes by year.	103

LIST OF FIGURES

Figure 3.1: # of seasonal homes in the Upper Great Lakes States, 1970-1990.	31
Figure 3.2: Percentage of the total housing stock that is seasonal	33
Figure 4.1: Basic Architecture of a MLP Neural Network.	55
Figure 5.1: Variogram of the dependent variable (distance in meters).	59
Figure 5.2: Variogram of the residuals of the dependent variable (in meters).	59
Figure 5.3: Logit and Neural Network Residuals.	72
Figure 5.4: Logit Residuals with Public Land Boundaries Overlayed.	73
Figure 5.5: Selected Residuals.	74
Figure 7.1: # of seasonal homes in the United States, 1940-1990.	104

Chapter 1: Introduction

Research Questions and Objectives

The notion of having a second residence, one that is only used part-time, is extremely attractive to many people. This concept of a part-time residence encompasses a spectrum of housing types, ranging from a lakefront house, to a time-share condominium, to a simple hunting cabin. These second (or seasonal) homes are frequently used as vacation homes, perhaps for weekend getaways or longer trips. Seasonal homes are also used as a base from which to enjoy amenities (such as hunting, fishing, or boating) or a place to store materials to be used at another time (such as a snowmobile or a boat). Seasonal homes also have the potential to transition to a new permanent home for retirees.

Numerous factors play into the location of seasonal homes. Some of these depend on the personal needs, wants, and desires of the person buying the seasonal home, including income level, land price, and reasons for why this second home is desired. However, several other destination-based factors play a role in the location of seasonal homes, involving landscape features and land use characteristics. For instance, the growth of urban land use may be seen as a “push” factor for people to leave a highly populated urban area for more rural settings. Likewise, land cover features such as forests, water bodies, and more natural environments may be seen as “pull” factors to bring seasonal homes into these areas.

This dissertation research will bring these concepts of seasonal home location together with land use in a geographic context. This will aid in answering the questions

of (1) what are the principal predictors of seasonal home location and (2) can this link between the pattern of land use and landscape features and the process of seasonal home distribution be modeled in a spatial GIS-based framework?

There are several objectives to this project:

1. To describe the spatial distribution of seasonal homes within the Upper Great Lakes States region of Michigan, Wisconsin, and Minnesota.
2. To identify the principal predictors of seasonal home distribution in this region.
3. To test a GIS-based modeling framework and various modeling approaches in examining the relationship between the patterns on the landscape and seasonal home distribution.

Definitions

In the 1990 census, second homes are labeled as homes kept for “seasonal, recreational, or occasional use.” This is defined as: “vacant units used or intended for use only in certain seasons or for weekend or other occasional use throughout the year. Seasonal units include those used for summer or winter sports recreation, such as beach cottages or hunting cabins. Seasonal units may also include quarters for herders or loggers. Internal ownership units, sometimes called shared ownership or time-share condominiums, also are included here” (Census 1990). (The Census Bureau seems to label these factors as “vacation homes” in the current analysis (Census Bureau 2000)). Other categories not included under this definition are housing units classified as: rent, for sale, rental or sold – not occupied, migrant workers, and other vacant (Census 1990).

There are several concerns in defining secondary and seasonal homes including:

- (1) determining the difference between rental properties and permanently owned ones;
- (2) distinguishing between recreational homes and migrant housing; and (3) conversions of second homes to permanent ones.

The third situation is one of individual choice and can change over the course of time. The temporal or psychological factors involved with this change are outside the scope of this project (see Stewart 1994). To adequately address the problems of the first and second situations, as well as to provide a clearer definition of second homes for this project, the definitions and framework used by Spotts (1991) will be used here. The US Census Bureau contains two labels for housing units in its 1980 census: (1) “vacant seasonal and migratory” and (2) “vacant year round – held for occasional use.” Spotts (1991) combines the counts from these two categories (which includes recreational cabins, cottages, condominiums for rent, migratory worker housing, time-share condos, shared-ownership condos). The Census Bureau also notes that seasonal homes may fall into either category, thus the numbers from both must be used. Of these, only migratory housing is not needed in the analysis. However, as Spotts (1991) notes, there are so few of these type of housing units that their inclusion in a statistical analysis is inconsequential. For instance, in the 1970 census, migrant housing made up only 3% of all secondary homes in the entire US. However, this issue of migratory housing in the seasonal home definition was excluded in the 1990 Census.

For the purposes of this dissertation, the term “seasonal homes” will be used to refer to the object of study, and can be considered interchangeable with “second home” and “vacation home.”

Format of the Dissertation:

The chapters of this dissertation are organized as follows:

Chapter 2 discusses where this literature that this project draws from. First, it examines the role of how this research fits in the overall geographic theory and literature. It then examines the development of seasonal homes in the context of the literature. Several other branches of literature pertaining to seasonal home development are touched on as well, including amenity migration, real estate, and public lands. The role of land use and land cover in the seasonal home development process is examined as well. The linkage between seasonal homes and areas of recreation and tourism is examined as well.

In addition, further literature describing recreational tourism as a land use and the associated connections between tourism development and land use is examined, as well as a brief look at human / environmental interaction based on the paradigms of landscape ecology and its relationship to this project is also reviewed. The role of Geographic Information Systems as well as its connection to seasonal homes and tourism modeling issues is also examined.

Chapter 3 discusses the study area of the Lake States of Michigan, Wisconsin, and Minnesota in terms of their potential for seasonal home and recreational tourism development. The development of seasonal homes over time in these states is examined, as well as the geographic and land use factors in these states which contribute to the process.

Chapter 4 discusses the methodology and approach to the problem, in terms of the development of variables and models. The GIS framework used in this project is

examined, along with descriptions of the variables and their construction, along with examination of the modeling approaches.

Chapter 5 looks at the results of the models and what these results mean in terms of the objectives stated above.

Chapter 6 provides a discussion that places this study back within the literature and gives further insight into the process of seasonal home location. The role of this project in the greater body of theory is examined, as well as the importance of this project in practical application. In addition, the broader implications of this project in terms of concepts of development and land use change are examined. The modeling approaches are examined and conclusions drawn about them. Some limitations of the project are also discussed.

Finally, Chapter 7 details future directions that can be developed from this research.

Chapter 2: Literature Review

Introduction

This project employs many concepts from the discipline of geography, but is grounded within three bodies of literature and theory: landscape ecology, tourism development, and geographic information science. Each of these broad areas of knowledge are not easily defined or bounded, leading to considerable overlap between fields. The object of this review is not to provide an in-depth look at all literature available on these areas, but to examine the research in them pertinent to this project.

Tourism development literature encompasses a wide variety of topics and fields. Seasonal home literature is primarily included under this umbrella, yet there are other branches of literature pertinent to the seasonal home process, including amenity migration.

Likewise, landscape ecology is a wide field with a multitude of definitions, which represents the study of human / environment interactions. The focus here is to touch on some of the relevant theory of this field and how it applies to the problem at hand in terms of analyzing pattern and process.

Finally, geographic information science encompasses a wide variety of fields as well. The aim of this section is to examine the nature of the “science” and the “system” of GIS, as well as its applicability to problems of spatial distribution, tourism, and seasonal home questions.

Theory in Geography

This project has connections to geographic theory as well as several broader implications and links to other branches of theory. In Explanation in Geography, David Harvey (1969) examines Haggett's (1965) groupings of the five major themes in geographic theory. These are: areal differentiation, the landscape, man-environment, spatial distribution, and the "geometric" theme. As Harvey (1969) notes, the five themes are "not mutually exclusive or entirely inclusive" of all branches and studies of geography, but provide some operational parameters and definitions from which to form concepts. The spatial distribution theme can be seen to encompass areas of locational analysis and providing a description and explanation of the distribution of a phenomenon. In other words, this theme may be seen as trying to describe and explain why things are located where they are.

Of these broad areas, this study fits best within the "spatial distribution" theme. It focuses on the concepts of where seasonal homes are located (or not located) and why they are where they are. However, this project has relations to the "man-environment" theme of the literature, typified as a view of geography as human ecology. Harvey (1969) notes that this may be the true overriding theme that lies at the heart of geography. However, Harvey does recognize the need for interdisciplinary research in the overall advancement of knowledge. The link between humans and the land is often associated with the locations of man in relation to land use.

Early work in linking location and land use came from David Ricardo's theories of land rent (referenced in Barlowe 1986). Here, the concept of rent referred to a surplus gained from a factor of production that is a result of a locational advantage. Ricardo

attributed land rent to issues of the fertility of the land – in which only the most fertile lands would be used and rent would arise on these fertile lands only when increases in demand lead to the use of less fertile lands (referenced in Barlowe 1986).

This connection between theories of land use and location continued in the work of Von Thunen (von Thunen 1826, referenced in Haggett et al. 1977). Von Thunen's model of rural land use examined the city in terms of an "isolated state" with no outside influences (a pre-industrialization notion). The model shows concentric rings emanating outward from the monocentric city to denote distinctive "zones" of land use (including dairying and farming, forest, field crops, and ranching). This model assumes that the farmers surrounding the market will produce crops with the highest value (or the highest "land rent"). Though outdated today, the Von Thunen model was an early theoretical model of the location of the city and the role of the area surrounding that location.

This location theory was followed by Weber (1909, referenced in Haggett et al. 1977). Weber's theory focuses on locating industrial sites to minimize distance and unnecessary movement. This theory showed that industrial sites will be located nearest the centers for raw materials, thus reducing cost and movement for industrial development.

Location theory and analysis was furthered by the work of Walter Christaller (1933) who developed Central Place Theory. This concept was used to explain the size and location of cities that specialize in selling goods and services. As cities would be centrally located, they are assumed to be "central places." He introduced two concepts: threshold (the minimum size market needed to keep a city in business) and range (the

maximum distance people will travel to purchase goods). Once a threshold size is established, the central place will try to expand out to the limit of the range.

Christaller also makes the distinction between “higher order” central places (those areas which provide more goods and services than other areas) and “lower order” central places (which provide less goods and services, but whose goods and services are consumed more frequently). He posits that higher order places are more widely distributed and fewer in number than lower order locations.

Losch (1940) built off of Christaller’s work by examining areas as a continuous series of centers rather than steps in a hierarchy. It is noted (Haggett et al. 1977) that Losch’s model produces patterns more like what exists in reality than Christaller’s. Likewise, Pred (1966) extended Christaller’s theory by examining the movement of goods and services between the ranks of cities rather than between cities themselves.

Another major contribution to location theory came from William Alonso (1964) with his theories of land rent. Alonso’s theory postulated that each of three types of land use (commercial, industrial, and housing) have a certain yield of rent for any given location, and that these three land uses are in competition for one another. The theory shows that commercial and industrial land uses will often be closer to the central city, along with lower-income housing. More expensive housing, which does not need a strong connection to the central city, will be found on the outskirts of the city where the cheaper land may be found. Thus, a segregated land use system will be created.

All of these theories relate land use and location in some way. According to theory, where certain types of areas (i.e. residential or industrial) are found are reliant about the locations of surrounding patterns of land use. Degrees of urbanization and

varying types of urban development rely on spatial patterns and usually a link to the landscape. It is this basic theory of spatial distribution, location theory, and its relation to land use that this project draws upon. The historical work in geography provided studies of “idealized” environments to draw conclusions about linkages between land use and location. However, this work has been expanded on by more recent theorists to encompass “real world” concepts and ideas into the theory. It is this linkage between the location of places and their relationship with elements of the landscape or land use that is part of the theoretical background of this project. In this fashion, this dissertation is set within the context of geography as a project to explain why things are where they are. This over-arching umbrella of location can be seen as one of the core tenets of geography (i.e. Harvey 1969, Haggett et al. 1977).

Land Use and Landscape Ecology

The study of land use in geography often falls under the umbrella of landscape ecology. The term “Landscape Ecology” was first coined by Carl Troll (1939) in his paper examining landscapes of Africa through aerial photography. He described it as “...the study of the entire complex cause-effect network between the living communities and their environmental conditions which prevails in a specific section of the landscape... and becomes apparent in a specific landscape pattern.” (Troll 1939, quoted in Thomas 2000). Zonneveld (1995) notes that this was an attempt by Troll to create “a marriage between biology and geography.” Later works would further embellish Troll’s basic definition and aims. Vink (1983) describes landscape ecology as “the study of relationships between phenomena and processes in the landscape or geosphere including

the communities of plants, animals, and man.” Forman and Godron (1986) define landscape ecology as “a study of the structure, function, and change in a heterogeneous land area composed of interacting ecosystems.” Naveh and Lieberman (1984) describe landscape ecology as “the recognition of the dynamic role of man in the landscape and the quest for the systematic and unbiased study of its ecological implications.” Likewise, Sanderson and Harris (2000) note that “it (landscape ecology) deals with them (humans) explicitly as entities and forcing functions on the landscape.” From these definitions, landscape ecology theory can be broadly interpreted as a study of the interaction between humans and the environment.

By combining the science and theory of these two broad fields, landscape ecology seeks to define the spatial patterns of the landscape and to also explain the processes behind these patterns, as well as the interactions resulting from them. The “ecology” in the field comes from the study of “the house of the Earth” (Zonneveld 1995). As Urban et al. (1987) note, “pattern is the hallmark of the landscape.” As Davis (1976) notes, geography remains a strong force in controlling land development and use. This project focuses on the issues of spatial distribution of land use and seasonal homes.

Seasonal Homes and Land Use

A wide variety of landscape features contribute to the location of seasonal homes. Coppock (1977) notes that, across the world, the main controlling factors of seasonal home distribution appear to be: the distance from major population centers; the quality and character of the landscape; the presence of sea, rivers, or lakes; the presence of other recreational resources; the availability of land; and the climates of importing or exporting

regions. Spotts (1991) notes that seasonal homes are concentrated in areas with lakes, rivers, forests, and other natural resources.

Water bodies provide access for a wide variety of outdoor recreational activities such as fishing, boating, and swimming. Lakes are a strong factor in seasonal home development (Coppock 1977, Styne et al. 1995). Chubb and Chubb (1981) note that the availability of sites along water-oriented areas in the United States and Canada such as lakes, rivers, or the Great Lakes has resulted in the development of seasonal homes at these locations. In a study of Michigan seasonal home owners, Tombaugh (1970) found that 55% owned a seasonal home on an inland lake, 24% on the Great Lakes, and 10% on a river or stream. In a study of three counties in the northern Lower Peninsula of Michigan, Gartner (1987) found that 57% of seasonal homeowners have their homes located on a water resource. Also, the Great Lakes provide a wide range of recreational opportunities (Chubb 1989), while lakes with houses built around them may have private boat docks or public launching facilities. Styne and Safronoff (1982) found that 30% of Michigan boat owners and 80% of those from outside the state owned seasonal homes within Michigan.

Second homes are found primarily in areas away from the “big cities” and heavily populated / dense urban areas (Coppock 1977). Coppock (1977) notes that the majority of seasonal homes appear to be 100 to 150 miles away from population centers. In Michigan, Tombaugh (1970) found that 70% of a sample of seasonal homes were within 200 miles of the owner’s first home. Thus, seasonal homes are often found outside the more urban areas in places of low population, i.e. rural or forested areas (Coppock

1977). The remoteness of these rural areas (Keller 2000) provides part of the draw of the area.

Previous models of second and seasonal home use in Canada (Lundgren 1974) identify seasonal home developments appearing away from an urban center and slowly being subsumed by growing urban development over time. Seasonal homes are often developed at areas away from crowded or overly developed urban areas (Lundgren 1974). Thus, the surrounding few miles is rural, but modern amenities are within a short drive. This boundary between urban and rural areas is referred to as “living space” around the cities (Cracknell 1967) or “recreation hinterlands” (Mercer 1970) for enjoyment of non-urban recreation. These more rural communities feature low population but still have some amenities and services on hand that residents of an area would require – i.e. grocery stores, gas stations, garages, bars, small restaurants, or local stores (Hart 1998). Nearness to medical treatment (i.e. health care facilities such as hospitals) plays a part in seasonal home development, particularly in retirement home communities (Hart 1998). An overabundance of “commercial tourism” sites (i.e. hotels, theme parks, or prime commercial tourist attractions) may be a detriment to seasonal home development (Hart 1998). It is also noted that few seasonal home developments are occurring in primarily agricultural regions (Chubb and Chubb 1981).

Travel to and from seasonal homes also plays a part in the location of seasonal homes. Coppock (1977) notes that accessibility depends on the quality and density of the road network. The use of major highways reflects the movement of a large number of people on direct routes (Wilder 1985). This could also be seen as reflective of the access needed to areas or a “distance to markets” function, allowing second home owners the

privacy of their seasonal home while still maintaining access to nearby roads or health-related facilities. “Easy travel” was rated the most important of the variables in the selection process by seasonal home buyers (Coppock 1977). Travel to and from seasonal homes can generate frequent trips, leading to a necessity for high-volume travel accessibility (Page and Getz 1997).

Seasonal homes are used in conjunction with winter sports, including snowmobiling, cross country skiing, downhill skiing, sledding, tobogganing, and ice skating. This has led owners of seasonal homes to keep their homes open and available well into the winter months to use as a place to stay while snowmobiling or skiing (Hart 1998).

Natural settings play a role in seasonal home location. Forested areas provide a great deal of outdoor recreation opportunities, and proximity to them is part of seasonal home development (Stynes et al. 1995). There are numerous forest-oriented seasonal homes within the United States (Chubb and Chubb 1981) providing homeowners with a variety of recreational or relaxation opportunities. Many people look to nature and outdoor recreation for peace of mind and a return to health (McHarg 1969). Factors such as these are difficult to quantify as landscape variables, but simulations of this aesthetic quality can be created. For instance, Loomis and Walsh (1997) note that size of an area (such as forests) is a good measure of its attractiveness. Similarly, forested areas provide a great deal of outdoor recreation opportunities, and proximity to them is part of seasonal home development (Stynes et al. 1995). Higher elevations (Chubb and Chubb 1981) and the variability of the landscape are factors in influencing the choice of location of seasonal homes. This desire for aesthetically pleasing landscapes may be linked to

seasonal home purchasers seeking a feel for a rural lifestyle or an alternative to an urban one (Page and Getz 1997).

Several studies have examined the attributes that contribute to an aesthetically pleasing landscape. Shafer et al. (1969) note the importance of vegetation within a view along with water being important factors. These two factors were also prominent in the Landscape Preference Model (Shafer and Tooby 1973) that identified the quantitative variables in a natural landscape that were significantly related to public preferences for that landscape. In a study of Lake Michigan beaches, Peterson and Neumann (1969) identified scenic natural beaches and the presence of trees and natural growth as factors contributing to the aesthetic quality of an area.

More recent studies in attempting to quantitatively model the landscape factors contributing to an aesthetically pleasing scene have incorporated GIS techniques. Bishop and Hulse (1994) present predictive equations of an area's aesthetic quality based on GIS-derived variables. Among the factors they designate as aesthetically pleasing are: rivers, amount of high slope, orchards, amount of forest, and range of relief. Similar factors were found by Crawford (1994) through the use of remote sensing techniques: rugged terrain, heavy tree cover, natural activities, outlook, diversity of landscape features, and orderly arrangement of these factors. In the visual preference model (Steinitz 1990) it was found that coastal development with a "historical character," water bodies, long distance views, a folded landscape (such as mountains or islands) and a diverse and well maintained vegetation distribution were all factors contributing to aesthetically pleasing landscapes. Also, Steinitz (1990) notes the sense of "mystery" being of great importance in a pleasing scene. This "mystery" feature was later mapped

(Bishop and Hulse 1994) and found to be made up of “locations on roads or trails with curvature and the edge zones between forest and the open land or water.”

Many factors have been identified that are associated with negative impressions of the aesthetic quality of an area. Large industrial structures and areas of industrial or commercial development (Crawford 1994) have been cited as being displeasing to view in a natural environment. Likewise, Steinitz (1990) notes that people do not wish to see a developed or urbanized landscape, evidence of crowded use, or tourist-oriented commercial development. Burley (1997) found that an abundance of humans, vehicles, buildings, and polluted and eroded land were considered undesirable factors.

Seasonal Homes and Public Lands

The amount of public land (inferring related outdoor recreation amenities such as hunting, fishing, hiking, etc.) available is a factor in seasonal home location (Coppock 1977). These lands are classified as National Parks, National Forests, National Lakeshores, State Parks, State Wildlife areas, and County parks. However, development of urban land use occurs within these areas – in fact, whole sections of the Huron National Forest contain towns, restaurants, bars, homes, farms, etc. Public lands provide a host of opportunities including hiking, hunting, boating, fishing, and camping.

There are several considerations to account for when dealing with public lands. First, urban and commercial development is often allowed on public lands. Thus, public lands cannot be excluded from consideration for seasonal home locations. For instance, the National Forests Ownership Adjustment Program (USDA 1986) was created to rearrange the ownership pattern of privately held lands within national forest boundaries

into more cost-effective and useable forms. The program seeks to take areas that are “low-priority” (i.e. not semiprimitive wilderness, recreational areas, or wildlife habitats) and exchange them for privately held lands that meet these criteria (USDA 1986). These non-priority lands are then used to facilitate community expansion, commercial, industrial, and residential purposes (USDA 1975). Thus, active building will be occurring in certain sites within public land boundaries.

Second, there may be some question as to the actual location of public land boundaries. Lands within the Lake States were originally surveyed by the General Land Office more than 120 years ago. However, due to natural decay, fires, and human destruction, many of the originally surveyed corner monuments have been lost (USDA 1975). With the questionable or unknown locations of legal markers, as well as the errors that may have been produced with the original surveying methods, a degree of uncertainty exists as to the true boundaries of public lands.

Lastly, public lands contain seasonal homes and thus must be included in any analysis of the subject. Seasonal homes within public land boundaries are officially classified as “recreation residence sites” (USDA 1986). Permits are annually issued (and re-issued) for existing summer home tracts (USDA 1975). In fact, there are specific regulations concerning how far recreational amenities (such as snowmobile or ATV trails) may be constructed from seasonal residences (USDA 1986).

However, there are strict controls on the amount of land that can be developed within public land boundaries. Also, one of the goals of the National Forests Landownership Adjustment Program is not only to acquire priority lands, but to acquire contiguous areas of 2500 acres or more (USDA 1986). Thus, while development is

allowed (for instance, towns containing schools, fast food restaurants, and other commercial developments exist in the heart of the Huron National Forest), it is limited to certain areas.

Seasonal Home Development

The notion of owning such things as a summer cottage, time-share condo, hunting cabin, or a part-time residence in another location away from home is a driving force in new development (Coppock 1977). These properties can also be referred to as “temporal residences” (Ragatz 1969). As such, seasonal homes need to be taken into account when dealing with issues of tourism development. Stynes et al. (1995) cite Waters’ (1990) findings that a failure to account for seasonal homes in recreational tourism studies may omit as much as 50% of domestic tourist activity.

Qualities such as market factors, supply and demand (Loomis and Walsh 1997), land valuation, and housing prices are all part of the decision making process of selecting a seasonal home. However, many of these choices are subjective and specific to the individual buyer. As Stewart (1994) notes, “real estate developers may aggressively market certain developments as ‘seasonal home areas,’ but a decision maker’s preferences for natural and social settings, recreational activities, their willingness to travel, and their financial resources have more to do with determining what is and what is not a seasonal home area.” While these type of variables are certainly part of the decision making process, their effect is a highly individual one – for instance, a wealthy person looking to purchase a seasonal home will not be constrained by land price or market fluctuations.

Manson and Groop (2000) have noted that migration patterns in the United States are of people moving out from the cities. They first move from the cities to the suburbs and then to rural (or exurban) areas. Likewise, Stewart (2000) notes a high rate of net in-migration to rural parts of the United States containing scenic or recreational amenities. This form of amenity migration is restructuring part of the rural landscape in the United States. No single factor can be distinguished as spurring amenity migration, but these include individual decisions (Williams and McIntyre 2000), choices made by the elderly or retired (Longino 2001), or tourism and seasonal home purposes (Stewart 2000).

Development of seasonal homes may be traced back to the construction of primitive hunting cabins or lodges (Hart 1998). These were places for hunting, fishing, and “getting away” from the busy lifestyle of civilization. As owners aged, these seasonal retreats were often winterized and turned into permanent residences for retirement (Hart 1998). Retirees have seen seasonal homes turning into their permanent homes and have sometimes purchased their second home with this purpose in mind (Stewart 1994). Similarly, seasonal homes can be seen as a shelter for access to nearby amenities or the home can be the actual focus of the recreational activity (Coppock 1977). In a classic feedback loop, some areas may have developed in response to seasonal homes, or some areas may have seen increased seasonal home development as a result of increasing development and popularity of a site.

Purchasing a seasonal home brings with it a number of changes, not only to a destination, but also to a person. For instance, the purchase of a seasonal home will likely influence the travel plans of a family (Stewart 2000). With the median price of a seasonal home (in 1999) being \$127,800 (NAR 2000) this represents a significant

investment for the buyer, beyond the other costs of property taxes and maintenance. The very nature of a seasonal home indicates that it will only be used for part of the year, adding on further costs for upkeep and security for the property during the times of the year that it stands vacant. Thus, the purchase of a seasonal home is not a lightly-made investment, and will probably become the vacation destination of choice for the purchaser for some time to come. This would cause a multitude of vacation trips to and from the seasonal home, or the use of the seasonal home for an extended period of time throughout the year.

The National Association of Realtors categorizes seasonal home buyers in 1999 as (NAR 2000): having a median income of \$68,000 and a median age of 43. Of these, 79% were married couples. From purchases made by the “baby boom generation”, an estimated 100,000 to 150,000 seasonal homes will be added each year.

As the number of seasonal homes in an area increases, there is an increase in population, and thus an increase of growth and development of the area. For example, in 1990 there were estimated to be more than 223,000 seasonal homes in the state of Michigan (Stynes et al. 1995), approximating 6% of all housing units in the state. Stynes et al. (1995) estimate that the economic impact related to these seasonal homes is between \$60 to \$75 million per year. In part, this may be due to seasonal homes facilitating longer stays in regions (Page and Getz 1997). Wisconsin seasonal home residents spent an estimated \$2500 - \$3500 per homeowner annually (beyond property taxes) in the area surrounding their seasonal home (Fedgazette 1997).

Seasonal home development brings with it a variety of impacts. Rural forest land is being consumed to make way for new seasonal homes (Vasievich 1999). Seasonal

homes built on a lakefront are contributing to problems including erosion, destruction of shoreline vegetation, and loss of wildlife habitat (Gartner 1987). Seasonal homes in the form of large recreational subdivisions cause strains on an area's water supply and the creation of extended road networks (Stroud 1983).

Seasonal Homes' Link To Recreation and Tourism

As Stewart (2000) notes, seasonal home ownership is a step between tourism and migration. To utilize seasonal homes, people are migrating from their permanent residence to a secondary one for a certain amount of time. This could be for a day trip, a weekend, a longer vacation, or for entire seasons (i.e. spending the winter in a warmer climate or the summer in a cooler one). Travel is involved with this temporary migration, which creates a link to tourism. As Chadwick (1994) notes, travel and tourism are words to describe "1) the movement of people, 2) a sector of the economy of an industry, and 3) a broad system of interacting relationships of people, their needs to travel outside their communities, and services that attempt to respond to these needs by supplying products."

This broad concept is broken down further with several definitions of what really constitutes "tourism." The World Tourism Organization (WTO) defines tourism as "the activities of a person travelling outside his or her usual environment for less than a specified period of time and whose main purpose... excludes only migration for temporary work" (Chadwick 1994). The National Tourism Resources Review Commission defines a tourist as a person who travels away from home at least 50 miles (referenced in Chadwick 1994). Similarly, the United States Travel Data Center uses 100

miles as the cut-off for travel and requires at least one night's stay in a destination to be considered a tourist (Chadwick 1994).

While these definitions have different demands on factors such as the distance traveled and the time spent at a destination, they all agree that tourism involves a movement (or migration) of people outside their usual territory for short-term (or non-permanent) travel. Thus, as seasonal homes constitute a temporary migration away from one's permanent residence (generally for purposes of recreation), seasonal home owners can be classified as tourists. Issues of tourism location and development are related to those of seasonal homes.

Tourism Development and Land Use

Recreational tourism-based activities can be seen as a specific type of land use. However, the use of land for recreation or tourism purposes can sometimes be in conflict with other needs or requirements for the land. McKercher (1992) describes the competition for land use between logging companies and wilderness outfitting industries. He notes that unlike harvesting operations that may only have a short term impact on land use, "once a wilderness area is accessed, it is accessed forever" (McKercher 1992). Roehl and Fesenmaier (1987) also point to land use conflicts occurring between areas of designated wilderness and tourist-based land uses. Construction of structures requires a permanency of land utilization (McKay 1987) that is often irreversible. Davis (1976) notes that land uses such as urbanization, draining of coastal areas or inland lakes, and mining and extractive practices carry an irreversible consequence for the environment.

Butler (1980) notes that recreation and tourism areas tend to follow a life-cycle which begins with initial interest and usually ends with the stagnation and decline of the area. Tourist destinations advance through this life-cycle as time progresses. Thus, initially, there tends to be an increase in the usage of land for recreational tourism purposes over time. At a certain point of development, some areas stagnate to the point of decline due to the overdevelopment of the area (Wheatcroft 1991). Many areas face the degradation of the resource that was the original initial attraction. In some ways, this eventual decline can be seen as “self destructive” (Holder 1988): as tourists came to the area the growing development in response to tourism brought the area to the “decline” stage of the life-cycle. Damage to these resources can threaten the economic viability of tourism in an area (Butler 1991). However, the stagnation of an area can often lead to redoubled efforts to try and attract new tourists (Cooper and Jackson 1989) by reinventing the area as a new concept.

Modeling Recreation and Tourism Issues

Frameworks have been constructed to model the changes that areas undergo in response to tourism development, but the spatial land use component is not present. The Destination Life Cycle Model (Butler 1980) uses the concept of product life cycle to reflect the changes that a tourist destination goes through from its inception to the ‘end’ of its ‘lifespan.’ The Destination Life Cycle is configured as a plot with time on the x-axis and number of tourists on the y-axis, with the plotted line reflecting the many stages a destination passes through: from exploration through consolidation and finally to eventual decline or rejuvenation.

Similarly, Miossec's (1976, referenced in Pearce (1989)) model of development reflects the changes to a destination, but focuses on the result of tourism development. By taking into account factors such as the number of facilities, their transportation infrastructure, the behavior of tourists, and the attitudes of decision makers, Miossec was able to provide a model of the changes that a destination will go through over a series of stages. Thurot's (1973, referenced in Pearce (1989)) process is similar; it typifies the phases a destination undergoes from the discovery of the area through its saturation. Gormsen (1981, referenced in Pearce (1989)) developed a representation of the increase in the number and type of accommodations in international seaside areas. Plog (1973) conceived a typology of tourists ranging from allocentrics (explorer types) on one end of the continuum to psychocentrics (stereotypical mass tourists) on the other end. In this way, the growth and development of an area is typified by the personality type of tourists it attracts. Also, the Community Options model (Talhelm et al. 1991) uses the STELLA modeling software to predict changes to a community based on possible recreation or tourism-related developments over a set time span (30 years). However, the Community Options model (and STELLA itself) are non-spatial in nature. None of these models deal specifically with spatial relationships in the areas under development pressure.

Beyond tourism-specific models of development, several other applications of statistical techniques to travel-based models exist. Wilkinson (1973) lists several distance-based economic and transportation based models (including time series projection and gravity models) and their relevance to tourism modeling. Similarly, Wolfe (1972) describes a variation on the gravity model called the inertia model, but notes it is only useful within a narrow range of distances. The gravity model was also applied to

describing and predicting intraurban recreational travel (McAllister 1977). Modeling of the spatial nature of recreation and tourism also comes out of applications of discrete choice modeling. Haider and Ewing (1991) apply discrete choice modeling techniques to simulate the choices made by tourists among several Caribbean islands. Stopher and Ergun (1982) used two-way segmentation to analyze the behavior of persons participating in urban recreation. Stynes and Peterson (1984) provide a review of logistic modeling techniques and possible applications to modeling choices of recreation. There are many more examples beyond these, but this paragraph simply provides a broad look at some representative publications.

Geographic Information Science

Geographic Information Systems have various definitions (Coppock and Rhind 1991, Johnson et al. 1987, Laurini and Thompson 1992, Maguire 1991), yet there are several common themes among them. Geographic Information Systems are defined as the hardware, software, and related materials that deal with issues of spatial or geographic analysis and spatial data handling. Marble and Sandhu (1994) state that GIS deals explicitly with spatial information. Issues such as spatial data capture, manipulation, overlay, visualization, and analysis are all part of GIS.

While the system itself, such as the Arc/Info software package, can be seen as a tool, and Goodchild (1990) calls the system a “toolbox” from which to solve spatial problems, the distinction must be made between the Geographic Information System and the Geographic Information Science that underlies it. A tool relies upon some sort of knowledge, science, or theoretical underpinnings to function. By way of analogy,

mathematics would be the science, while a calculator would be the tool used to aid in the implementation of that science.

Goodchild (1992) discusses the main issues or components of a perceived Geographic Information Science as follows:

1. Data collection and measurement
2. Data capture
3. Spatial statistics
4. Data modeling and theories of spatial data
5. Data structures, algorithms, processes
6. Display
7. Spatial analysis
8. Management and ethics

In all cases, these subjects are studied from a spatial perspective. Issues such as spatial autocorrelation (e.g. Griffith 1987), spatial visualization and rendering (e.g. Monmonier 1993, Orland 1994), and algorithm design (e.g. Openshaw and Clarke 1996) would all be components of Geographic Information Science. Spatial models make frequent use of Geographic Information Systems (GIS) to compile, analyze, and display the modeling results (Fischer 1996). Likewise, the theories and design behind the development of the system (Marble and Sandhu 1994) would be part of this as well.

GIS and Recreational Tourism

GIS is not commonly applied to tourism based research. When it is, it is used as a planning tool (Wilks et al. 1993), a database management system (Sheldon 1993, Vlitas-

Rowe 1992), or simple mapping utility (Filippakopoulou and Nakos 1995, May and Schmidley 1999, Phillips and Tubridy 1994). Other GIS-like technological approaches to tourism include expert systems and decision support (Crouch 1991, Gamble 1988) or management and planning applications (Poon 1988).

GIS, Seasonal Homes, and Real Estate

The literature is silent on the subject of the use of GIS to study the processes of seasonal home development. One link between GIS and this field comes in the form of GIS' impact on real estate markets. The use of GIS in site selection strategies (Daniel 1994) has long been the primary linkage between GIS and real estate. Also the Multiple Service Listing (MLS) tool used by all realtors to identify and analyze available properties (Longley and Clarke 1995) has begun to include mapping capabilities. GIS is also used with the real estate market in matters of property tax assessment and management of publicly owned real estate assets (Castle 1994).

Conclusion

This chapter provides an overview of the pertinent literature surrounding the three major divisions of this project. The foundations of geographic theory that this project draws from were examined, in terms of location analysis and the relevancy to land use. A brief description of the intersections of the literature of tourism development, landscape ecology, and GIS was also examined. It can be seen that while these three areas have their commonalities, they rarely (if ever) significantly overlap with one another. While GIS is an ideal tool for examining the spatial patterns of both landscape

ecology and tourism development, it is not often used for this purpose. The following chapter will take a closer look at the three states making up the study area of this project and their relation to seasonal homes.

Chapter 3: Study Area

The Upper Great Lakes States and Seasonal Homes

The Upper Great Lakes States region consists of: Michigan, Wisconsin, and Minnesota. The Great Lakes form a control for climate in the area, causing increased intensity of storms in and around the lakes in the winter, but decreased intensity of weather in the spring and summer (Albert 1995). The Upper Great Lakes States feature a diverse physical landscape with thousands of lakes and depressions across the landscape (Albert 1995).

The northern reaches of the Great Lakes region of the United States have seen tourism become their principal source of income (Chubb 1989). Declining agriculture and sawmilling, combined with manufacturing plant closures have led areas to shift their economies to tourism-based ones. Hart (1984) notes that the 3-state region gained over 935,000 people between the 1970 and 1980 censuses; the fast growth areas targeted the northern portion of Michigan's Lower Peninsula, northern Minnesota, and north-central Wisconsin.

In the Upper Great Lakes states of Wisconsin, Minnesota, and Michigan, approximately \$9 billion was spent on tourism-based activities in 1995, generating \$3.2 billion in income and supporting 214,000 jobs (Stynes 1996). The United States Travel Data Center (1998) estimates that in 1997, a total of \$10.1 billion was spent in total travel expenditures in the state of Michigan, generating over 157,000 jobs and nearly \$3 billion in wages and salaries, consistent with Stynes' (1996) research.

The Upper Great Lakes Region features an overwhelming number of inland lakes with opportunities for waterfront property development. Michigan contains more than 11,000 lakes, Wisconsin more than 15,000, and Minnesota over 10,000 inland lakes. The region contains an abundance of natural and public land as well. Much of Michigan's scenic areas are preserved from development as state or federal lands. There are over 224,000 acres of national parks and lakeshores in the state, over 111,000 acres of national wildlife refuges, and over 2.7 million acres of national forests (Spotts 1991). Another 260,731 acres are set aside for state parks, 28,925 acres for state boating and fishing, nearly 300,000 acres for state game and wildlife areas, and over 3.8 million acres for state forests. Wisconsin boasts over 59 state parks or forests (WDNR 2000) totaling 60,570 acres of state parks and 471,329 acres of state forests (Wisconsin.com 2000). There are 70 state parks in Minnesota, encompassing more than 240,000 acres of land (MDNR 2000).

With these types of natural features and the amount of outdoor recreation they represent, it is not surprising that seasonal home development has been steadily increasing in the Upper Great Lakes States over the last several decades. In 1940, seasonal homes were originally concentrated near major cities such as Detroit and Milwaukee (Hart 1984). However, over time, rural areas in the Lake States began to steadily gain more seasonal homes, while those areas around major cities were overshadowed by the growth of more permanent residences (Hart 1984).

Figure 3.1 shows the number of seasonal homes per county in the three states for the years 1970, 1980, and 1990. The northern parts of Minnesota and Wisconsin (particularly areas in Adams, Vilas, and Oneida counties) have seen steady growth along

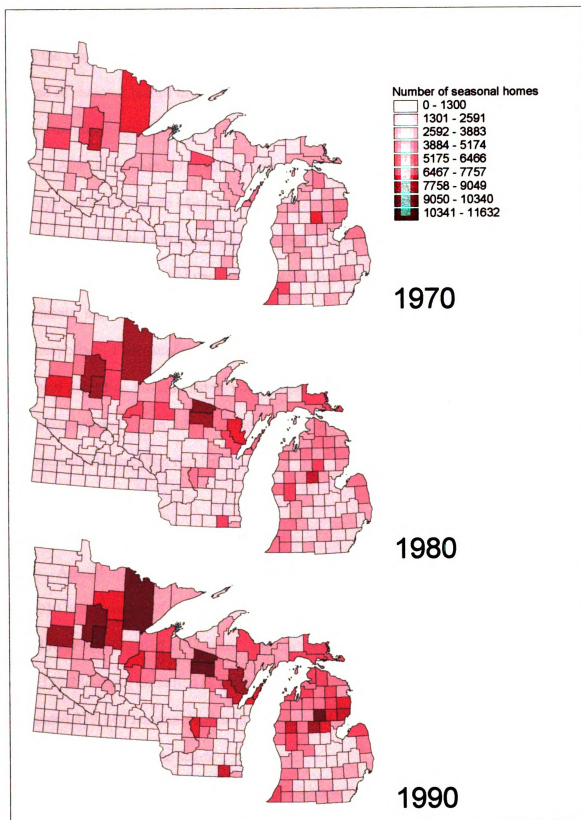


Figure 3.1: Number of seasonal homes in the Upper Great Lakes, 1970 - 1990

with the Upper Peninsula and northern Lower Peninsula of Michigan. This northward spread of seasonal homes can be partially attributed to lower land prices and the distance away from the larger cities in the southern portions of the states.

Also of note are the counties in southeast Wisconsin (particularly Walworth county) and southwestern Michigan, which are steadily growing over time. The proximity to the huge urban market of Chicago in the immediate vicinity is certainly playing a part in the level of vacation homes in those counties (Preissing et al. 1996). In Minnesota, the northern-most counties are seeing individuals from the Twin Cities of Minneapolis-St Paul purchasing seasonal homes there (Preissing et al. 1996). This seasonal migration has also led to a population redistribution in the Lake States (McDonough and Parker 1995), as more people are making the move to second homes and recreational vacation residences.

Spatial Distribution of Seasonal Homes

According to 1990 Census records, Michigan had 3,847,926 housing units with 224,030 being classified as seasonal. Minnesota had 1,848,445 housing units with 104,838 being classified as seasonal. Finally, Wisconsin had 2,055,774 housing units with 150,280 being classified as seasonal. Thus, over the three state area, there were an estimated 479,148 seasonal homes in 1990. These account for more than 6% of the total housing stock in the three states.

Figure 3.2 shows the proportions of the total housing stock that is seasonal of the Upper Great Lakes Region by Minor Civil Division (MCD) in 1990. The highest

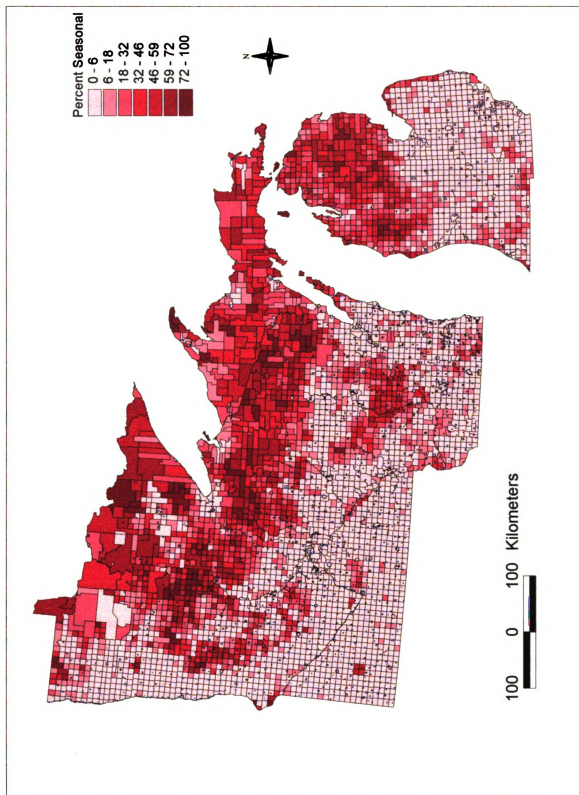


Figure 3.2: Percentage of the total housing stock that is seasonal, 1990

concentrations are at a distance away from the metropolitan urban centers of the states such as Detroit, Lansing, Milwaukee, and Minneapolis-St. Paul. A closer examination of the data shows that while these areas may be low in percentages of seasonal homes, the metro areas contain relatively high numbers of seasonal homes. Table 3.1 shows the number of seasonal homes in each of the MCDs located at the heart of the metro area.

MCD Name	Number of Seasonal Homes
Detroit City	174
Lansing City	137
Minneapolis City	287
Milwaukee City	260
St. Paul City	216

Table 3.1: Metro MCDs and their number of seasonal homes, 1990

However, the seasonal homes rank as an insignificant percentage of the total housing stock for each of these MCDs (the data lists them as being 0%). Thus, while seasonal homes may exist within highly metro areas, they are overshadowed by the number of permanent residences.

Examination of Figure 3.2 shows clusters of seasonal homes in areas away from major cities and closer to the lakeshore of the Great Lakes. High proportions of seasonal homes can also be found in areas within public land boundaries or near waterfront property. Analysis of this distribution shows certain patterns associated with higher- and lower- proportion clusters on the landscape. Table 3.2 shows the average proportion of seasonal homes for certain groupings of MCDs. For areas containing waterfrontage of major water bodies, a minimum of 1 ha of waterfront property was assumed. For public lands, if the centroid of the MCD was contained within the public land boundaries, it was considered part of the analysis. “Large cities” are defined as those places with a population of more than 100,000 persons.

Type of MCD	Average % Seasonal
Contains waterfront	18%
Contains Great Lakes lakeshore	24%
Distance of 25 miles from GL lakeshore	15.9%
Distance of 25 miles from large cities	2%
Distance of 100 miles from large cities	8.1%
Distance of 200 miles from large cities	12.1%
Population > 5000 persons	1%
Population > 10000 persons	0%
Population < 1000 persons	16%
Within public land boundaries	37%

Table 3.2: Types of MCDs and the average % seasonal they contain

Higher proportions of seasonal homes can be found in MCDs containing more natural features, such as water and Great Lakes lakeshore, or within relatively close proximity to the Great Lakes shoreline. The higher percentages of seasonal homes found within public lands' boundaries can possibly be attributed to the nature of public lands in the Lake States as outlined in Chapter 2. Select areas of public lands are being sold off as private lands, and thus only limited amounts of space are available for building. This could account for the higher proportions of seasonal homes in these areas. Lower average proportions of seasonal homes are found in areas with higher populations, while higher average proportions are found in less populated MCDs.

Conclusion

This chapter provides an overview of the study area and a description of the distribution of seasonal homes in the region (to fulfill objective 1). The Lake States contain a large array of landscape features which lend themselves to outdoor recreation. Seasonal home distribution in the Upper Great Lakes states is largely situated in areas away from large cities and has been historically growing in more remote, northern

locales. The presence of natural features, water bodies, and the Great Lakes adds to a higher proportion of seasonal homes found in those MCDs.

Chapter 4: Methods

Introduction

This chapter will outline the methodology used in the project and provide a look at the overall modeling approach. The modeling framework will be described, along with the description of the variables and how they were constructed in the GIS. Lastly, the types of models used in the study will be specified.

Modeling Framework

The modeling framework and technique used in this research project will be based upon some of the tenets of the Land Transformation Model (LTM) (Pijanowski et al. 1995). The LTM is a GIS-based modeling approach, which is used to predict land use change. While land use change is not being performed in this study, some of the basic functions of the LTM are utilized. The basic six steps of the modeling framework that the LTM utilizes are as follows (Pijanowski et al. 2001a):

1. processing and coding of data
2. applying spatial transition rules
3. applying integration method (such as the neural network)
4. using a Principal Index Driver to forecast over the temporal period being modeled
5. examination of spatial error
6. dynamic forecasting of future time steps through the use of population, per capita use demands, and economic trends.

The framework used in this project utilizes some of these tenets, including the

processing of data, application of spatial transition rules, application of the neural network for prediction (of one time period rather than a transition between two time steps), and examination of the spatial error signal.

The LTM operates on spatial data. Predictor variables are generated in the GIS from a series of base layers, representing land uses (such as agriculture parcels or urban areas) or features on the landscape (such as roads, rivers, or lakeshores). The predictor variables are built out of four distinct spatial interactions: neighborhood functions, patch size, site specific characteristics, or distance from a feature.

In land use change modeling, the LTM learns about land use transitions through different integration methods, including the use of an artificial neural network (Pijanowski 2001a), the Stuttgart Neural Network Simulator (SNNS). An interface written in C code links the output from the GIS to the neural network. The LTM's neural network capabilities have been used to examine land use changes within Grand Traverse County and then predicting subsequent land use changes within the six counties (Benzie, Leelanau, Grand Traverse, Antrim, Kalkaska, and Charlevoix) making up the Grand Traverse Bay's watershed (Pijanowski et al. 2001a). The LTM has also been applied to forecasting urban developments in the Minneapolis-St Paul and Detroit metropolitan areas (Pijanowski et al. 2001b). In addition, the LTM has been utilized in a variety of circumstances for prediction of land use change: exploring the linkages between land use/cover and subsequent impacts to surface and groundwater quality in an area of the Great Lakes Basin (Boutt et al. 2001); dynamically modeling predictors of land use change in the Saginaw Bay Watershed (Pijanowski et al. 2000a); and examination of land use change and urban development in Kuala Lumpur (Pijanowski et al. 2000b).

Given computing limitations, these predictor variables can only be constructed on a three-state-wide level at no less than 100m cell size. All computations were performed on a Pentium III Dual Processor PC, with access to roughly 45 Gb of storage space. Arc/Info, ArcView, SNNS, and C programming tools were used as tools for design of the variables, models, and model interfaces.

Unit of analysis

The unit of analysis in this project is the Minor Civil Division (MCD) level of the US census. MCDs are primary divisions of counties established under state law. According to the US Census, MCDs are defined as “A type of governmental unit that is the primary legal subdivision of a county in 28 states, created to govern or administer an area rather than a specific population. The several types of MCDs are identified by a variety of terms, such as town, township, and district, and include both functioning and nonfunctioning governmental units. Many MCDs represent local, general-purpose governmental units, which makes them required areas for presentation of decennial census data” (Census 2000). MCDs are identified by a five digit FIPS (Federal Information Processing Standards) code. In this project, MCDs for the three states of Michigan, Wisconsin, and Minnesota will be utilized.

The initial MCD database was put together by the Great Lakes Ecological Assessment project and the US Forest Service. The dataset required some fixing of inaccuracies within it (such as digitizing errors or numerous sliver polygons). Consequently, a few small polygons were lost in the “clean up” process. The MCD

database used in this project consists of 6,112 unique polygons spread across the three states.

Dependent Variable

The dependent variable used in this study is the percentage of housing units per MCD that were seasonal in 1990 (shown in Figure 3.2). As MCDs vary in size, using the straight number of seasonal housing units as the dependent variable may provide skewed results. Predictions for a particular MCD may be higher simply due to its size. Likewise, using the number of seasonal homes divided by the MCD size (i.e. seasonal housing density) may also be inaccurate, due to the varying sizes of MCDs.

Using the percentage of housing units that are seasonal per MCD as the dependent variable reflects a better measure of the number of seasonal homes in a particular MCD. However, there are some caveats that need to be accepted with this variable. Firstly, this proportional value of seasonal homes will vary according to the number of permanent homes in the area. A low (or zero) value will usually indicate an area with few seasonal homes, little seasonal migration, and a large number of permanent residences (such as predominantly urban areas). A high value will usually indicate those areas which have few permanent residences, or those areas where seasonal migrants outweigh the year-round populace.

Spatial Autocorrelation

Spatial autocorrelation is a measure of the correlation of values of the same variable at different spatial locations (Bailey and Gatrell 1995). This measure will aid in

determining where seasonal home distribution is clumped together, and how far one observation of seasonal homes has on other observations. Spatial autocorrelation will also aid in testing to accept or reject the null hypothesis that there is no detectable spatial variance in the distributions of seasonal homes.

Spatial autocorrelation analysis can be measured through the usage of the Moran's I indice. The results of Moran's I range from +1 meaning strong positive spatial autocorrelation, to 0 meaning a random pattern to -1 indicating strong negative spatial autocorrelation. Thus, the results of this analysis would aid in rejecting the null hypothesis and explaining the patterns of seasonal home distribution.

Variogram analysis is another method used in measuring spatial dependence. A semivariogram is used to compare the similarity between pairs of points a given distance apart and then mathematically expresses the average rate of change over a distance (Bailey and Gatrell 1995). Variograms are used to compute the average spatial dependence of the data and are fit through computing three parameters: the range, the sill, and the nugget. The sill is the point of variance where the variogram flattens out. The range is the distance at which the sill is reached (i.e. the limit of spatial dependence). The nugget is the intercept value. A flat variogram (one that is "pure nugget") would have no spatial dependence.

Independent Variables

From the literature review in Chapter 2, several variables that aid in the prediction of the percentage of seasonal homes within the total housing stock can be identified. However, the choice of predictor variables is limited by several factors. Availability of

data plays a strong role in the choice of variables that are used. As noted before, origin-based data would reveal much about the socio-economic characteristics (i.e. age, race, income level) of who is purchasing or utilizing a second or seasonal home. As this is a landscape-based study, such origin data was unavailable. Likewise, land values play a strong role in the choice of seasonal home location. However, without origin-based data of when seasonal homes were purchased, this type of variable cannot be taken into account.

Land prices for many locations have changed dramatically over the last several decades. What could have been bought cheaply in the 1960's is now very expensive in the 1990's. Thus, using 1990 land valuation prices for an independent variable would not provide much of explanatory power of what the variable represents. While factors such as age, education, and income are key demographic estimates of evaluating the tourist usage of a site (Loomis and Walsh 1997), they are unfortunately inapplicable in this instance. Characteristics such as these reflect the qualities of the tourists or users of the second homes, yet this study is a land use-based one. Origin information is unavailable. Factors such as age, education, and income level can be compiled only for the destination MCD. These factors will more than likely have no direct correlation to the age, education, or income level of those arriving to utilize the second homes. Also, the scale of this project limited the types of factors that could be used. Factors such as locations of boat launches or marinas may help to explain the locations of high proportions of seasonal home distribution, however such data were unavailable.

Many of the independent variables used in this study are based on land uses or natural features on the landscape. As noted in chapter 2, in many instances it is the land

use and land cover features that provide the impetus for seasonal home development, along with the numerous outdoor recreation activities that can be derived from them. For instance, public lands not only contain much in the way of unspoiled nature, but are also the home to popular activities such as cross county skiing, hiking, boating, and hunting.

Development of Independent Variables

The independent variables were estimated for all MCDs within the three state region. Some variables (based on census data) will assign a single value to an MCD. However, variables that cover a wide spatial area (such as the distance and density functions) will have numerous values associated with them. In cases such as these, an average of the values within the MCD will be taken and a single value assigned to the MCD as a whole. For most variables, the mean value will be used so that all townships, regardless of size, may be evaluated equally.

Variables are created in the GIS (using ArcView) as follows:

- (1) put the MCD theme in a view containing the predictor variable theme;
- (2) Use the “Summarize By Zones” function – this computes statistics for each polygon in the MCD theme using the predictor variable grid. The FIPS code variable in the MCD theme is used to insure that zones are summarized by individual MCDs. A variety of statistics are created, including “mean” and “sum.” The “mean” option represents the average value of the predictor variables per MCD. In most cases, except where noted below, the mean value for each MCD will be used to give an average estimate of the variable for each MCD.

- (3) Join the table produced by the Summarize Zones function to the table of original MCD theme. This will result in the statistical value table being joined to the spatial data of the MCD theme.
- (4) Convert the theme to a grid, using a statistical attribute and the extent of the predictor variable grid (and 100 m cell size). This will create a grid of the requested value (usually mean) of the predictor variable per MCD. Thus, each MCD has one value for each predictor variable.

Each of the tables created per variable can be exported in dbase format, which is useable in several statistical based products for further analysis in this project.

Each of the independent variables can be constructed in the GIS for the models as follows:

1. All landscape density functions: The particular classification of the data may be extracted into a separate layer, and a neighborhood operation (with a window size of 1.1 km and a “focal sum” option chosen) can compute a measure of density for that land use (this follows similar variable construction in Pijanowski et al. 2001a and Pijanowski et al. 2001b).
2. All distance functions: The distance option of the LTM can be used on the necessary layer. Euclidian distance is computed from the base layer to all other cells.
3. All amount functions: Land use data is available for each of the three states. By overlaying the MCD layer on the land use layer, a count of the cells in the area (which can be computed to acreage – one grid cell of 100 meters square

is equivalent to 1 ha) for each MCD may be obtained. In these cases, the “sum” option for the summarize zones will be used.

In Chapter 2, several factors contributing to the locations of seasonal homes were described. The best available sources of data were used to create the independent variables in this project. With data from commercial or government sources, or data processed by others and used in previous projects, some inaccuracies may be found in the datasets, especially when working at such a broad three-state scale. Maps of each independent variable can be found in Appendix A. Specific independent variables were created as follows:

1. Water related

Three water-related variables were created. The source of the water data for the three states was derived from USGS 1:100,000 DLGs of hydrographic sources developed in conjunction with other projects, which included a polygon coverage of major water bodies in the three states, including lakes and rivers. The polygon file was also built as a line file to represent the water frontage rather than the water area. This new “water front” coverage was converted to a grid and the “sum” option as described above was used to find a summation of the amount of waterfront per MCD. The variables created are as follows:

- distance from major water bodies: this shows the distance from water bodies to all other areas.

- density of major water bodies: this reflects the abundance and clustering of water bodies and thus implies more area for recreational opportunities in the area.
- amount of waterfront per MCD: this variable provides a measure of the amount of water frontage, and thus the amount of area potentially available to build on.

2. Great Lakes related

The most accurate Great Lakes shoreline boundary file was derived by taking the polygon file for the MCDs and editing out all arcs not on the lakeshore. Two variables were created from this data:

- Distance from Great Lakes Lakeshore: this reflects the distance from the lakeshore to all other areas
- Amount of Great Lakes Lakeshore: reflects the amount of lakeshore available per MCD to potentially build on. Like the “amount of water front” variable described above, this is a summation of the amount of Great Lakes lakeshore each MCD contains.

3. Urban areas and population

The US Census Bureau maintains a “place names” list for places throughout the United States. This contains the latitude and longitude coordinates, names, and 1990 populations for towns and cities throughout the country. This provided the best available measure of the locations of towns and cities, as well as their populations. Places within Michigan, Minnesota, and Wisconsin were extracted, along with the databases for the neighboring states of Iowa, Illinois, Indiana, and Ohio. These extra states were chosen as

there may be significant markets from these neighboring cities (especially in Chicago) that may be contributing to recreational use, or seasonal home use or purchase in Michigan, Wisconsin, or Minnesota.

Population density figures were available from US Census records on a per MCD basis, measured as population per square mile. The best available source for hospitals in the region came from ESRI's Street Map product, which stored hospital locations as point files (hospitals in Michigan, Wisconsin, and Minnesota were used along with hospitals in Indiana, Iowa, Illinois, and Ohio). The following variables were created from this data:

- Stratified distance from cities: This variable is an attempt to simulate the effect of proximity to amenities in more rural settings, as well as the distance from larger metropolitan areas. The place names database was used as the source of the locations and population counts. Cities were broken up into the following five categories:
 - (a) Distance from places of less than 500 persons
 - (b) Distance from places between 500 and 1000 persons
 - (c) Distance from places between 1001 and 10000 persons
 - (d) Distance from places between 10001 and 100000 persons
 - (e) Distance from places with more than 100000 persons
- Distance from hospitals: This variable reflects access and proximity to major health care facilities.
- Distance from designated "tourist attractions." This is intended to be a simple list (not a detailed all-inclusive one) that is representative of some major attractions within the three state area. These places were selected from a variety of sites including Spotts (1991), Wisconsin.com (2000), Office of Minnesota Tourism (2000),

and Delorme (2000a, 2000b, and 2000c). The locations of tourist sites were selected from the “places” database and distance functions were computed from them. The tourist sites selected for use in this project are as follows:

- MI: Birch Run Village, Bridgman City, Dearborn City (Henry Ford Museum / Greenfield Village), Detroit, Flint, Frankenmuth, Holland, Mackinaw City, Mackinac Island, Munising, Petoskey, St. Ignace, Saugatuck, Sault St. Marie, Traverse City
 - WI: Baraboo, Green Bay, Milwaukee, Sturgeon Bay, Wisconsin Dells
 - MN: Bloomington (the Mall of America), Duluth, Shakopee, Minneapolis, St. Paul, Taylors Falls
- Population Density: reflecting the number of persons per square mile.

4. *Public lands*

Minnesota public lands data came from the Minnesota branch of RESAC (Regional Earth Science Applications Center) in the form of GIS layers digitized from 7.5 minute DRGs. Wisconsin public lands information came from the Wisconsin DNR (Department of Natural Resources) fashioned from a variety of sources (including 1:24000 DRGs, 1:100,000 DRGs, and 1:15940 data sources). Michigan public lands information came from the Michigan DNR, fashioned from the Michigan Public Lands survey system. All of these public lands files represent the most accurate public lands information available for this project. Included in these data are state forests, national forests, state parks, national parks, and protected wildlife areas. Two variables were created from this data:

- Distance from public lands: reflecting access and proximity to public lands
- Density of public lands: this variable reflects the abundance and clustering of public lands.

5. *Land Use / Land Cover Related*

1990 NLCD (National Land Cover Database) data were available in a rectified format for the three state area. Derived from early-to-mid-1990's Landsat TM satellite imagery, the NLCD is a 21-class land cover classification scheme applied consistently over the United States. The spatial resolution of the data is 30m. Land covers were extracted and used as the source for the land use and land cover related variables used in this project. Land cover types were resampled to 100m resolution to keep them in line with other variables used in this project. This allowed the best available access to data about the locations of land use / land cover features such as agriculture, forest, water, and wetlands. This was the source of the agriculture-related variable used in this project, as well as two of the aesthetic quality related ones. The following variables were derived from this data:

- Density of agriculture: This variable reflects areas high in agricultural land use (using the land cover categories of pasture / hay, row crops, and small grains).
- Density of natural areas: Natural areas in this case are defined as open water, all forested land cover types, and woody and non-woody wetlands from the NLCD database. This variable simulates the natural (aesthetic) quality of the area.

- Amount of forests: The amount of forests (all forested types and woody wetlands) in an area is an important variable for aesthetic quality. A summation of all forested cells within an MCD is used in this variable.

6. *Terrain related*

A 90m DEM (Digital Elevation Model) of the three-state area was obtained from USGS sources in conjunction with other projects. This was resampled to 100m to be consistent with the other variables used in this project, and used to construct a measure of landscape variability. A neighborhood operation similar to the density functions was performed using a 1.1km window, but using the standard deviation option, rather than a summation on (a similar method was used by Pijanowski et al. 2001a). The following variable was created from this:

- Landscape variability: this simulates an aesthetic measure of the “hilliness” of the area’s topography.

7. *Accessibility*

The best available measures of accessibility came from data derived from ESRI’s Street Map product (which is in turn derived from US Census TIGER files). This provided a road network for the three state area of Michigan, Wisconsin, and Minnesota, along with the northern portions of Illinois, Indiana, Iowa, and Ohio. The road network file classifies each road segment according to its CFCC (Census Feature Class Code). Thus, the road layers used in this project are those selected according to their CFCC classification. The following variables were created from this data:

- Distance from local roads: This reflects access to places via local roads, rather than more mass-movement road systems.
- Distance from US highways: This reflects access to places from major US highways.
- Distance from interstates: This reflects access to places from major US interstates.
- Distance from vehicular trails: As no concise database of the locations of snowmobile trails for the three states is available, this measure is used as a surrogate for snowmobile or off-road vehicular trails.

Correlation of the Dependent and Independent Variables

Bivariate correlations between the dependent and the independent variables were run to examine the strength of their linear relationships. The bivariate correlation results are shown in Table 4.1.

Principal Components Analysis

With 23 variables and 6,112 observations per variable, there is likely to be a high degree of correlation between some of the variables. Some of them will be reflective of similar properties and have high correlation.

Principal Components Analysis (PCA) was used (with the SPSS software package) to reduce the number of variables to a set of new independent variables (Taylor 1977). This will “reduce a large number of variables to a smaller number of factors for modeling purposes, where the large number of variables precludes modeling all the measures individually” (Garson 1999). PCA will combine two or more correlated variables into a single component. The concept of PCA involves a variance maximizing

VARIABLE	Correlation
DIST MORE 100K	.266
DIST TOURISM	.005
DIST 10K TO 100K	.289
DIST HOSPITALS	.226
DIST GREAT LAKES	-.262
DIST INTERSTATES	.257
DIST 1K TO 10K	.450
DENS NATURAL	.702
DENS AG	-.518
DIST PUB LANDS	-.321
DENS PUB LANDS	.493
DIST VEHICULAR	-.361
DIST LESS 500	.289
DIST 500 TO 1K	.294
AMOUNT FORESTS	.426
AMOUNT GL SHORE	.190
DIST LOCAL RDS	.034
POP DENSITY	-.251
DIST HIGHWAYS	.364
DENS WATER BODIES	.273
DIST WATER BODIES	-.222
AMT WATERFRONT	.314
VARIABILITY	.015

Table 4.1: Bivariate correlations

(“varimax”) rotation of the original variable space (Taylor 1977). The rotation maximizes the variance captured by the new output variable (a component) while minimizing the variance around this new variable (Statsoft 2000). These new components will be orthogonal in n-dimensional space and uncorrelated with each other (Taylor 1977). Each component produces a set of factor scores, which can be used as the new set of independent variables in the modeling procedures.

Logistic Regression and the Logit Model

Logistic regression is based on the known values of the dependent variables being represented by the presence or absence of an occurrence, and applies maximum likelihood estimation to estimating the probability of an occurrence. Logistic regression does not assume a linear relationship between the dependent and independent variables (Garson 1999). However, logistic regression usually requires the dependent variable to be a binary response function (i.e. the dependent variable is either a value of zero or one). This is generally used to indicate the presence or absence of a phenomenon. However, as the dependent variable in this project is a proportional value ranging between 0 and 1, binary logistic regression cannot be performed.

The solution to this is to calibrate a logit regression model. The output of the logit regression model will be bounded between 0 and 1 regardless of the values of the independent variables (Statsoft 2000). The logit model takes the following form:

$$y = \exp(b_0 + b_1 \cdot x_1 + \dots + b_n \cdot x_n) / \{1 + \exp(b_0 + b_1 \cdot x_1 + \dots + b_n \cdot x_n)\}$$

The coefficient values are obtained through maximum likelihood estimation using the quasi-Newton algorithm (Statsoft 2000). Maximum likelihood is a technique to

maximize the likelihood that the dependent variable values occur within the sample (Statsoft 2000). The greater the value for the likelihood, the better the fit of the model.

The maximum likelihood estimation for logit models takes the form of:

$$\log(L) = \sum_{i=1}^n [y_i * \log(p_i) + (1-y_i) * \log(1-p_i)]$$

where L is the loss function to be minimized. Seeds for the initial beta coefficient values are necessary. They were obtained by running a linear regression on the variables (Gupta 1999).

With these initial coefficient values in hand, the Solver program in Microsoft Excel was used to provide the best fit to the model via maximum likelihood estimation. Once the best fit was found to minimize the loss function, Solver stored the new values of the coefficients. These new coefficients were then used in the logit model to find an estimate for the dependent variable.

The Artificial Neural Network model

Artificial Neural Networks are tools designed to simulate the learning processes of the interconnected neurons of the human brain. Neural networks are able to sort patterns and learn from trial and error about data. One of the first neural networks was the perceptron (Rosenblatt 1958), a simple linear machine consisting of a single node capable of receiving weighted inputs and thresholding them according to a defined rule. This type of neural network can classify linear data.

Current neural network technology processes data in a non-linear fashion. The MLP (multi-layer perceptron) neural network algorithm is non-linear. The MLP neural net consists of three layers: inputs, hidden, and output and thus can separate data

that are non-linear in nature (Skapura 1996). Figure 4.1 shows a diagram of the design of

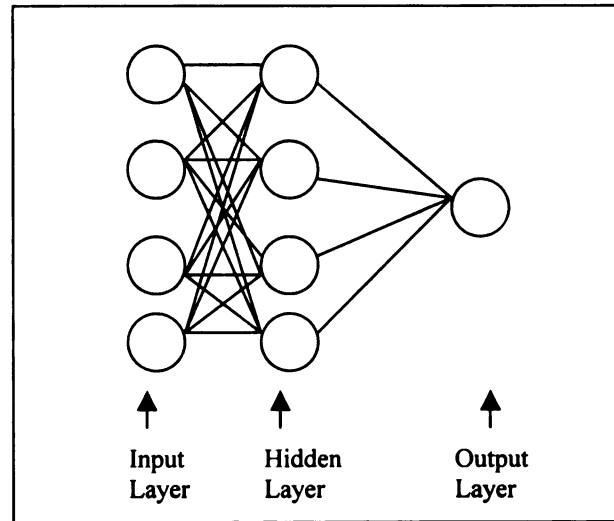


Figure 4.1: Basic Architecture of a MLP Neural Network

a simple neural network. Neural networks feed information forward from the inputs through the hidden layer and finally to the output layer. Weights are associated with the connections between nodes, which modify the signal propagating through the network. Weighted summations are computed at each node. Each hidden node computes a sum based on the input from the previous layer. This sum is then altered, usually using a logistic sigmoid function (Bishop 1995) to reduce it from the range (+infinity, -infinity) to values between (0,1). The output from this node is computed to be the “activation function” for that node to “fire” and send its data forward through the network. This process continues until the output node is reached. In this fashion, the data are learned.

The weights in a neural network are determined by a training algorithm, the most prevalent of which is the backpropagation algorithm (Skapura 1996). Initial weights along the connections are randomly determined. The given value for a particular observation is compared to the estimated value. The difference between observed and

expected for all observations is calculated as a mean squared error. The total error value is then distributed backwards through the system, resetting the weights to new values. The signal is then fed forward through the network again. This process of feeding input forward and then propagating the error backwards through the system continues for several iterations until some specified threshold value for the error is reached. At this point, the neural network is considered sufficiently “trained” on the data.

The use of neural networks as models comes when the training data is applied to another data set to test upon. Once the network has learned about the patterns in one set of data, it will apply what it has learned to another set of data. These two phases, the “training” and “testing” stages are what make up the functionality of using neural networks in modeling.

In this project, the input layer of the neural network represents the independent variable inputs, while the output represents the values of the dependent variable (see Figure 4.1 for a graphic of the network). The Stuttgart Neural Network Simulator (SNNS) software package is used to create and implement the network. The backpropagation algorithm is used for the training procedure and the neural network takes the form of a set of input layers (one for each independent variable), a set of hidden layers (where the training occurs), and 1 output layer (representing the estimate of the dependent variable). A stratified random subset of the data were used for the training, and the remaining data used for the testing.

The structure of the neural network used in this project is modeled after the architecture used in Pijanowski et al. (2001a) and Pijanowski et al. (2001b). These other projects utilize the same neural network software package (SNNS), and software

interface, large data sets with a neural network trained at 500 cycles, a structure of one hidden node for each input node, and the use of the logistic sigmoid function. Simple investigation into reducing the number of hidden nodes showed that the mean squared error result did not significantly change. The MSE began oscillating at a point just after 500 cycles. For this project, the structure outlined by Pijanowski et al. (2001a) and Pijanowski et al. (2001b) was used, while the number of training cycles was selected as 1000.

Conclusion

This chapter has presented the general format for the procedures used in modeling the relationship between the predictor variables and the dependent variable. Spatial autocorrelation techniques will be used to examine the dependent variable to draw conclusions about the relative clustering and pattern of seasonal homes on the landscape. Principal Components Analysis will be used to reduce the number of independent variables to a smaller set of new independent variables to be used in the modeling processes. Two nonlinear models, a logit regression model and an artificial neural network model, will be used to evaluate the relationship between the dependent and independent variables. Chapter 5 will present the results of these models, as well as methods for model evaluation.

Chapter 5: Model Results and Analysis

Introduction

This chapter examines the results from the modeling approaches described in Chapter 4. The results of the models are discussed, along with metrics for analyzing those results. A comparison of the modeling approaches is given, as well as a means for identifying those variables that best explain the relationship between the independent and dependent variables. Images in this dissertation are presented in color.

Spatial Autocorrelation

Spatial autocorrelation analysis was performed using the S-PLUS program, which computes a Moran's I value on the data. The centroids of the MCD polygons were computed and used as point data in S-PLUS. Thus, each point contained the value of the dependent variable. Spatial autocorrelation was calculated using the four nearest neighbors for each point. The Moran's I was 0.625, indicating highly positive spatial autocorrelation. This can be interpreted as MCDs with similar proportions of seasonal homes (whether high or low) are found clustered together in a smooth and continuous fashion. In this case, the null hypothesis of no detectable spatial variation can be rejected.

The semivariogram of the dependent variable was fit using functions in S-PLUS (shown in Figure 5.1). The best fit was found with a range = 105500, a sill = 365, and a nugget = 70. This indicates that 105.5 km is the average limit of spatial dependence and the "drop-off" point for the autocorrelation effect of a particular MCD.

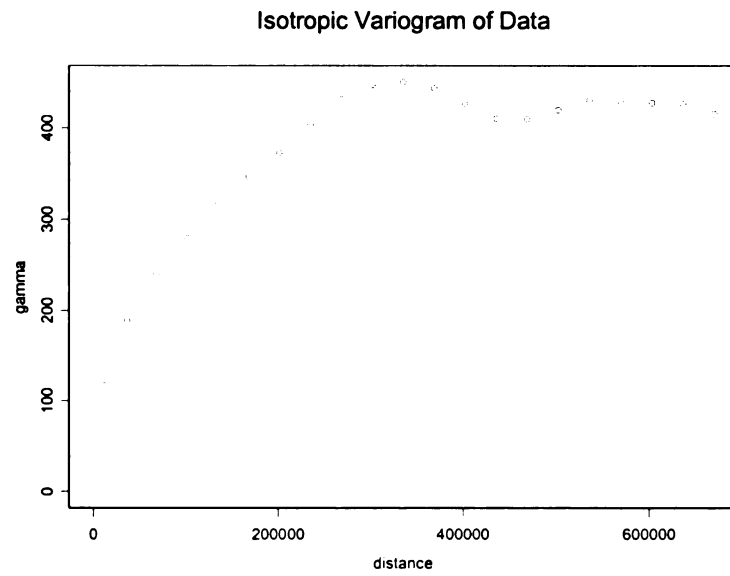


Figure 5.1: Variogram of the dependent variable (distance in meters)

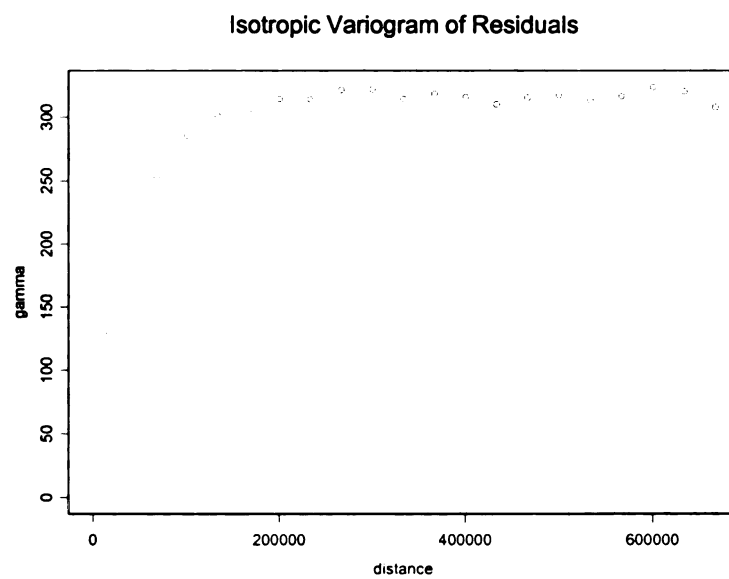


Figure 5.2: Variogram of the residuals of the dependent variable (in meters)

After removing the first order trend surface, a new semivariogram of the residuals (what cannot be explained by the trend surface itself) may be created (see Figure 5.2). The best fit of the residual variogram was found with range = 41.5, sill = 318, and nugget = 1. In this case, those factors that are unexplained (the residuals) have their “drop-off” point of spatial dependence at 41.5 km. The percentage of seasonal homes has a spatial dependence of over 105 km, but to find those factors accounting for the unexplained variation, more local trends (such as those within 41.5 km) must be examined.

PCA Results

The Principal Components Analysis produced 6 components with eigenvalues greater than 1 (see Table 5.1), explaining 63.1% of the variance in the independent variables. According to the Kaiser rule, components with eigenvalues of 1 should be kept and others discarded (Guttman 1954). The six components are orthogonal and uncorrelated with each other. Each component has a set of standardized factor scores (having a mean of 0 and a standard deviation of 1) for each observation that can be used as new independent variables in further analysis.

Component	Total	% of Variance	Cumulative %
1	5.218	22.686	22.686
2	3.812	16.572	39.258
3	1.658	7.210	46.468
4	1.495	6.499	52.967
5	1.288	5.601	58.568
6	1.053	4.579	63.147

Table 5.1: Initial Eigenvalues

The rotated factor loadings for each variable in the components are shown in Table 5.2. By examining the rotated factor loadings, each component can be assigned

VARIABLE	C1	C2	C3	C4	C5	C6
DIST MORE 100K	.899					
DIST TOURISM	.850					
DIST 10K TO 100K	.787					
DIST HOSPITALS	.685					
DIST GREAT LAKES	.541	.512	-.430			
DIST INTERSTATES	.526					
DIST 1K TO 10K	.510	-.410				
DENS NATURAL		-.824				
DENS AG		.708				
DIST PUB LANDS		.693				
DENS PUB LANDS		-.567				
DIST VEHICULAR	-.438	.511				
DIST LESS 500			.703			
DIST 500 TO 1K			.655			
AMOUNT FORESTS			.599			
AMOUNT GL SHORE			.545			
DIST LOCAL RDS				.796		
POP DENSITY				-.686		
DIST HIGHWAYS				.571		
DENS WATER BODIES					.761	
DIST WATER BODIES					-.625	
AMT WATERFRONT			.412		.619	
VARIABILITY						.914

Table 5.2: PCA Rotated Component Matrix

an interpretation. As Burley and Brown (1995) point out, the names of the new components are somewhat subjective, but are very useful in interpreting the results. The new variables are described as follows:

1. Distance from large cities: the highest positive factor loadings come from the variables of the distance from areas of more than 100,000 persons, distance from areas of between 10,000 and 100,000 persons, distance from hospitals, distance from tourism sites, and distance from interstates.
2. Absence of natural areas / presence of agriculture: the variable with the highest factor loading in this component is the density of natural lands. However, this component is loading highly negative. The next highest loading comes from density of agriculture, a high positive loading. This component also includes high positive loadings for distance from public lands, and distance from commercial tourism sites, as well as a high negative loading for density of public lands. This can be interpreted as “absence of natural lands” or areas low in natural lands and public lands, with a high density of agriculture.
3. Distance from small towns: Distances from towns of less than 500 persons and from places of between 500 and 1000 have the highest factor loadings, along with the amount of forest and the amount of Great Lakes lakeshore.
4. Distance from local roads / accessibility: This component is categorized by a high positive loading for distance from local roads, and a high negative

loading for population density, with a high positive loading for distance from highways.

5. Presence of water: The highest factor loading is for the density of water variable. Also loading high is the negative loading for distance from water.
6. Landscape Variability: This variable is categorized by a high loading in landscape variability.

Logit Model Results

As noted in Chapter 4, the seed values for the coefficients of the logit model were obtained using a linear regression of the variables. After applying maximum likelihood estimation to calibrate the logit model, the coefficients changed considerably. Table 5.3 shows these changes in coefficients.

A correlation between the observed and estimated results of the logit model was found to be 0.727. This value can be squared to calculate a “pseudo- R^2 ” value for the overall fit of the model. The pseudo- R^2 value for the logit model was 0.53.

Beta	Seed value	Solved value
b0	.126	-2.47
b1	0.068	0.29
b2	-.092	-1.05
b3	0.047	0.28
b4	0.043	0.26
b5	0.058	0.41
b6	-0.008	0.07

Table 5.3: Initial beta values and solved values

Model verification and validation was performed to test the robustness of the model. A random sample of 20% of the MCDs was chosen to calibrate the model, while

the remaining 80% was used to test the model. The model was fit by running a linear regression using the 20% sample to obtain the initial seed values for the beta coefficients. The model was then tested on the remaining 80% of the data, resulting in a correlation coefficient of 0.72, or a pseudo- R^2 of 0.53. These values are very similar to the overall fit of all the data, indicating a robust model.

Neural Network Model Results

As noted in Chapter 4, the neural net model needs data to train on and to test on. A stratified random sample of 80% of the MCDs was chosen and separated from the remaining 20%. The neural network was trained on the 20% sample and then tested on the remaining 80%. This gave a correlation value of 0.79 between the observed and estimated values, resulting in a pseudo- R^2 value of 0.63 for overall fit of the model.

The 20% sample size was chosen due to concerns of “overtraining” the data, or giving the neural network too much to work with. Another concern with a large training sample is that if the model is learning about the majority of the data, intuitively one would think that it should have a relatively decent fit to the smaller testing sample. For instance, a model presented with 99% of the data should be able to provide a good estimate of the remaining 1%. Therefore, a smaller sampling size was chosen to test the model.

Examination of Independent Variables

To fulfill the second objective of this dissertation, the independent variables contributing the most to the overall model must be found. The PCA culled the list of the

initial 23 variables down to a set of 6 new variables. However, to fulfill objective 2 as laid down in Chapter 1, the principal predictors of the proportion of seasonal home distribution must be identified.

Logit Model

In the logit model, the principal predictors of the dependent variable can be identified by examining the beta coefficients of the independent variables. These give a measure of not only the importance, but also the direction (negative or positive) that the variable takes. Table 5.3 shows the beta coefficients for the model. As can be seen, the highest values associated with a factor is that of the second variable, the one representing the absence of natural areas and presence of agriculture with a value of -1.105 . This indicates that the absence of natural areas has a very strong negative effect on the model, meaning that the presence of natural areas has a very strong positive effect on the model. Also, this indicates that the presence of agriculture land has a strong negative influence on the distribution of seasonal homes.

The variable with the next highest beta coefficient is the first one, representing the presence / density of water, with a value of 0.41 . Ranked next with a value of 0.29 is the variable representing the distance away from urban areas. These results indicate that these two variables have a strong positive effect on the overall fit of the model.

Neural Network Model

Unfortunately, neural networks do not produce coefficients to fit to structured equations, and thus one cannot directly infer from neural networks the explanatory power

of the model. All of the processing occurs within the “invisible” portion of the model, namely the hidden layers and their associated weights moving to and from the inputs, hiddens, and outputs. Unlike the logit model, the importance of each variable to the overall model cannot be extracted.

However, there is a way of examining the relative importance of each independent variable in the overall model. Sensitivity analysis allows for the model to be run multiple times with a single variable removed from the model. Thus, with 6 independent variables, there will be 6 new models created, each containing 5 variables. Each of these “reduced variable” models represents the model without the contribution of one of the variables. The changes in pseudo- R^2 values between the reduced variable models and the overall model simulate the relative contribution of each independent variable to the model. Table 5.4 shows the pseudo- R^2 values of each reduced variable model.

Variable removed	NN Model's Pseudo-R^2
1: Distance from cities	0.59
2: Absence of natural areas	0.47
3: Distance from small towns	0.557
4: Accessibility / Local roads	0.56
5: Presence of water	0.525
6: Landscape variability	0.63

Table 5.4: Reduced variable runs of neural network

As can be seen, the overall fit of the model drops to a pseudo- R^2 value of 0.47 when the second factor (absence of natural areas) is removed from the model. This indicates that the absence of natural areas variable contributes the most to the overall fit of the model. The second largest drop in model fit comes when the fifth factor (water) is

removed from the model, reducing the overall fit to 0.525. These variables correspond with the top two predictors found by the logit model.

The Principal Predictors

Both the logit model and the neural network model identify first the absence of natural areas (or presence of agriculture) as the principal predictor of seasonal home distribution and secondly the presence of water. The third through fifth ranked variables of distance from large cities, distance from rural towns, and distance from local roads / accessibility are close together and their ranking varies only slightly from each model. Finally, both approaches agree that the landscape variability factor contributes the least to the overall model.

Table 5.5 shows how each variable is ranked in its importance to the model. Note that while the “reduced variable” model gives some approximation of the contribution of that variable to the overall fit of the model, they cannot tell if the variable has a positive or negative effect, as the logit model coefficients can.

Logit model	Neural network model
1. Absence of natural areas	1. Absence of natural areas
2. Presence of water	2. Presence of water
3. Distance from large cities	3. Accessibility to local roads
4. Distance from small towns	4. Distance from small towns
5. Accessibility to local roads	5. Distance from large cities
6. Landscape variability	6. Landscape variability

Table 5.5: Comparison of model rankings of principal predictors

Spatial Distribution of Seasonal Homes

This project has provided a measure of the principal predictors of seasonal home distribution across the Upper Great Lakes states of Michigan, Wisconsin, and Minnesota.

From the results of the logit and neural network modeling approaches, the “absence of natural lands density / presence of agricultural density” variable had the strongest effect on the model. In the logit model, this was shown to have a high negative coefficient, indicating that absence of natural lands (forest, water, and wetlands) and the presence of agriculture has a strong negative effect on seasonal home distribution. This indicates that the presence of natural lands and the absence of agricultural lands has the strongest positive influence on the model. This fits in with the general theory and literature outlined in chapter 2, outlining that an impetus for utilizing a seasonal home is to remove oneself from an urban environment out to a more rural one. Natural areas of forest and water provide a wide range of recreation opportunities, either in maintained public lands areas or in other natural areas. Also, the natural areas present a more pleasing aesthetic quality than urban areas. Similarly, high density of agriculture has a negative effect on seasonal home location, similar to findings by Chubb and Chubb (1981), possibly indicating the land being used for more farming than seasonal homes.

The next principal predictor is presence of water bodies. The literature indicates that water plays a key role in the location of seasonal homes. Water bodies provide a wide range of outdoor recreation opportunities, including boating, fishing, and swimming. Also, the presence of water indicates a more aesthetically pleasing, or perhaps even a “soothing” effect from more urban settings.

The next three principal predictors of seasonal home location revolve around urban areas, cities, towns, and access to and from them. Their order of importance is ranked differently between the two models, although they are still ranked at a level of third through fifth, with coefficients in the same general range. The literature notes that

seasonal homes are purchased to remove oneself from an urban environment into a more rural setting. Thus, distance away from cities should be a key factor here. Similarly, the distance from smaller towns (reflecting distance from more rural locations) and a distance from local roads provide similar explanations. Smaller towns reflect more rural settings that may still provide basic amenities such as food, lodging, gasoline, or other necessary items. However, while they provide amenities, they are very different from the heavy urbanization of larger cities. The local roads / population variable reflects two types of access: residential roads in urban areas and more local backroads to rural destinations. This provides a measure of access from all types of housing areas. These factors reflect some of the ideas in the literature that seasonal homes may be found in less urbanized and more rural areas, but that still have easy road access to them.

The final factor ranked last in importance by both models is categorized by landscape variability. The changing character of the landscape has little role in explaining the distribution of seasonal homes in the area. A possible explanation for this may lay in the concept that it is not the abruptly changing landscape that contributes to the area's aesthetic quality, but rather what can be seen from certain vantage points. For instance, a high elevation may allow a wide view, but the quality of what can be seen may not be pleasing to the observer. The approximation of this variable may be a crude way of showing the nature of the landscape.

Model Comparison

Several methods were used to compare these two modeling approaches. The first of these are the pseudo- R^2 values (or the square of the correlation coefficient). These

provide a measure of the overall fit of the model. The logit model's pseudo- R^2 of 0.53 is less than the neural network's pseudo- R^2 of 0.63. Judging solely from this measure, neural networks give a better predictive power.

However, a pseudo- R^2 is not the final word in appraising model performance. For instance, neural networks are not explanatory models where the logit model is. If the only model used was the neural network model, the power to explain the strength and direction of variables would be very limited, and thus objective 2 could not be accurately met. This can be seen in examining the strongest variable in the model, absence of natural lands. From the logit model, the coefficients indicate the overall strength of the model. More importantly, they indicate that this variable has a strong negative effect on the model, which can thus be interpreted that the presence of natural lands have the overall strongest effect. The neural network has no such similar ability. The "reduced variable" models give some approximation of how the model reacts without that variable in place, but nothing about its effect, positive or negative. As removing the "absence of natural lands" variable from the neural network corresponds to the greatest drop in its predictive power, one may be led to believe that it is the absence of natural lands that has the greatest effect on explaining seasonal home distribution. However, with the sign of the coefficient supplied from the logit model, one can more accurately infer that it is the presence of natural lands that has the strongest influence on the model.

Another technique that can give additional insight into the models is mean absolute deviation (MAD). The value for MAD gives a way to measure the variability or spread of a set of numbers by computing their average distance to the mean (Sclove

1998). The MAD will compute the mean of the absolute value of the differences between the supplied values and the arithmetic mean of all values. The formula for MAD is:

$$\text{MAD} = (\text{Sum } (X - \mu)) / n$$

Where: n = the number of observations, X = the supplied values, and μ = the arithmetic mean of all the values.

Table 5.6 shows the results of the MAD for the observed and expected values used in the logit and neural network models. As can be seen, the spread of values is relatively close, with both models having small differences between the observed and expected values. The logit model computes a slightly lower deviation between the MAD for observed and expected values.

Logit model	Neural network
.1088	.1327

Table 5.6: Mean Absolute Deviation of the modeling approaches

The Spatial Error Signal and Residual Analysis

To better examine the differences between the estimated and the observed values, the model residuals were spatially mapped. The value of the observed minus estimated score was taken and mapped in the GIS. Figure 5.3 shows the spatial error of the logit and neural network models (note that the missing areas on the neural network map are from the 20% random sample of the MCDs removed for the training phase). As can be seen, the largest clusters of over or under prediction of values are occurring in sections of northern Minnesota, Wisconsin, and the Upper Peninsula and northern lower peninsula of

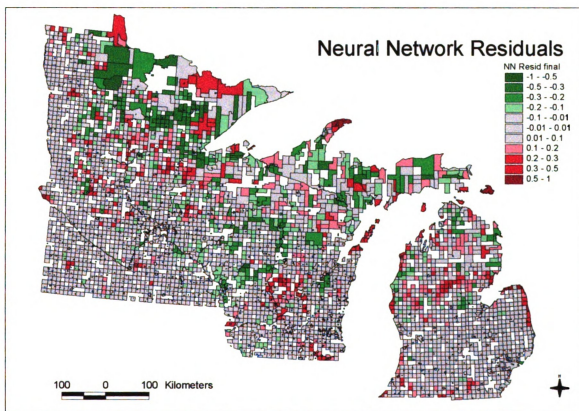
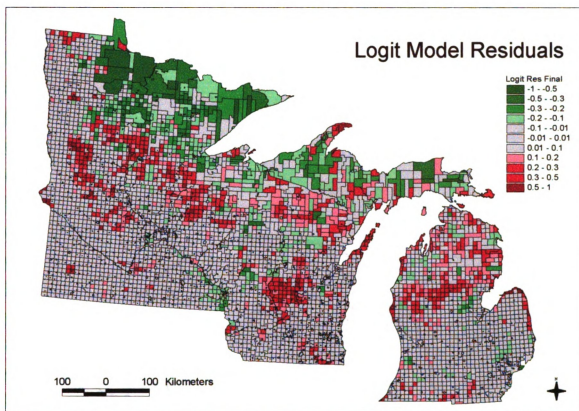


Figure 5.3: Logit and Neural Network Residuals

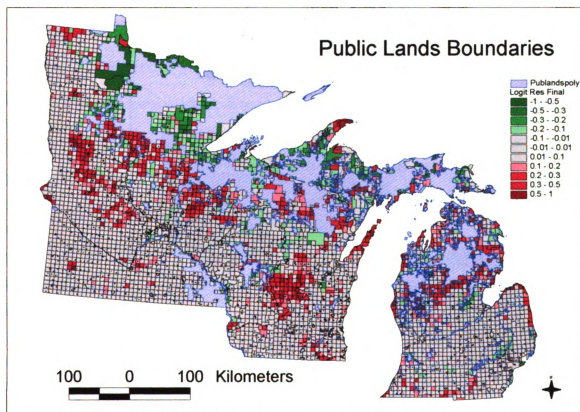
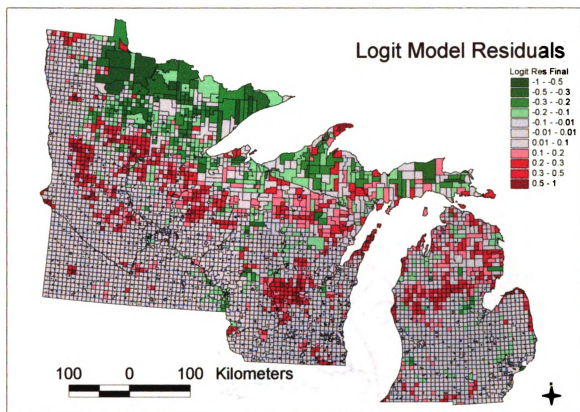


Figure 5.4: Logit Residuals with Public Lands Boundaries Overlaid

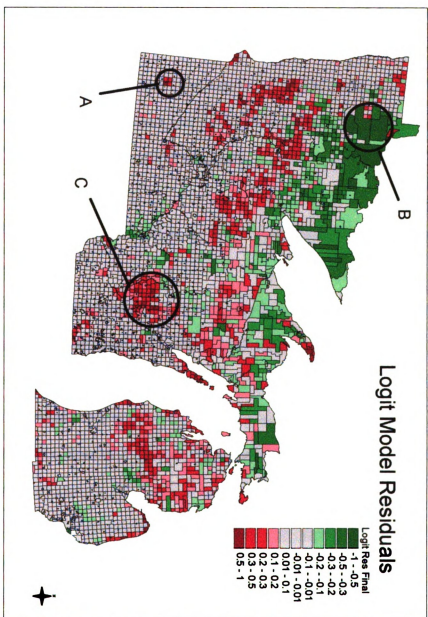


Figure 5.5: Selected Residuals

Michigan, with clusters of the highest positive and lowest negative values spread across the three states.

The model residuals appear to be spatially correlated. The first pattern to the residuals relates to the boundaries of public lands. Figure 5.4 shows the logit model residuals again with the public lands file overlayed on top of them. An analysis between the residuals and the public lands showed the following:

- Of the 567 MCD polygons which have their center within public lands boundaries (the most accurate measure of fit between the two overlays), 565 of them contained an MCD with at least some residual (observed minus estimated) greater than zero.
- Of the 567 public land MCD polygons, 196 of them contained a residual value of greater than 20.

Thus, MCDs falling within public land boundaries are being over or under estimated by the logit model. As noted in Chapter 2, public lands are areas that need to have special consideration taken when dealing with them. Not only does seasonal home development take place within their boundaries, but also commercial and residential development as well. Strict controls are placed on these developments, and they can only occur in certain parcels of land that are privately owned. Thus, the proportions of seasonal homes within public lands boundaries are a special case, as can be explicitly identified through mapping of the spatial error signal of the models.

Other clusters of residuals were examined as well. Three examples were chosen to demonstrate different effects of residuals in the analysis. Figure 5.5 indicates these circled on the map of logit model residuals.

The first of these (marked with an “A”) on the map, is a cluster of townships in southwestern Minnesota. Upon closer examination, these four townships are part of Murray County and centered upon a body of water and a public land polygon. From the Murray County website (Murray County 2000), some key information about this residual cluster can be obtained. Lake Shetek is home to a number of church camps, including the Shetek Lutheran Bible Camp and Lakota Retreat Center. This facility can hold up to 100 persons in the winter and up to 180 in warmer times. These facilities (along with a Baptist Bible Camp and a Boy Scout Camp) may be considered seasonal residences, and thus are factors unaccounted for by the model.

The second cluster of residuals (marked with a “B”) is located in the upper northern portion of Minnesota. This area is predominantly covered by land for the Red Lake Indian reservation (with a population of 5000). Although the area seems to fit many of the criteria for seasonal home development, none seems to be happening on Native American territorial boundaries.

The third cluster of residuals (marked with a “C”) is located in lower central Wisconsin. This grouping, predominantly centered on townships in Waushara County, contains many residuals. A closer look at the area in the Delorme Gazetteer of Wisconsin (Delorme 2000b) shows that there are numerous designated public boat launches ringing many of the lakes. In addition, there are designated hunting areas posted throughout the region. Also notable are the number of landing strips marked on the map, perhaps indicating many rural pilots in the area. A downhill ski resort is also located in the area. Most notably, however, are several small public land boundaries marked as state wildlife areas that are missing from the public lands database. This shows a number of factors

unaccounted for in the model (such as ski resorts and boat launches), as well as pointing to some holes in the available statewide data supplied from the Wisconsin DNR.

Logit Estimates vs. Neural Network Estimates

One final test was done for model comparison. The estimated results of the logit model and the neural network model were compared with one another to see just how similar the results of each model were. To keep the comparison accurate in terms of numbers and which observations were being compared to what, the 80% sample of the logit model used in the model validation procedure (the same 80% used in the testing of the neural network model) was used for comparison. The estimated values of the two models were found to have a correlation coefficient of 0.87. Furthermore, the values were placed into a Principal Components Analysis (another measure of linear correlation). The results of the PCA found the two sets of results rotated into the same component. Thus, the two models are explaining very similar things.

Conclusion

This chapter provided an analysis of the results of the various stages of the modeling approach. PCA reduced the predictor variables to a new set of independent variables, while the results of the logit and neural network model were compared and contrasted. This also achieved the last of the objectives that were laid out in Chapter 1: to identify the principal predictors of seasonal home location as well as to test the various modeling approaches. The next chapter places these results in the context of the theory and literature, and provides a more in-depth discussion of these results. Some of the

limitations of the data, variables, and modeling approaches are also discussed. Finally, some greater implications for areas of growth beyond this project are examined.

Chapter 6: Discussion

Introduction

This chapter will take a closer look at the results of this dissertation in the context of the overall theory and literature, and the contribution that this research makes to that theory. This dissertation encompassed a wide range of ideas and concepts, drawing from many different areas of theory and literature and combining them together. The nature of the model results and how they relate back to the literature and theory will be discussed, as well as some thoughts on the nature of the variables.

Also, this dissertation provided an exercise in model building in terms of defining variables and the formation of models to explain their contribution to the distribution of seasonal homes. However, there are some limitations to the study, not the least of which is the large data set used as the dependent variable, and the data sets used to construct the independent variables. A greater discussion of the modeling approaches and the limitations of linking huge data sets together with them are examined. Lastly, the use of geographic information systems provided a means at looking at the science underlying the system.

Dependent Variable

The dependent variable used in this project is the percentage of the total housing stock per MCD that is seasonal homes. This provides a good approximation of the influence of seasonal homes over permanent homes in a particular area, especially when taking into account the varying sizes of townships. A larger township may have a greater

number of homes due to its size, just as a much smaller township will have less. The proportion of seasonal homes to the total housing stock provides an attempt to adjust for this discrepancy in MCD size.

Another way of approaching the problem that future researchers may wish to examine is the use of the total number of seasonal homes per MCD. While this would give the best measure of the exact number of seasonal homes in the area, using this as the dependent variable introduces new problems and potential sources of error into the model. First, this does not account for the varying size of MCDs. A larger MCD would potentially contain more seasonal homes than average-or-smaller sized MCDs. A larger MCD in a “hot” area for seasonal home development would contain an inordinate number of seasonal homes when compared to other regions and would almost become an outlier in the model, even though it is an important factor. Also, as noted in Chapter 3 some areas with high numbers of seasonal homes exist within the township boundaries of large cities. Even though these areas would represent an insignificant portion of the total housing stock, they would be much higher than other areas. Lastly, the methods used in this project would require significant changes or adaptations to be used with such a study. The logit model and the neural network are designed to produce output scaled between 0 and 1 as a way to approximate the proportion variable. Different methods to fit this new dependent variable would need to be evaluated and selected, or several changes made to the data and methods being used here.

Another limitation of the data is the definition of what constitutes a seasonal home. Census definitions of seasonal homes cover a wide range of housing types including small cabins and cottages, to fully stocked and winterized homes, to large scale

condominium developments and time-shares. All of these varied types of homes are encompassed under the Census definition of “homes held for seasonal or occasional use.” This can lead to certain clusters of outliers in the model that were unaccounted for. For instance, the townships in Murray County, MN contain several seasonal cabins for numerous camps that may have increased the numbers of what is considered “seasonal homes” in that area.

The dependent variable contains over 2000 observations of “0” and another 1600+ observations between “1” and “5” indicating an overwhelming number of permanent homes in each township when compared to the number of seasonal homes. This could be due to areas of very little housing development, or heavily urbanized areas where permanent homes far outrank the seasonal ones. One hypothesis for dealing with the nature of the variable would be to remove all urbanized areas from the dependent variable as the models should be trying to exclude areas of heavy urban usage.

Some simple analyses were done by removing all observations from the data for townships that had a population of greater than 5000 persons and had a seasonal home percentage of less than 5. 535 townships fit this criteria, corresponding to urban areas in the three states including Detroit, Milwaukee, and Minneapolis-St. Paul. The models were re-run with this new dataset. The logit model slightly improved to a pseudo- R^2 of 0.535. The rankings of the coefficients changed slightly, as distance from cities lost some of its strength, reducing from an 0.29 to an 0.20 beta coefficient. The neural network model dropped from a pseudo- R^2 of 0.63 to a value of 0.60.

These results would indicate that there is a definite spatial pattern of seasonal home usage on the landscape occurring in urban areas. The neural network was

“learning” about the patterns of very low values in the dependent variable, and its overall fit dropped when those patterns were removed from its learning process. Similarly, it can be seen that the “distance from cities” variable was partially responsible for explaining the values of seasonal homes within heavily urban areas. One conclusion that can be drawn from this is that seasonal home distribution within urban areas is not an outlier, but part of the overall pattern of the dependent variable.

Independent Variables

This leads back to the general landscape ecology and human / environment theories on the nature of pattern and process (e.g. Forman and Godron 1986, Navah and Lieberman 1984, Skole and Gage 2001). In this theory, it is the patterns on the landscape (such as the arrangement of lakes, forests, and elevation) that lead to some process (in this case, the distribution of seasonal homes). The variables in this project were chosen to reflect the landscape patterns and they in turn are related to this process.

The choice of a seasonal home is dependent on a variety of factors, including several socio-economic ones, as well as personal choice. However, no matter if the reason for a second home purchase is eventual retirement, a weekend getaway, a cabin used for hunting, a time-share condo, or a place to park the snowmobile, there are numerous landscape features which are prevalent in some aspect of the choice of seasonal home location. Chapter 2 reviewed a number of land uses that relate to seasonal home development (or the lack of it) including waterfront, urban areas, natural areas, transportation, and public lands. It is the pattern of these land uses that in turn leads to the process of seasonal home location.

This is not to say the list of variables is all-inclusive. The fit of the models indicates some variables are missing from the analysis. The addition of origin-based data would likely also improve the overall fit and explanatory power of the models. Both Tombaugh (1970) and Gartner (1987) utilized survey data to gain origin-based information about socio-economic characteristics of seasonal homeowners and their purpose for purchasing a seasonal home. However, as noted in Chapter 2, this type of data was unavailable for this research. Land price is certainly a factor of where seasonal homes would be located. However, using 1990 land valuation variables would not get at the true nature of the dependent variable in the models. Without origin based data, there is no way to tell when seasonal homes were bought: several decades ago when the land they now reside on was relatively inexpensive, or in more recent years when the land is much more expensive. Also, the addition of a measure of township zoning would be beneficial to the model. The amount of land per township that can be developed would be another useful indicator of the distribution of seasonal homes. Areas with ample available land or lacking zoning regulations may be seeing greater numbers of homes, both seasonal and permanent.

Weather is another factor unaccounted for in the model. Areas that receive heavy snowfall may be high in proportions of seasonal homes whose owners wish to partake in winter sports. Snowmobiling and skiing both rely on snowfall for enjoyment. Likewise, areas experiencing milder climes in the summer may be more plentiful in seasonal home proportions since they would be better suited for more days of outdoor recreational activities such as boating, fishing, or swimming. The nature of weather and climate in these models is one that requires much further investigation.

As Stewart (1994) points out, individual choice also has much to do with the location of where to purchase a seasonal home. Also, origin-based data such as the race, age, income level, and reason for purchasing the second home would also be useful in providing more explanatory power, along with a measure of how far the purchasers would be traveling to utilize their new second home. At the three state level of analysis, these values would have to be averaged by township to be utilized in the modeling framework.

However, these are additional choice variables to help the predictive or explanatory power of the models. Returning to the idea of landscape pattern leading to a seasonal home process, perhaps “predictor variables” are not the best name for the set of independent variables used in this project. Perhaps “process driver” variables would be more apt to describe their function. In this way, the link to land use would be more explicitly laid out.

Bias in the system

The nature of the dependent variable will be causing bias in the models. Not all MCDs are the same size – they range from small villages dotted throughout the three states to near-county level sized MCDs in reaches of northern Minnesota or Michigan’s Upper Peninsula. The construction of the independent variables attempts to account for this by using a mean value for most of the variables. For example, the distance-related variables compute the euclidian distance from each feature to all other cells in the study extent. For the boundaries of each MCD, the mean value of the variable is computed in

the GIS. This was an attempt to not only assign a single value to each MCD, but to also try to account for the varying sizes of each MCD.

A source of bias in the data comes with the introduction of variables that do not represent the mean value for each MCD, most particularly the three “summation” variables – amount of waterfront, amount of Great Lakes lakeshore, and amount of forest per MCD. In these cases, the MCDs with the biggest size will be receiving higher values simply because they are larger. Likewise, smaller MCDs will be receiving much smaller values simply because of size. This problem becomes compounded with some of the larger MCDs in Michigan’s Upper Peninsula or areas of northern Minnesota which have copious amounts of waterfront or forest, or places like the Drummond Island MCD off the eastern coast of Michigan’s Upper Peninsula which has a very large amount of Great Lakes lakeshore. In these cases, these MCDs will have much higher values for each variable than other large MCDs and much higher values than smaller MCDs. In effect, due to their size, they have become outliers, or sources of bias, in the model.

These sources of bias then become perpetuated through the system. For instance, the Principal Components Analysis rotates variables together that are highly correlated, and then constructs standardized z-scores to be used as factor scores. Components containing high loadings for variables containing the largest values for the summation variables are receiving some observations of inordinately high z-scores due to their having values so much higher than the other observations for each variable. The components in which these variables load highly will be receiving this “bias” influence of these larger MCDs (or the extremely large ones). These high factor scores can be seen as

a source of bias in the logit model. Other extreme values could lead to similar sources of bias.

This source of bias also contributes to the neural network model. Neural networks can theoretically accept input values of any size. However, due to the nature of the neural network transfer (or “squashing”) function (such as the logistic sigmoid function), the input values should ideally take on the range of approximately (-1, +1). Input values significantly beyond this range will still be included in the analysis, but the neural network will be less sensitive to their presence and some saturation may occur (Statsoft 2000). With the standardized scores being slightly skewed by the extreme values of some summation observations, this could be changing how the neural network trains and tests on the data. The neural network could potentially then be learning less about those observations with the highest z-scores (in effect, those observations with the most extreme values anyway) and learning more about the majority of the observations.

This effect on the neural network could be interpreted in two ways. First, the neural network could be learning less about the most outlying patterns in the data, and learning more about general trends and patterns. In this way, the neural network itself is acting as a sort of “filter” on the data, learning less about the sources of bias and more about the non-noisy patterns. This could account for the higher pseudo R^2 of the neural network when compared to the logit model. However, secondly, the neural network could be potentially missing out on learning about certain necessary patterns of seasonal home distribution, and applying only more general learning to the process. “Overfitting” of the data may be occurring. This is one possible hypothesis for why the neural network’s pseudo- R^2 is higher than the logit model.

These are sources of bias that future researchers should be cautious about in dealing with this type of data, scale, and variables. However, there are some potential alternatives to removing the bias from the system. First, the largest-sized MCDs could be removed from the analysis. This, however, is not solving the problem as the size of the various MCDs has been attempted to be accounted for by using the “mean” value for the independent variables. The varying sizes of townships are something that must be accounted for, not simply eliminated. Another alternative is to drop the extreme outlier values from the procedure entirely. This is much the same as dropping the MCDs, only dropping them after the variables have been computed and finding which ones are contributing these overly high values.

Another choice would be to remove the summation variables from the analysis, or replace them with other surrogates. However, the theory and literature concerning seasonal home distribution indicates that these are important variables in the analysis. For instance, the amount of waterfront is the best indicator of the potential building space around a lake or river. Simply using the amount of waterfront per MCD (a ratio) may give skewed results. For example, a 100 square mile MCD containing 10 miles of waterfront would produce a value of 0.1. A 10 square mile MCD containing 4 miles of waterfront would produce a value of 0.4. In this case, the MCD with less waterfront would be weighted higher than the MCD with more waterfront. Thus, using a ratio value to produce a percentage will not give accurate results. Much cruder estimators potentially could be constructed and used as surrogates (such as percentages of MCD perimeters), however, these may have problems associated with them as well.

A final alternative would be to not use the factor scores from the Principal Components Analysis in future modeling efforts, but rather to use the highest loading variable in each component. However, this introduces a new host of different potential problems to the modeling process. The object of using the factor scores was that they represented values for orthogonally rotated components in n-dimensional space. These components, being orthogonal, were uncorrelated with each other, and thus the new variables created from the factor scores represented observations of truly uncorrelated variables. These new variables could safely be used in a reconstituted regression in the logit regression model without violating the assumption of no multicollinearity. By using the highest loading variable, this introduces multicollinearity into the model and violates one of the assumptions. This could cause other problems with the model that are accounted for by using the factor scores.

Similarly, using the variables with the highest loading in each component could cause problems with the neural network as well. The variables would have to be linearly normalized (selecting the highest value among the observations and dividing all observations by that value), thus reducing the range of the variable to between 0 and 1. However, any significant outliers in the observations will throw off the range of the data. For instance, if one of the summation variables (i.e. amount of waterfront) was the highest loading variable, it would have a small number of observations that were significantly higher than the others. Normalizing this variable for use in the neural network would mean dividing by the highest value (i.e. an extreme outlier with an observation value much higher than all other values), and thus significantly altering the variable. It would be left with a handful of high values, with the majority of the other

values reduced so low they would probably be seen as constants when contrasted with the other observation patterns the neural network would be learning about. Thus, the neural network would be learning very little, if anything, about that particular variable. Similar results could occur if an anomalous value (or a significantly higher value) for one of the other variables was part of the dataset.

While these are some examples of the effect of bias in the data used in this project, they are not unique when dealing with large data sets. Examining the seasonal home phenomenon at such a broad three-state-range, while dealing with the non-uniform sizes of MCDs will cause problems of bias to occur throughout the analysis.

With large data sets, it is often difficult to pinpoint exactly where errors or sources of bias are coming from. Sources of bias may be uncovered, but are difficult to locate or determine due to the large size, nature, and complexity of the data set. Anomalous or missing values may prove difficult to locate or correct. Also, when using large data sets from such a wide variety of sources, mismatches between data sets may occur. As such, the results emanating from large data sets cannot simply be taken at face value – investigation into potential sources of bias and inaccuracy is required as part of the examination of the validity of the results. Future researchers should be aware of these potential problems when working at the MCD level or dealing with large ecological datasets.

Methods and Model Building

The second and third objectives of this dissertation revolved around model building concepts, namely testing approaches to predict the seasonal home distribution on

the landscape and provide an explanation of this process. The two non-linear modeling formats (the logit and the neural network models) provided a means of coupling the initial GIS-based framework of the LTM to the outputs. In this sense, the movement from pattern to process is again followed: the GIS is used to identify the shapes and patterns of landscape phenomena, while the models are used to explain the process. However, this use of the term “explanatory power” should truly be applied to the logit model. While the logit model had less predictive power than the neural network, its explanatory value (as outlined in Chapter 5) is much greater than the neural network.

Neural networks provide no explanatory power about the data. Due to their structure, neural networks do not make assumptions about how data and the relationships between them are structured. The nature of the neural net is such that it can examine data without needing assumptions about this structure, while statistical models do make certain assumptions about what relationships the data will take.

Neural networks are often labeled as a “black box” since the processes of the learning and distribution of the weights, and thus the manner of just what contribution the variables are making to the end result of the model are all invisible to the viewer. The changes in weights that occur during the training cycles are also invisible. Weights may become altered in unexpected ways, growing larger or smaller than expected, but these functions cannot be viewed by the user of a neural network. Likewise, changes in the number of hidden nodes, or the number of hidden layers can theoretically affect the outcome of the model, as the number and variety of the weightings would be changed. In essence, tinkering with the structure of the hidden layer(s) is forming a new model. A problem with this process is that neural network theory is at such an embryonic stage that

there really are no set guidelines on how many hidden nodes or hidden layers are the optimal number to provide the best fit of the model. In this case, statistical models may be preferential, as they follow specific rules and utilize specific equations for which parameters may be specified. In addition, these coefficients and parameters of the model may be specified and examined to find explanatory power in the statistical model.

Neural networks have similar qualities to logistic regression as both are non-linear in nature. Neural networks and logistic regression are commonly used techniques in medical research (Piantadosi et al. 1994). It is noted that “neural processes will always give you results as good as statistical techniques and usually better” and that “they (neural networks) get you a much closer fit of the data in a lot less time” (Statcon 2000). Sclove (2000) notes that while logistic regression and neural networks are similar approaches, “standard and non-standard statistical techniques need to be compared with neural networks to understand the relative advantages and trade-offs among these different tools.” Also, the two methods share the use of the logistic transfer function. As noted in Chapter 4, the neural network used in this project utilizes a sigmoid function that converts the summations of the hidden nodes from their values to fit a logistic curve of values between 0 and 1. In this fashion, the neural nets are attempting to approximate the data in the hidden nodes to the non-linear form of the logistic curve. Thus, both methods are using similar methods to fit the data to the functional logistic form.

The Science Under The System

From the theories of pattern to process to model building, the question must be raised of where does this fit in to the nature of GIS? Is GIS just a toolkit for processing

information, or is there an underlying science behind using a GIS-based framework for this analysis? This is currently a much-debated topic in the field of geography: is GIS a geographic information system or a geographic information science? This debate seems to polarize into two camps – those who see GIS as a tool and those who view it as a science. The GIS internet discussion group GIS-L featured many responses dealing with the subject of GIS as a tool or a science (Wright et al. 1997b) indicating a great deal of interest by the research community in relation to the nature of GIS. The study by Wellar et al. (1994) indicated a large number of published journal articles concerned with the link between GIS / Quantitative procedures and scientific inquiry.

Goodchild (1992) notes that “science is an understanding of the issues that underlie a system.” Likewise, Fisher (1997) defines a science as “the state of fact of knowing ... knowledge or cognizance of something specified or implied ... knowledge acquired by study ... a particular branch of knowledge.” Similarly, Fisher (1997) sees a “system” “an organized or connected group of objects ... a set or assemblage of things connected, associated, or independent so as to form a complex unity ... a whole composed of parts in an orderly arrangement according to some scheme or plan.” In the case of GIS, the “system” (or tool) would be the hardware, software, and related tools used to implement tasks, while the “science” would be the underlying theory and design which the system is implementing.

The validity of Geographic Information Science has been championed by many in the field (Goodchild 1992, Openshaw 1998, Fisher 1997, Wright et al. 1997). Historically, there has been a strong underpinning of the science of spatial data handling beneath the system. Spatial manipulations such as overlay operations have been

documented back to 1912 (Steinitz 1976) with uses in environmental quality issues or land use planning. Likewise McHarg (1969) used spatial overlay operations to try and maximize social benefit and minimize environmental cost in landscape planning operations. The science has been there all along, while computer programs such as ESRI's Arc/Info and ArcView are all GI Systems for doing more modern and advanced forms of these spatial data operations. When the original Geographic Information System (the Canada Geographic Information System) was developed by Roger Tomlinson in the 1960s (Marble and Sandhu 1994) it was not technology driven (given the state of computer systems technology at that time). Rather, there was an underlying science behind the system.

This research can be seen as an examination of some of the science underlying the system. The spatial construction of the variables in the GIS shows some underlying science to the system – analyses such as distance or neighborhood functions incorporate spatial techniques. Likewise, studying this type of phenomenon at a spatial level adds another component rather than just tabulating statistics. For instance, the ability to spatially map out residuals provides a link between the underlying spatial science (in terms of visualization and analysis) and the software of the system.

Implications for Land Use and Potential Land Use Change

Though this project provided a look at one point in time, the examination of areas of high proportions of seasonal homes carries some implications beyond one time frame. The concepts and ideas outlined in this section are beyond the scope of the results of this project (and fodder for future research), yet there is some room for inference and

speculation partly based on the modeling results. Tombaugh (1970) similarly made predictions of future developments based on an analysis of one point in time, though he notes that a strong degree of caution should be exercised in making such forecasts.

The development of a new seasonal home at a particular site brings with it a number of effects, the first among these is the onset of further development. The results of the spatial autocorrelation analysis partly drive this idea. Areas of the higher proportions of seasonal homes may be areas where more seasonal homes may be found in the future. This fits in with Butler's (1980) life cycle concept, which seems applicable to this aspect of recreational tourism. Butler posits that destinations will grow until a stagnation level is reached and then will plateau and either reinvent themselves or fall into decline. Some areas with higher proportions of seasonal homes may grow in numbers of both permanent and second homes, becoming more of an area catering to increased numbers of tourists and seasonal migrants.

As the natural areas draw in more seasonal home development, the area becomes despoiled as new seasonal homes are built. Further land use changes occur when the seasonal population increases, and a need for basic amenities is created. Construction of grocery stores, gas stations, and medical facilities can soon lead to development of restaurants, shopping facilities, and hotels for tourists. Thus, areas which are highly clustered in seasonal homes could see development of both seasonal and permanent homes in the adjoining area or properties. This can also be seen as a continuance of the "self destruction" idea of tourism (Holder 1988).

In some ways, seasonal homes can be seen as endemic of some of the problems associated with ecotourism. The concepts of ecotourism revolve around tourists and

visitors arriving into an area to enjoy some natural feature, whether wildlife or “untouched” landscape features. The frequent slogan behind ecotourism is for visitors to “take only pictures, but leave only footprints,” implying that while tourism may take place in a locale, that location will remain untouched by the wake of the tourist. This type of “green” or “low-impact” tourism is frequently identified as the future of tourism and an ecologically viable alternative to mass tourism and development (Mader 1988, Romeril 1989, Hunter and Green 1995). However, as other authors point out (Butler 1990, Wheeler 1991, Jarvilouma 1992) ecotourism is bringing tourists and visitors into sites that were otherwise inaccessible to them. As these areas grow in popularity and word gets out about them, more visitors will arrive, perhaps more than the area can handle. With this increased tourism to the site, new developments must be made to accommodate the rising influx of visitors, thus leading to new developments, and the eventual rise of the area on Butler’s life cycle model (Butler 1980) from untouched to developed. Strapp (1988) found similar results when examining a predominantly seasonal home area in the Sauble Beach region of Ontario, Canada, and found it to be in the stagnation or decline stages of the life cycle while developing into a more permanent home area.

This same sort of theory can be applied to the development of seasonal homes. Frequently, seasonal homes are developed in areas away from large cities and in proximity to natural areas. In fact, as seen in Chapter 5, the presence of water and the density of natural areas were identified as the principal variables involved with the proportion of seasonal homes in an area, with factors such as distance from cities playing a part. However, it is development at these natural or more distant areas that leads them

to become steadily less natural or less remote. Lundgren's (1974) model of seasonal home development in Canada showed the initial seasonal home (or cottage) region at some distance outside the city. As the city grew and expanded, this seasonal home region is gradually overtaken by the expansion process, leading to it being subsumed within the city proper. Thus, a new seasonal home area is established farther out from this new city boundary. The implication is that as the city continues to grow, this new seasonal home area will be consumed into the city environs, leading to the establishment of a third seasonal home region further out from the city. Like ecotourism, development of seasonal homes can be seen as having short term benefits but long term effects.

An example of this type of scenario can be seen in Door County, Wisconsin. The Door Peninsula juts into Lake Michigan from the eastern shores of Wisconsin, and with its 250 miles of shoreline is one of the top tourist destinations in the state. The earliest visitors came to the area from Chicago and Milwaukee to enjoy the scenic landscapes, picturesque beaches, and quiet lifestyle of the peninsula (Hart 1984). Door County rose to prominence after being described as "A Kingdom So Delicious" by National Geographic magazine in 1970 (Bosselman et al. 1999). Today, Door County is a prime destination for all manner of tourists, boasting itself as one of the ten best places in the world for cross country skiing (Hart 1984) as well as numerous opportunities for boating, swimming, hiking, sightseeing, as well as unique shopping and dining experiences. In addition, Door County has become a popular choice for seasonal home construction, particularly large condominium developments which are used as time-shares (Bosselman et al. 1999).

According to the 1990 Census, Door County has a year-round population of 26,000 persons (Scheberle and Pagel 1999). However, during peak tourist season, there are approximately 40,000 tourists and seasonal residents per day in the county (Scheberle and Pagel 1999). Door County has seen its number of seasonal home units increase by more than 50% between 1970 and 1990, with seasonal homes accounting for roughly 34% of the total housing stock in the area (Scheberle and Pagel 1999). It is estimated that second homes in Door County are priced between \$100,000 and multi-millions (Ronan 1999). Frequently, one can see beachfront homes (with fenced or roped off areas declaring the land to be private property) crowded in and among resorts and areas jammed with tourists.

This level of seasonal and tourist development has brought many land use changes to Door County. For instance, residential land use in the county is on the rise, slowly but steadily, along with growing values for developed land use. Agriculture land use is in decline, along with the total natural land use of the area (Scheberle and Pagel 1999). With much shoreline development under protective private ownership or public control, development has grown in the inland rural areas. Door County appears to have been an area once used as a seasonal vacation getaway that is now slowly becoming a more commercial tourist area along with a growing and rising trend in seasonal homes. One small example of the growing tourist influence is the presence of the unique “fish boil” seen at numerous restaurants in the area and billed as a Door County specialty. However, as Hart (1984) points out, this is a definite invention to draw tourists, as this “ancient Scandinavian tradition” (as it is billed) was actually invented at an early stage of the 20th century.

Conclusion

This chapter provides a final discussion of some points of the dissertation and how its place in the overall theory of geography. Seasonal home development may largely be a tourism-related phenomenon, but it can be traced to other branches of theory and development, and is heavily influenced by the patterns of use and cover of the landscape. Similarly, the model-building portion of this project sheds new insight on the nature of statistical models when compared to the new theories of neural networks. Limitations of the use of large data sets and the varying sizes of MCDs have been examined as a source of bias that gets uniquely perpetuated through the system. Also, the role of science of GIS is examined. Finally, some speculations as to potential implications and extensions of this study are examined. This project folds in on itself, drawing from several branches of theory and returning to those theories once more.

The final chapter of this dissertation uses the results of this project and some topics raised in this discussion to outline a series of future directions for research that can use this project as a starting point to build from.

Chapter 7: Future Directions

A Baseline Approach

The project has value and significance as original research, both in contributing to theory and to application. This project is relevant in its contribution to the branches of theory of landscape ecology, tourism development, and geographic information science. This project brought together three disparate fields of theory that have been rarely considered together. As a landscape ecology project, this study aids in identifying the weight of various landscape factors to the distribution of seasonal homes. Likewise, the field of tourist development will benefit from the focus placed on this project. As noted previously, little research has been done in the area of spatially predicting areas of seasonal homes, or in spatial modeling of tourism and seasonal home development. The project can also contribute to geographic information science as it has tested the strength of the modeling approaches in an overall GIS-based modeling framework. In regards to the theory of spatial distribution, this project shows a validation of those theories in the notion that the patterns of the landscape are factors in driving overall human processes.

This project also provides an applied test of the GIS-based framework and the LTM's functions in development of independent variables, and while no true land transformation is being measured, the LTM's modeling techniques are being tested in the prediction of seasonal home areas. The results of this project may prove of use to state, county, or township planners to aid in identifying areas of seasonal home distribution. From these results, future modeling efforts of examining temporal change may be done.

By identifying the principal factors of second home location, as well as the weight and contribution of each factor, the ground is readied so that future predictions of seasonal home developments can be made. Spatial residuals from the model are targets for learning; planners will have direction where to go for more information to learn why these areas are over or underestimated by the model. This notion can be especially seen in the spatial error signal analysis presented in Chapter 5. Many residual areas were seen to be highly correlated with the distribution of public lands. As noted in Chapter 2, while public lands play host to a wide range of natural features and outdoor recreation, there are numerous problems involved with them, including misidentification of borders and growing privatization and development within their boundaries.

This research also contributed to the science of model building. Though it is not a heavily methodological project, several results can be made concerning the notion of explanatory vs. predictive power of models. Also, the combination of several factors (extremely large data sets, statistical techniques, and basic neural network approaches) pointed out some limitations and shortcomings of the various applications. This project has identified several areas requiring further study or development, and hints at other techniques that could be used.

The use of this project should be considered a baseline for things to come. The theory behind the project and its link to geography has been addressed, as well as development of independent variables through GIS. Various modeling approaches have been defined, and their abilities and shortcomings noted. As this project utilizes such a large scale, it should be seen as a “broad brush” in the sense that several of the project’s limitations are created by setting the study in such a large region. While this project is

complete unto itself, it provides numerous future directions from which to proceed.

These are described as follows:

Regionalization

The broad sweep of seasonal home distribution at the three-state level is captured in this project. At such a level, many generalizations had to be made, and the varying geography was one source of noise in the models (as noted in Chapter 6). With this baseline study complete, more detailed work can be done at a smaller regional scale. At a smaller scale, better data sources could be utilized for some of the variables. For instance, if the study was being done at a county or multi-county level, perhaps more accurate public lands data could be obtained. In a smaller region, more accurate data could be created by hand or obtained from local sources, things that were unfeasible at a three-state level. In addition, a smaller, regional, data set would be much more manageable to work with; data observations could be closely examined and sources of noise or error would be more readily identifiable.

Similarly, additional data could be obtained to utilize in the model. Information that was unavailable at a three-state level could be collected at a more regional level. Data such as locations of public boat launches, airstrips, ski resorts, or snowmobile trails could be obtained at a smaller scale. Working at a regional scale would also allow more detailed information about the types of urban areas, their extents, and information about local seasonal home patterns or developments. In addition, more localized data such as zoning regulations, potential zoning plans, or information regarding the amount of available land that can be developed on can all be collected at a smaller scale.

Working at a smaller scale could also allow for the collection of origin-based data. Survey data taken at a community level could provide insight into many of the missing variables as outlined in Chapter 6. Socio-economic data about individuals, patterns of use, and travel time to and from a seasonal home. Furthermore, information about the purpose of the home, the year bought, and the value of the land could all prove beneficial to the models' overall fit and explanatory power.

This regional scale is adjustable. Studies could be done within a largely seasonal community, requiring work at the parcel level scale. An entire township could be modeled as well. Scaling up to the county level would allow for an adaptation of the analyses presented here, where controls could be developed to account for variations in township size. Counties such as Door, WI, Grand Traverse, MI, or Otter Tail, MN would be sites to perform such analysis. Finally, a more broad regional scale could encompass a multi-county area, such as the six counties comprising the Grand Traverse Bay Watershed. At this type of regional scale, the techniques used in this project could be incorporated with a host of new variables that are missing from the original study.

Temporal Modeling

With the imminent arrival of the 2000 Census, data will be available to perform temporal modeling. This would take the form of modeling changes between townships to see if they went up or down in their percentage of seasonal homes between 1990 and 2000. From there, development trends could be modeled. Once again, a broad baseline study could be performed, or this could be used to follow up on one or more regional

studies. In this sense, a regional study at 1990 could be performed, another study of the same area with 2000 data, and the results could be compared.

There is a definite need for studying the trends of seasonal home development, as it is on the rise. In 1999, 377,000 seasonal homes were purchased or built, up from 345,000 units in 1997 (NAR 2000). This figure is up considerably from sales of 296,000 in 1995 (NAR 2000). This rise in seasonal home development and sales comes at the same time as a tax-law change in 1997, which “essentially did away with the capital gains tax penalty for most buyers wishing to trade down to a smaller primary residence and also use some of their equity to purchase a second home” (NAR 2000). With the favorable economy that the 1990’s have enjoyed and changes in the tax law structure to allow for an easier purchase of seasonal homes, there has been a corresponding boom in seasonal home sales. Approximately 6% of all homes sold each year are seasonal homes (NAR 2000).

There is a definite growing trend of seasonal development, as the number of seasonal homes in the United States has increased over time. Table 7.1 shows the total number of seasonal homes in the US from 1940 through 1990.

YEAR	1940	1950	1960	1970	1980	1990
Seasonal	739,594	1,050,466	2,024,381	2,020,087	2,794,054	3,116,867
All Homes	37,325,470	45,983,398	58,326,357	68,679,030	88,411,263	102,263,678

Table 7.1: US Housing and Seasonal Homes by year (source: US Census Bureau)

Figure 7.1 shows the geographic distribution of seasonal homes in the US for the same time period. As can be seen, the numbers of seasonal homes in the US is growing

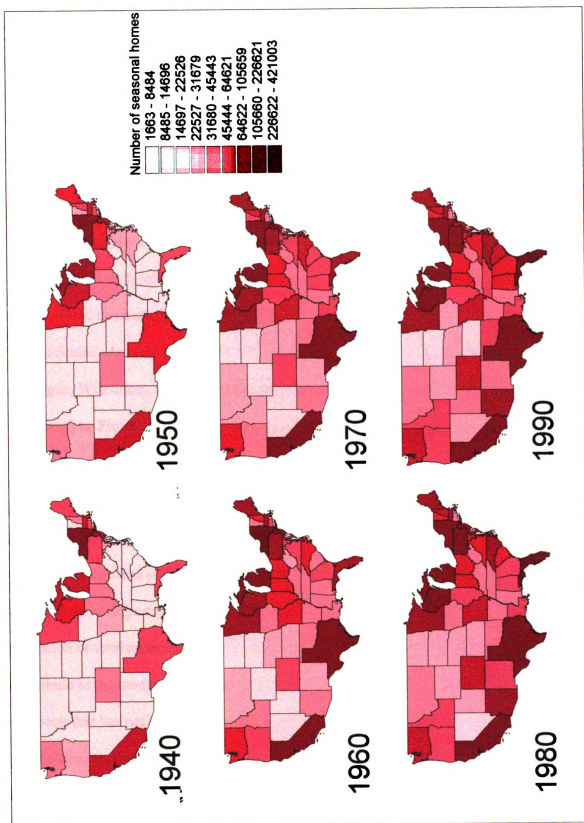


Figure 7.1: Number of seasonal homes in the United States, 1940-1990

at a steady rate. By 1990, seasonal homes account for roughly 3% of the total housing stock in the US. Areas in the south (including Florida, which leads the US in the number of seasonal homes) and the mid-Atlantic are growing, and also regions in the north, including parts of New England and the Lake State region of Wisconsin, Minnesota, and Michigan. California has seen a boom in seasonal home development. In fact, in 1999, Sugarloaf, California saw 93.3% of all housing sales in the area come from seasonal home sales (Fogarty 2000).

Thus, there are definite growing trends in seasonal home development in the United States. As can be seen, this type of study has applications to areas outside the Michigan, Wisconsin, and Minnesota area. Places in California (the aforementioned Sugarloaf) or Florida (in booming seasonal home markets such as Naples) would be prime candidates for other regional and temporal studies.

Adapting for Land Use Change

An extension of the temporal modeling project would be to utilize the data to do true land use change modeling. The literature lacks models that explicitly focus on recreation or tourism as land use change drivers, so this project would be able to contribute to the theory of land use change as well as develop a new type of modeling approach that has not been done as yet. As Campbell (1998) notes, there is more to human-environment interaction than simply the biophysical level of study. For example, the Kite framework (Campbell and Olson 1991) provides a means of encapsulating the

dimensions of human interaction, the nature of power, and the environmental changes occurring.

Future uses of the land (in relation to seasonal home development) may also be forecasted through the use of these modeling techniques. As one of the goals of this study is to identify the predictor variables of seasonal home location, this portion of future analysis is already completed. Thus, future research for true land transformation studies can begin at a higher plane.

The use of the LTM in these types of studies takes advantage of its full capacity as a model of land use change. By examining land use at two time steps, the change between the two can be modeled. This would build off of the “temporal modeling” approach previously outlined, and would probably be best to study at a regional scale. Gathering data about locations of seasonal homes could be best done at a parcel level, where tax records could give indications of which homes were held for seasonal use and which were permanent at two time steps. Process-driver variables of housing development could be constructed and the resulting land use changes (in the form of further seasonal home development) could be examined.

Another application would be to take the results of the temporal change study and identify areas high in seasonal home development (i.e. areas which showed a strong increase in seasonal homes between the 1990 and 2000 census), and do land use change modeling on those areas. The results of this project point to some of the principal predictors of seasonal home development that could be used as variables in the model. Areas that changed to a residential land use would then be predicted by the model.

A limitation to this type of study is the reliance on historical patterns to forecast future trends. A regression style land use model (Theobald and Hobbs 1998) will be doing just this. If seasonal home development areas have fallen into the stagnation or decline stages of Butler's life cycle model (similar to a case study examined by Strapp 1988), then population trends may give spurious results. In cases such as this, on site evaluation may be necessary to determine the type of changes occurring. Otherwise, the model may be forecasting a saturation effect of urbanization in the area where no such development is occurring. However, if the area is experiencing fast growth due to seasonal home development (or other factors) then population forecasts may be too low for the area.

Land use change modeling approaches can be used for other forms of recreational based tourism as well, beyond seasonal homes. Tourism has been seen as an "engine of development" (Roehl and Fesenmaier 1987), as localities throughout the US have made the transition to a tourism-based economy. Several initial tourism developments occur to provide economic benefits to a particular area (Pearce 1989, Witt and Moutinho 1994), or to meet the demands of travelers to the region (Ryan 1991). In fact, tourism comprises the world's largest economic sector (WTO 1995) and carries with it significant changes to the landscape and physical environment (Cohen 1978, Pigram 1980, Stroud 1983, May 1991, Ryan 1991, McKercher 1992). The land is changed from its natural state to new usages specifically designed for tourism, such as resorts, amusement parks, shops, parking areas, airstrips, and boat docks. Tourism drives changes in population and the local economy. These factors in turn spur changes in the way land is utilized.

For instance, Walker (1991) notes that the natural flora at Heron Island, Great Barrier Reef, has been disturbed by tourism development to the point where exotic species now outnumber the natural ones. The Sikkim Himalaya region provides another example; tourism has caused a greater degree of pastoralization, and thus a marked increase in timber harvesting and forest destruction for fuel and fodder (Rai and Sundriyal 1997). Coastal areas have seen sewage disposal, desalinization, irrigation, infilling, and sediment from construction contribute to the degradation of coral reefs (Hawkins and Roberts 1994). It is these types of effects that application of land use change models to tourism-related phenomena would be of benefit.

Other Applications of the Framework

The modeling framework outlined in this project, at its most extremely basic level takes a number of spatially constructed variables and uses them as independent variables in varying modeling approaches to predict a dependent variable bounded by 0 and 1, rather than a binary response of 0 or 1. This basic modeling framework can be adapted and applied to other sources of data as well. For instance, one application of the model could be to explain the spread of farms across the landscape. Data would have to be collected to represent the proportion of farms of the total housing stock in a particular area. A new set of independent variables would be required to be built from new branches of theory and literature, but the basic modeling framework to undertake such a study is now in place.

Methodological Exploration

Much more can be done with neural networks than is shown in this project. Their use in this project is to examine another potential modeling approach and to provide a comparison with established statistical techniques. Numerous other forms of neural network learning algorithms are available to use, including bidirectional associative memory, adapted resonance theory, and counter-propagation networks (Freeman and Skapura 1992). Likewise, as noted in Chapter 6, altering the number of hidden nodes or hidden layers changes the structure of the network and its learning abilities. Too many or too few hidden nodes can cause generalization errors in the system (Sarle 2001). There is no set rule for the number of hidden nodes to use (Swingler 1996), and examples may be constructed to nullify some rules of thumb (Sarle 2001). The correct number of nodes seems to depend on much exploration of the data, the number of training cases, and the complexity of the function involved (Sarle 2001). Addition of extra hidden layers is another concept that requires investigation, but is beyond the capabilities of the neural network software used in this project.

The problems with the data as outlined in Chapter 6 were causing certain problems with the functionality of the neural network. The following analysis is beyond the range of the results, but one hypothesis is that this noise in the data could potentially be causing what is called “activation function saturation” in the neural network (Skapura 1996). If enough saturation occurs, the derivations for the weights may be reduced and learning may be slowed (Reed and Marks 1999). A typical cause of saturation is large external inputs (Reed and Marks 1999) that are the outlier data values. This could cause

the model to be overfitting or underfitting the data. Reed and Marks (1999) suggest that the data should be altered to fit to a reasonable range, yet the standardization procedure of the PCA has done just this. However, simply altering values will not solve the problem, as the weight initialization range to avoid saturation is dependent on many effects, including the number of inputs to a node, and the correlations between inputs (Reed and Marks 1999). Even accounting for this, Reed and Marks (1999) note that since the weights of a neural network change during the learning process, saturation may develop at later levels. Obviously, this is not a problem with a simple solution.

Neural network research is still at an embryonic stage of development. There are very few hard and fast rules that are applicable to questions concerning the usage of neural networks. Neural networks provide strong predictive abilities, but care must be given to interpreting those results. As this study has shown, a neural network suffers from a multitude of problems including how it handles data, the structure of its own architecture, and its label as a “black box.” Unlike statistical models, one cannot easily peer into the innards of the neural network and derive results.

A highly methodological paper on the architecture of neural networks and how they relate to spatial problems is suggested as a future direction for research. Many of the above concepts are speculations about how the neural network is handling the data and potential problems with the data. A heavy investigation into the structure and usage of neural networks would aid in clearing up several of these mysteries and may possibly find solutions to them.

Spatial Statistics Approaches

There are other modeling approaches that can be used to predict the distribution of a phenomenon across a landscape. Spatial statistics (and in particular spatial interpolation methods) can provide a measure of patterns on the landscape. A random sample of points on the landscape can in turn provide a measure of the values of nearby points. Numerous applications can be done with this type of data. Cross-validation (Bailey and Gatrell 1995) is a method involving removing one observation from a dataset and using all other points to interpolate the missing value. Ordinary kriging is a technique that estimates a local mean (Bailey and Gatrell 1995) and is a common tool in predicting values along a continuous surface.

To use these spatial techniques with seasonal home distribution, the MCD polygons would first have to be converted to point data. The centroids of each MCD were found, and as they are so clustered in space, the MCDs can be evaluated as points (in a similar method to the spatial autocorrelation method described in Chapter 5). From there, ordinary kriging can be performed on a subset of the data to interpolate the surface.

A simple experiment was done with ordinary kriging to determine the usefulness of this application in future studies. Ordinary kriging was chosen as makes no assumption of a local trends. However, ordinary kriging assumes a mean of 0 (Isaaks and Srivastava 1989). While the seasonal home distribution data does not meet this criteria, the residuals do, so kriging will be performed on them. The semivariogram shown in Chapter 5 was used to compute the range, sill, and nugget for the kriging. A randomly selected sample of 20% of the polygons (the same 20% used in the neural network testing) was used and ordinary kriging was done to interpolate out the remaining 80%.

The estimated output values were compared to the observed values and a correlation coefficient and pseudo- R^2 were computed to compare the estimates with the other results. Using a 20% sample produced a correlation coefficient of .69, which becomes a pseudo- R^2 of .47, indicating a good fit to the landscape. This is a simple experiment, but shows the applicability of additional techniques such as spatial statistics to this project. Future research could expand on these methods with more in-depth and detailed examinations of the applicability of spatial statistics to seasonal home or recreational tourism data.

Conclusion

This chapter has shown that this study, while a broad brush over three states, can be used as a baseline for a myriad of future projects to come. The results and conclusions can be adapted for use on a smaller, more regional setting. This type of environment may provide the best setting for studying a phenomenon such as seasonal home distribution as more local factors can be accounted for. In addition, greater amounts and more accurate data can be obtained at a regional scale.

With the advent of the 2000 Census, temporal change between 1990 and 2000 can be computed as well. This project also lends itself to land use change modeling. The framework of the study could be utilized in other environments to predict similar measures, such as farms. Finally, a more in-depth examination of neural networks is suggested, especially with regard to their format, design, architecture, and applicability to other projects within the social or natural sciences. Lastly, other techniques, such as spatial statistics, can be utilized in future extensions of this project.

BIBLIOGRAPHY

- Albert, D. (1995). "Regional landscape ecosystems of Michigan, Minnesota, and Wisconsin: a working map of classification." St Paul, MN, US Department of Agriculture, Forest Service, North Carolina Forest Experiment Station.
- Alonso, W. (1964). Location and Land Use: To A General Theory of Land Rent. Cambridge, Harvard University Press.
- Bailey, T.C. and A.C. Gatrell (1995). Interactive Spatial Data Analysis. Essex, Longman Group Limited.
- Baker, W.L. (1989). "A review of models of landscape change." Landscape Ecology 2(2): 111-133.
- Barlowe, R. (1986). Land Resource Economics: The Economics of Real Estate 4th edition. Englewood Cliffs, Prentice Hall.
- Bishop, C. (1995). Neural Networks for Pattern Recognition. Oxford, University Press.
- Bishop, I.D. and D.W. Hulse (1994). "Prediction of Scenic Beauty Using Mapped Data and Geographic Information Systems." Landscape and Urban Planning 30: 59-70.
- Bosselman, F. P., C. A. Peterson, and C. McCarthy (1999). Managing Tourism Growth: Issues and Application. Washington DC, Island Press.
- Boutt, D., D. Hyndman, B. Pijanowski, and D. Long (2001). "Identifying Potential Land Use Derived Solute Sources to Stream Baseflow Using Ground Water Models and GIS." Groundwater 39(1): 24-34.
- Bull, C., P. Daniel, and M. Hopkinson (1984). The Geography of Rural Resources. Edinburgh, Oliver & Boyd.
- Burley, J. B. and Brown, T. (1995). "Constructing Interpretable Environments from Multidimensional Data: GIS Suitability Overlays and Principal Components Analysis." Journal of Environmental Planning and Management 38 (4): 537-550.
- Burley, J. B. (1997). "Visual and Ecological Environmental Quality Model for Transportation Planning and Design." Transportation Research Record 1549: 54-60.
- Butler, R. (1980). "The concept of a tourist area life cycle of evolution." Canadian Geographer 24(1): 5-12.
- Butler, R. (1990). "Alternative Tourism: Pious Hope Or Trojan Horse." Journal of Travel Research 28(3): 40-45.
- Butler, R. (1991). "Tourism, Environment, and Sustainable Development." Environmental Conservation 18(3): 201-209.

- Campbell, D.J. (1998). "Toward an Analytical Framework for Land Use Change." Carbon and Nutritional Dynamics in Tropical Agro-Ecosystems. B. a. Kirschmann. Wallingford, UK, CAB International: 281-301.
- Campbell, D.J. and J.M. Olson (1991). "Framework for environment and development: the Kite." CASID Occasional Paper No. 10. Michigan State University, Center for Advanced Study of International Development.
- Castle, G. (1994). "The State of the Art in Real Estate." Business Geographics November / December.
- Census (2000). "Glossary." <http://www.census.gov/geo/www/tiger/glossary.html>
- Census, Bureau of the Census (1990). 1990 Census of Population and Housing. U. D. o. Commerce.
- Chadwick, R. (1994). Concepts, Definitions, and measures used in travel and tourism research. Travel, tourism, and hospitality research: a handbook for managers and researchers. J. R. B. R. a. C. R. Goeldner. New York, John Wiley and Sons: 65-80.
- Christaller, W. (1933). Central Places in Southern Germany (English translation), Prentice Hall Inc.
- Chubb, M. (1989). "Tourism Patterns and Determinants in the Great Lakes Region: Populations, Resources, Roads, and Perceptions." GeoJournal 19(3): 297-302.
- Chubb, M. and H. Chubb (1981). One Third Of Our Time. New York, John Wiley and Sons.
- Cohen, E. (1978). "The Impact of Tourism on the Physical Environment." Annals of Tourism Research 5(2): 215-237.
- Cooper, C., and S. Jackson (1989). "Destination Life Cycle: The Isle of Man Case Study." Annals of Tourism Research 16: 377-398.
- Coppock, J.T. (1977). Second Homes: Curse or Blessing? Oxford, Pergamon Press.
- Coppock, J.T. and D.W. Rhind (1991). The History of GIS. Geographical Information Systems: Principles and Applications. M. F. G. David J. Maguire, and David W. Rhind. Essex, Longman.
- Cracknell, B. (1967). "Accessibility to the countryside as a factor in planning for leisure." Regional Studies 1: 148.

- Crawford, D. (1994). "Using Remotely Sensed Data in Landscape Visual Quality Assessment." Landscape and Urban Planning 30: 71-81.
- Crouch, G. (1991). "Expert Computer Systems in Tourism: Emerging Possibilities." Journal of Travel Research 29(3): 3-10.
- Daniel, L. (1994). GIS Helping to Reengineer Real Estate. Earth Observation Magazine. November.
- Delorme (2000). Michigan Atlas and Gazetteer.
- Delorme (2000). Minnesota Atlas and Gazetteer.
- Delorme (2000). Wisconsin Atlas and Gazetteer.
- Fedgazette (1997). "State Roundups: Wisconsin." <http://minneapolisfed.org/pubs/fedgaz/97-01/wi.html>
- Filippakopoulou, V., and B. Nakos (1995). "Is GIS Technology the Present Solution for Creating Maps." Cartographica 32(1): 51-59.
- Fisher, P. (1997). "Editorial." International Journal of Geographic Information Science 11(1): 1-3.
- Fogarty, T.A. (2000). New wealth brings surge in two-home families. USA Today: D1-2.
- Forman, R. and M. Godron (1986). Landscape Ecology. New York, John Wiley.
- Freeman, J. and D. Skapura (1992). Neural Networks: Algorithms, Applications, and Programming Techniques. Reading, Addison-Wesley.
- Fresco, V. (1996). "CLUE-CR: an integrated multi-scale model to simulate land use change scenarios in Costa Rica." Ecological Modeling 91: 231-248.
- Gamble, P. (1988). "Tourism technology: developing effective computer systems." Tourism Management 9(4): 317-325.
- Garson, D. (1999). "Multivariate Analysis in Public Administration." <http://www2.chass.ncsu.edu/garson/pa765/logistic.htm>
- Gartner, W. (1987). "Environmental Impacts of Recreational Home Developments." Annals of Tourism Research 14: 38-57.
- Glig, A. (1985). Introduction to Rural Geography. Baltimore, Edward Arnold.

- Golley, F. (1995). "Reaching a landmark - Editor's comment." Landscape Ecology 10(1): 3.
- Goodchild, M. (1990). Spatial Information Science. Proceedings of the 4th International Symposium on Spatial Data Handling Volume 1, Zurich, Switzerland.
- Goodchild, M. (1992). "Geographical Information Science." International Journal of Geographic Information Systems 6(1): 31-45.
- Gormsen, E. (1981). The spatio-temporal development of international tourism: attempt at a centre-periphery model. La Consommation d'Espace par le Tourisme et sa Preservation. Aix-en-Provence, CHET.
- Green, H. (1995). Planning for sustainable tourism development. Tourism and the Environment: A Sustainable Relationship? C. Hunter, and Howard Green. London, Routledge: 93-121.
- Griffith, D. (1987). Spatial Autocorrelation: A Primer. Washington DC, Association of American Geographers Monograph.
- Gupta, V. (1999). SPSS for Beginners. Bloomington, 1st Books Library.
- Guttman, L. (1954). "Some necessary conditions for common factor analysis." Psychometrika 30: 179-185.
- Haggett, P, A. Cliff, and A. Frey (1977). Locational Analysis in Human Geography. Bristol, J.W. Arrowsmith Ltd.
- Haider, W. and G. Ewing (1991). "A model of tourist choices of hypothetical Caribbean destinations." Leisure Studies 12: 33-47.
- Hart, J. F. (1984). "Population Change in the Upper Lake States." Annals of the Association of American Geographers 74(2): 221-243.
- Hart, J. F. (1984). "Resort Areas In Wisconsin." Geographical Review 74: 192-217.
- Hart, J. F. (1998). The Rural Landscape. Baltimore, The Johns Hopkins University Press.
- Harvey, D. (1969). Explanation in Geography. London, Edward Arnold.
- Hawkins, J.P. and C.M. Roberts (1994). "The Growth of Coastal Tourism in the Red Sea: Present and Future Effects on Coral Reefs." Ambio 23(8): 504-508.
- Holder, J.S. (1988). "Pattern and impact of tourism on the environment of the Caribbean." Tourism Management 9(2): 119-127.

- Hunter, C., and H. Green, Ed. (1995). Tourism and the Environment: A Sustainable Relationship. London, Routledge.
- Isaaks, E. and R.M. Srivastava (1989). Applied Geostatistics. New York, Oxford University Press.
- Jarvilouma, J. (1992). "Alternative tourism and the evolution of tourist areas." Tourism Management 13(3): 118-120.
- Johnston, C.A. and R.J. Naiman (1990). "The use of a geographic information system to analyze long-term landscape alteration by beaver." Landscape Ecology 4(1): 5-19.
- Johnston, R.J., D.Gregory, and D.M. Smith (1987). The Dictionary of Human Geography Second Edition. Oxford, Blackwell.
- Keller, J.W. (2000). The Importance of Rural Development in the 21st Century - Persistence, Sustainability, and Futures. Land Use Conference of the Rocky Mountain Land Use Institute, Denver, Colorado.
- Laurini, R., and D. Thompson (1992). Fundamentals of Spatial Information Systems. London, Academic Press.
- Longino, C.F. (2001). Geographical Distribution and Migration. Handbook on Aging and the Social Sciences, 5th Edition. R. B. a. L. George. New York, Academic Press.
- Longley, P. and G. Clarke (1995). GIS for Business and Service Planning. Cambridge, GeoInformation International.
- Loomis, J.B., and R.G. Walsh (1997). Recreation Economic Decisions: Comparing Benefits and Costs. State College, Venture Publishing Inc.
- Losch, A. (1940). The Economics of Organization (English translation), Yale University Press.
- Lundgren, J. (1974). "On access to recreational lands in dynamic metropolitan hinterlands." Tourist Review 29(4): 124-131.
- Maguire, D.J. (1991). An Overview and Definition of GIS. Geographical Information Systems: Principles and Applications. M. F. G. David J. Maguire, and David W. Rhind. Essex, Longman.
- Manning, R.E. (1985). "Crowding Norms in Backcountry Settings: A Review and Synthesis." Journal of Leisure Research 17(2): 75-89.

Manson, G.A. and R.E. Groop (2000). "US Intercounty Migration in the 1990s: People and Income Move Down The Urban Hierarchy." Professional Geographer 52(3): 493-504.

Marble, D. and J. Sandhu (1994). Principles of Geographic Information Systems. Upper Arlington, Castlereigh.

Marcouiller, D., G. Green, and S. Deller (2000). "Recreational Homes and Regional Development: A Case Study from the Upper Great Lakes States." Madison, WI, University of Wisconsin Extension.

Mathieson, A., and G. Wall (1982). Tourism: Economic, Physical, and Social Impacts. Essex, Longman Scientific and Technical.

May, C., and R. Schmidley (1999). The Use of GIS Mapping to Analyze Advertising. 30th Annual Conference Proceedings: Travel and Tourism Research Association: "Navigating the Global Waters", Halifax, Nova Scotia, Canada.

May, V. (1991). "Tourism, environment, and development." Tourism Management 12(2): 112-118.

McAllister, D.M. (1977). "An Empirical Analysis of the Spatial Behavior of Urban Public Recreation Activity." Geographical Analysis 9(N2): 174-181.

McClendon, M. (1994). Multiple Regression and Causal Analysis. Itasca, F.E. Peacock Publishers Inc.

McDonough, M., and J. Parker (1995). "Status and Potential of Michigan Natural Resources." Michigan State University Extension, Ag Experiment Station Special Reports: 17.

McHarg, I. L. (1969). Design With Nature. Garden City, The Doubleday Company Inc.

McKay, L. (1987). "Tourism and changing attitudes toward land in Negril, Jamaica." Land and Development in the Caribbean. J. B. a. J. Momsen. London, MacMillan.

McKercher, B. (1992). "Tourism As A Conflicting Land Use." Annals of Tourism Research 19: 467-481.

MDNR, Minnesota Department of Natural Resources (2000). "Minnesota DNR - State Parks." http://www.dnr.state.mn.us/parks_and_recreation/state_parks/parksataglance1.html

Mercer, D. (1970). "Urban recreational hinterlands: A review and example." Professional Geographer 22(2): 74-78.

Miossec, J.M. (1976). Elements pour une Theorie de l'Espace Touristique. Les Cahiers du Tourisme C-36. Aux-en-Provence, CHET.

Monmonier, M. (1993). Mapping It Out: Expository Cartography for the Humanities. Chicago, University of Chicago Press.

National Association of Realtors (NAR) (2000). "Data Extrapolated (Second Homes / Recreational Property)." <http://nar.realtor.com/news/2000Releases/May/Data2.htm>

Naveh, Z., and A.S. Lieberman (1984). Landscape Ecology: Theory and Application. New York, Springer-Verlag.

Office of Tourism Minnesota (2000). "Explore Minnesota."
<http://www.exploreminnesota.com/>

O'Neill, R.V., J.R. Krummel, R.H. Gardner, G. Suighara, B. Jackson, D.L. DeAngelis, B.T. Milne, M.G. Turner, B. Zygmunt, S.W. Christensen, V.H. Dale, and R.L. Graham (1988). "Indices and landscape pattern." Landscape Ecology 1(3): 153-162.

Openshaw, S, and K. Clarke (1996). Developing spatial analysis functions relevant to GIS environments. Spatial Analytical Perspectives in GIS. M. Fischer, Henk J. Scholten, and David Unwin. London, Taylor and Francis.

Openshaw, S. (1998). "Towards a more computationally minded scientific human geography." Environment and Planning A 30: 317-332.

Orland, B. (1994). "Visualization Techniques for Incorporation in Forest Planning Geographic Information Systems." Landscape and Urban Planning 30: 83-97.

Page, S. J. and D. Getz, Ed. (1997). The Business of Rural Tourism: International Perspectives. London, International Thomson Business Press.

Pearce, D. (1989). Tourist Development. New York, Longman.

Peterson, G.S. and E.S. Neumann (1969). "Modeling and Predicting Human Response to the Visual Recreation Environment." Journal of Leisure Research 1(3).

Philips, A., and M. Tubridy (1994). "New Supports for Heritage Tourism in Rural Ireland." Journal of Sustainable Tourism 2(1&2): 112-129.

Piantadosi, S., S.R. Criley, S.L. Wu, A.W. Partin, P.C. Walsh, D.S. Coffey (1994). "Neural Network versus logistic regression for prediction of pathological stage or disease progression among men with clinically localized prostate cancer." Journal of Urology 151(5): 415a.

Pigram, J. J. (1980). "Environmental Implications of Tourism Development." Annals of Tourism Research 7(4): 554-582.

Pijanowski, B., T. Machemer, S.Gage, D.Long, W.Cooper, and T.Edens (1995). "A land transformation model: integration of policy, socioeconomics, and ecological succession to examine pollution patterns in watershed." Research Triangle Park, North Carolina, Environmental Protection Agency: 72.

Pijanowski, B., D. Brown, B.Shellito, S. Pithadia, and G. Manik (2000). The Land Transformation Modeling Project: Model Users Guide. Manila, Philippines, LUCC / LCUCC SE Asia Regional Researchers Workshop.

Pijanowski, B., S.H. Gage, D.T. Long, and W. Cooper (2000). A Land Transformation Model for the Saginaw Bay Watershed. Landscape Ecology: A Top-Down Approach. J. S. a. L. D. Harris. Boca Raton, Lewis Publishers: 183-199.

Pijanowski, B., B.Shellito, M. Bauer, and K. Sawaya (2001). Using GIS, Artificial Neural Networks, and Remote Sensing to Model Urban Change in the Minneapolis-St.Paul and Detroit Metropolitan Areas. ASPRS Annual Conference, St. Louis, MO.

Pijanowski, B., D.Brown, B.Shellito, and G. Manik (2001). "Using Neural Nets and GIS to Forecast Land Use Changes: A Land Transformation Model." Computers, Environment, and Urban Systems. In Press.

Plog, S.C. (1973). "Why destination areas rise and fall in popularity." Cornell H.R.A. Quarterly November: 13-16.

Poon, A. (1988). "Tourism and Information Systems." Annals of Tourism Research 15: 531-549.

Pred, A. (1966). The Spatial Dynamics of U.S. Urban Industrial Growth, MIT Press.

Preissing, J., D. Marcoullier, G. Green, S. Deller, and N.R. Sumathi (1996). Recreational Homeowners and Regional Development: A Comparison of Two Northern Wisconsin Counties. University of Wisconsin - Extension, Center for Community Economic Development.

Ragatz, R.L. (1969) "Vacation Homes: An Analysis of the Market for Seasonal-Recreational Housing." Department of Housing and Design. Ithaca. Cornell University.

Rai, S.C. and R.C. Sundriyal (1997). "Tourism and Biodiversity Conservation: The Sikkim Himalaya." Ambio 26(4): 235-242.

Roehl, W.S. and D.R. Fesenmaier (1987). "Tourism Land Use Conflict in the United States." Annals of Tourism Research 14: 471-485.

- Romeril, M. (1989). "Tourism and the environment - accord or discord?" Tourism Management 10(3): 204-208.
- Ronan, C. (1999). "Picturesque Peninsula Enjoys Booming Real Estate Business." http://realitytimes.com/rtnews/rtcpages/19990125_doorcounty.htm
- Ryan, C. (1991). Recreational Tourism: A Social Science Perspective. London, Routledge.
- Sanderson, J., and L.D. Harris, Ed. (2000). Landscape Ecology: A Top Down Approach. Boca Raton, Lewis Publishers.
- Sarle, W. (2001). "comp.ai.neural-nets FAQ." <http://www.faqs.org/faqs/ai-faq/neural-nets/part3/preamble.html>
- Scheberle, D. and R. Pagel (1999). Developing an efficient and effective wetland protection program in Door County. University of Wisconsin - Green Bay, Department of Public and Environmental Affairs.
- Sclove, S. (2000). "Logistic Regression Notes." http://www.uic.edu/classes/idsc/ids470/jw4e/nts11_8.htm
- Shafer, E.L., J.F. Hamilton, and EA Schmidt (1969). "Natural Landscape Preferences: A Predictive Model." Journal of Leisure Research 1(1): 1-19.
- Shafer, EL and M Tooby (1973). "Landscape Preference: An International Replication." Journal of Leisure Research 5(3): 60-65.
- Sheldon, P. (1993). "Destination Information Systems." Annals of Tourism Research 20: 633-649.
- Skapura, D. (1996). Building neural networks. New York, ACM Press.
- Skole, D. and S.H. Gage (2001). Creating a Research Enterprise on Land Use and Land Cover Change. Michigan State University.
- Smith, C. and P. Jenner (1989). "Tourism and the Environment." Travel and Tourist Analyst 5: 68-86.
- Spotts, D., Ed. (1991). Travel and Tourism In Michigan: A Statistical Profile 2nd Edition. East Lansing, Travel, Tourism, and Recreation Resource Center, Michigan State University.
- Statcon (2000). "Statcon: Statistische Software und Dienstleistungen." <http://www.statcon.de/vertrieb/neuralconnection/ncqa.htm>

Statsoft (2000). "Online Statistics Textbook."
<http://www.statsoftinc.com/textbook/stnonlin.html>

Steinitz, C., P. Parker, and L. Jordan (1976). "Hand-drawn Overlays: Their History and Prospective Uses." Landscape Architecture: 444-455.

Steinitz, C. (1990). "Toward a Sustainable Landscape with High Visual Preference and High Ecological Integrity: The Loop Road in Acadia National Park, USA." Landscape and Urban Planning 19: 213-250.

Stewart, S. (1994) "The seasonal home location decision process : toward a dynamic model." Department of Park and Recreation Resources. East Lansing. Michigan State University.

Stewart, S. (2000). Amenity Migration. Trends 2000, Lansing, MI.

Stopher, P. and G. Ergun (1982). "The Effect of Location on Demand for Urban Recreation Trips." Transportation Research 16A(1): 25-34.

Strapp, J.D. (1988). "The Resort Cycle and Second Homes." Annals of Tourism Research 15: 504-516.

Stynes, D, and Safronoff, D. (1982). 1980 Michigan Recreational Boating Survey. Ann Arbor, Michigan Sea Grant Publishing.

Stynes, D. and George Peterson (1984). "A Review of Logit Models with Implications for Modeling Recreation Choices." Journal of Leisure Research 16(4): 295-310.

Stynes, D., J. Zheng, and S. Stewart (1995). Seasonal Homes and Natural Resources: Patterns of Use and Impact in Michigan, Forest Service, US Department of Agriculture.

Stynes, D. (1996). Recreation Activity and Tourism Spending in the Lake States. East Lansing, MI, Michigan State University.

Talhelm, D.R. (1991). "The Community Options Model." Travel and Tourism Research Association, Censtates Chapter: 12th Annual Conference, Marc Plaza Hotel: Milwaukee, Wisconsin.

Taylor, P.J. (1977). Quantitative Methods in Geography. Tyne, University of Newcastle.

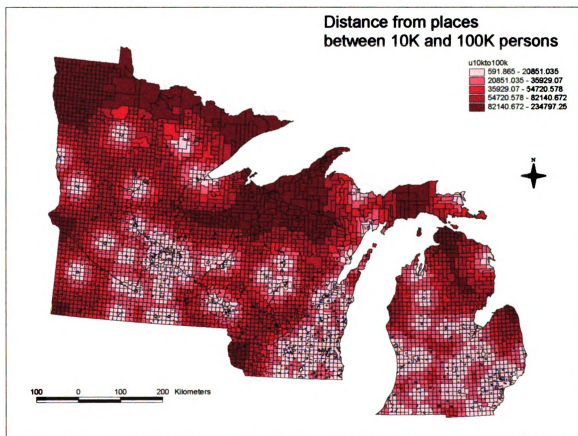
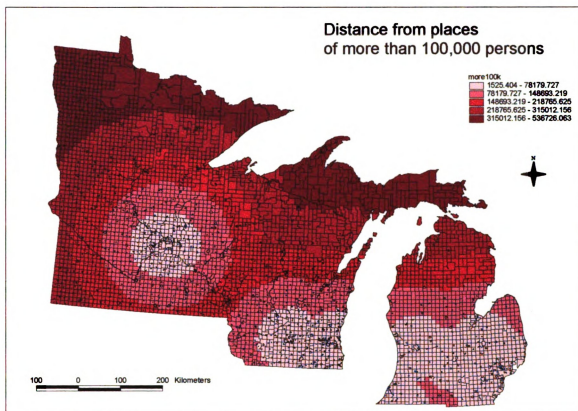
Theobald, D.M., and Hobbs, N.T. (1998). "Forecasting rural land use change: a comparison of regression and spatial transition models." Geographical and Environmental Modelling 2(1): 65-82.

Thomas, H. (2000). "Landscape Ecology as a geographical approach to the study of the environment." <http://humanities.newport.ac.uk/landsc~1.html>

- Thurot, J.M. (1973) "Le Tourisme Tropical Balnearie: le Modele Caraibe et ses Extensions." Aix-en-Provence. Centre d'Etudes du Tourisme.
- Tombaugh, L.W. (1970). "Factors influencing vacation - home location." Journal of Leisure Research 2: 54-63.
- Travel Association of America (1999). Impact of Travel on State Economies 1997. Washington DC, Research Department of the Travel Industry of America.
- Troll, C.V. (1939). Luftbildplan und ökologische Bodenforschung (Aerial photography and ecological studies of the Earth). Berlin, Zeitschrift der Gesellschaft für Erdkunde.
- Turner, M.G., R.H. Gardner, V. Dale, and R.V. O'Neill (1989). "Predicting the spread of disturbance across heterogeneous landscapes." Oikos 55: 121-129.
- Turner, M.G. (1990). "Spatial and temporal analysis of landscape patterns." Landscape Ecology 4(1): 21-30.
- Urban, D.L., R.V. O'Neill, and H. Shugart Jr. (1987). "Landscape Ecology." Bioscience 37: 119-127.
- US Travel Data Center (1998). 1997-98 Survey of State Tourism Offices. Washington DC, Travel Industry Association of America.
- USDA and Forest Service Eastern Region (1975). Guide for Managing the National Forests in the Lake States.
- USDA and Forest Service Eastern Region (1986). Land and Resource Management Plan: Huron-Manistee National Forests.
- Van Doran, C., G.B. Priddle, J.E. Lewis (1979). Land & Leisure: Concepts and Methods in Outdoor Recreation. Chicago, Maaroufa Press Inc.
- Vasievich, M. (1999). "Here Comes The Neighborhood: A New Gold Rush and Eleven Other Trends Affecting The Midwest." NC News, North Central Forest Experimental Station. August / September.
- Vink, A.P.A. (1983). Landscape Ecology and Land Use. London, Longman.
- Vlitros-Rowe, I. (1992). "Destination Databases and Management Systems." EIU Travel and Tourism Analyst 5: 84-109.
- von Thunen, J.H. (1826). Der Isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie. Hamberg.

- Wahab, S. and J. Pigram, Ed. (1997). Tourism, development and growth: the challenge of sustainability. New York, Routledge.
- Wall, G. (1979). Recreational Land Use in Southern Ontario. Waterloo, Department of Geography.
- Waters, J. (1990). Travel Industry World Yearbook: The Big Picture - 1990. New York, Child and Waters, Inc.
- WDNR, Wisconsin Department of Natural Resources (2000). "State parks and recreation." <http://www.dnr.state.wi.us/org/land/parks/>
- Wheatcroft, S. (1991). "Airlines, tourism, and the environment." Tourism Management 12(2): 119-124.
- Wheeler, B. (1991). "Tourism's Troubled Times." Tourism Management 12(2): 91-96.
- Wilder, M.G. (1985). "Site and Situation Determinants of Land Use Change: An Empirical Example." Economic Geography 61: 332-344.
- Wilkinson, P.F. (1973). "The Use of Models in Predicting the Consumption of Outdoor Recreation." Journal of Leisure Research 5(3): 34-48.
- Wilks, B.E., K.F. Backman, J. Allen, and D. Van Blaircom (1993). "Geographic Information Systems (GIS): A Tool For Marketing, Managing, and Planning Municipal Park Systems." Journal of Park and Recreation Administration 11(1): 9-23.
- Williams, D.R. and N. McIntyre (2000). Where Heart and Home Reside: Changing Constructions of Place and Identity. Trends 2000, Lansing, MI.
- Wisconsin.com (2000). "Facts About Wisconsin."
- Witt, S.F. and L. Moutinho (1994). Tourism Marketing and Management Handbook. New York, Prentice Hall.
- Wolfe, R (1972). "The Inertia Model." Journal of Leisure Research 4(Winter): 73-76.
- Wright, D., M. Goodchild, and J.D. Proctor (1997). "GIS: Tool or Science." Annals of the Association of American Geographers 87(2): 346-362.
- WTO (World Tourism Organization) (1995). "Steady recovery in world tourism in 1994." WTO News 1(1).
- Zonneveld, I. (1995). Land Ecology. Amsterdam, SPB Academic Publishing.

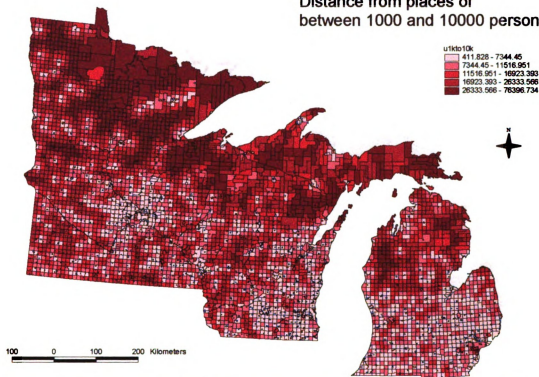
APPENDIX A
INDEPENDENT VARIABLE MAPS



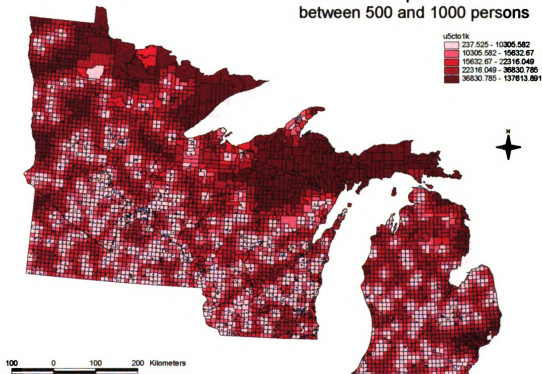
000000



Distance from places of
between 1000 and 10000 persons



Distance from places
between 500 and 1000 persons



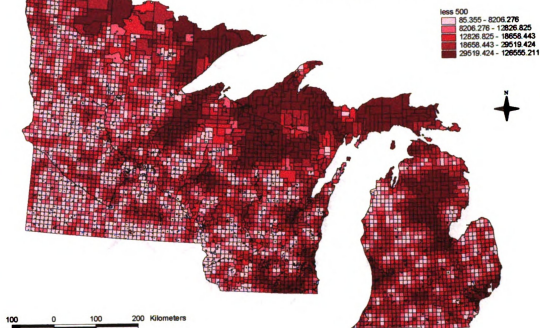
Below



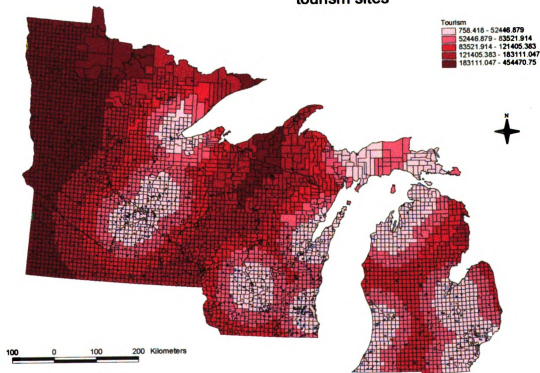
Below

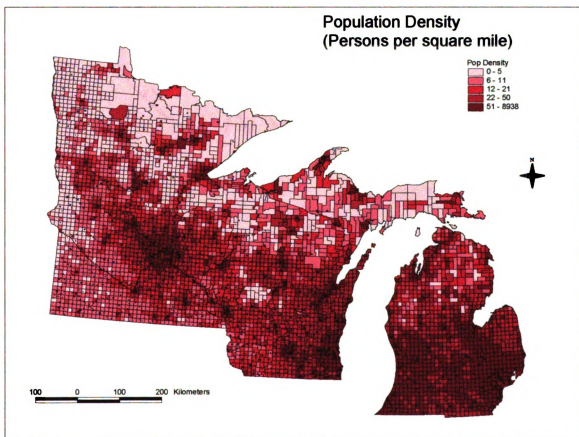
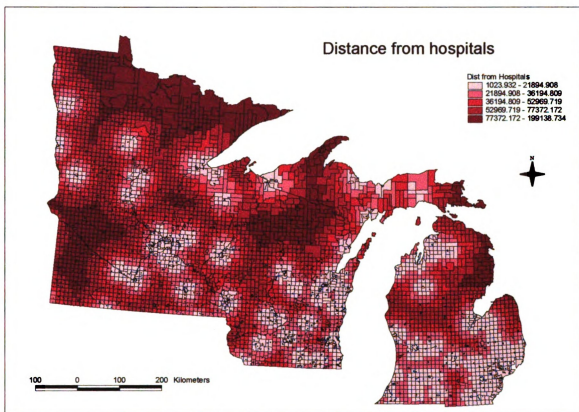


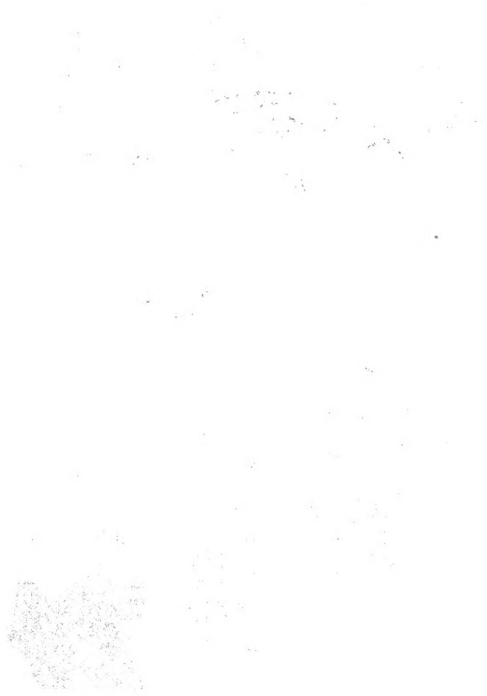
Distance from places
of less than 500 persons

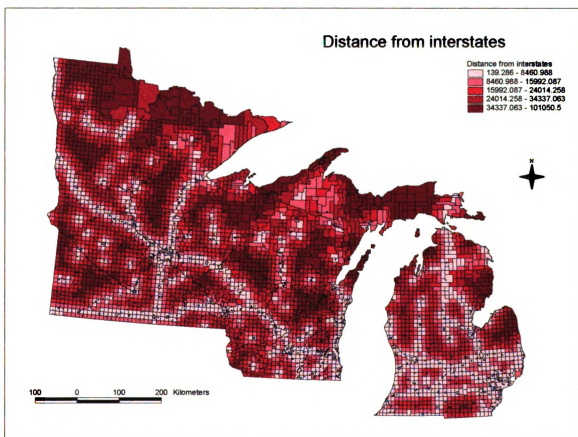
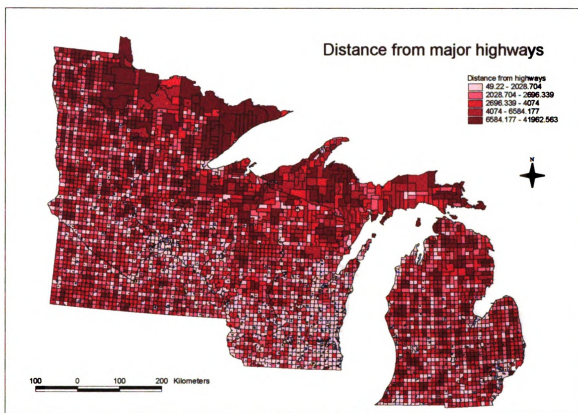


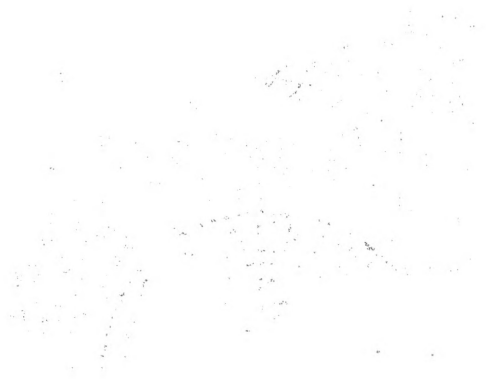
Distance from selected
tourism sites

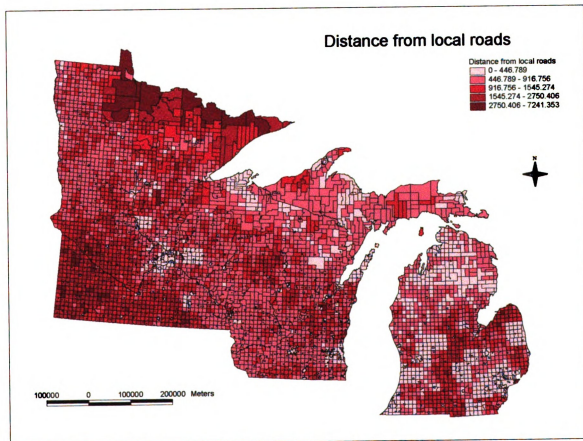
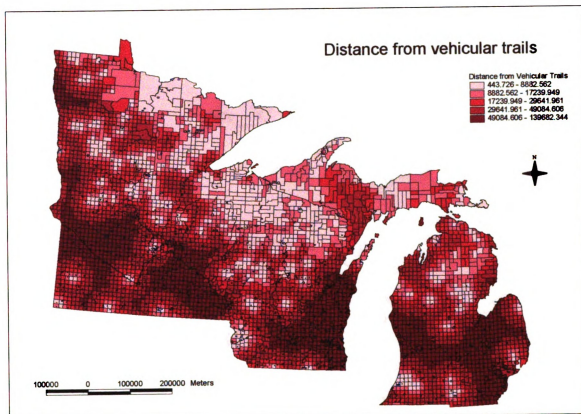


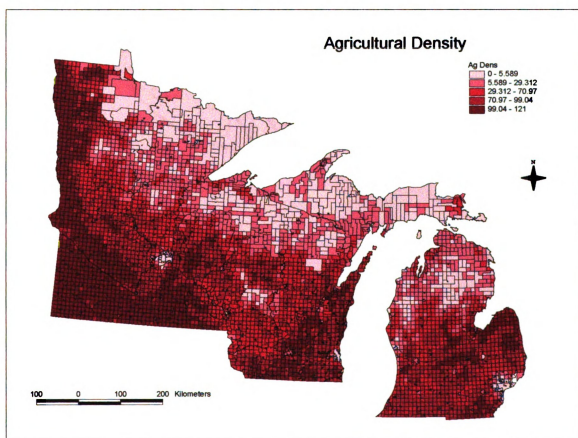
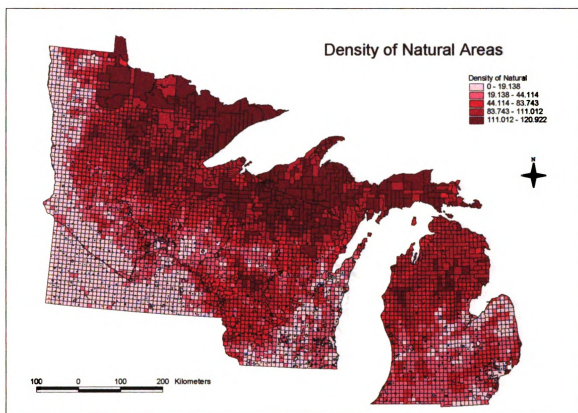












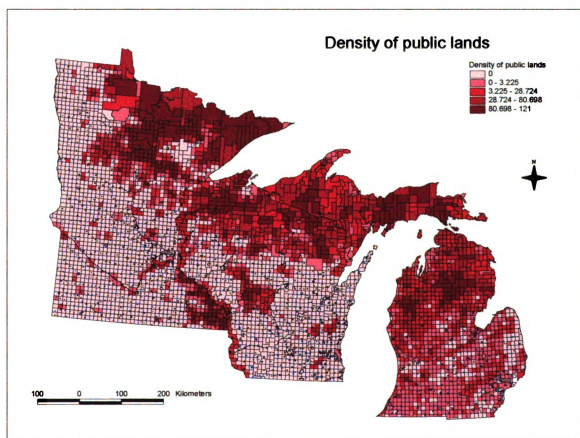
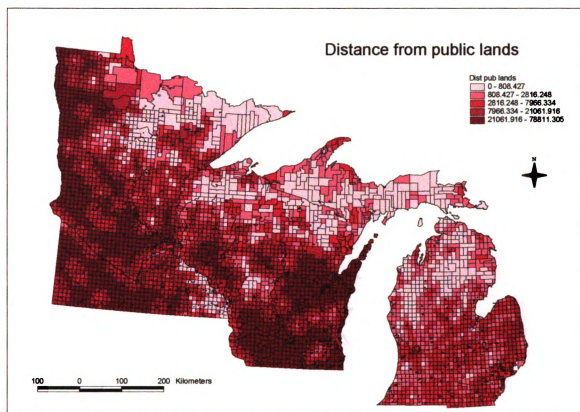
Dent



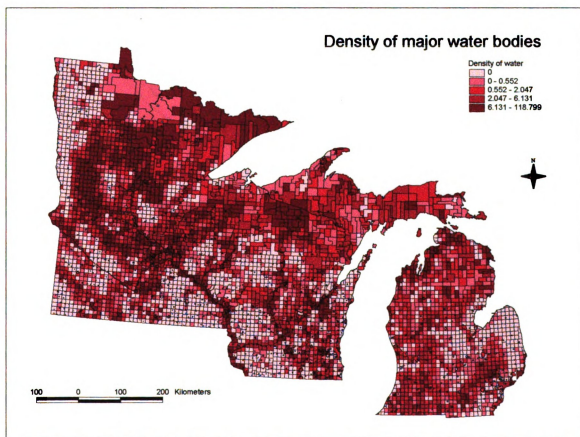
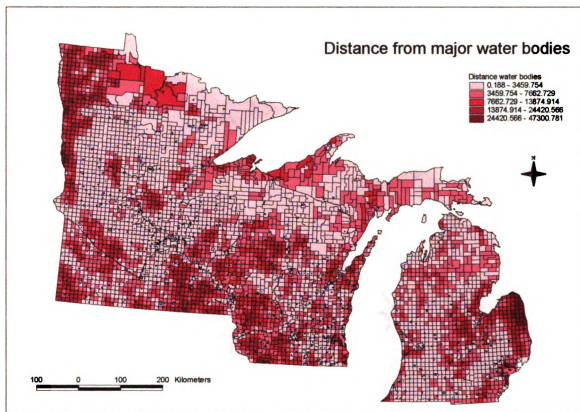
100

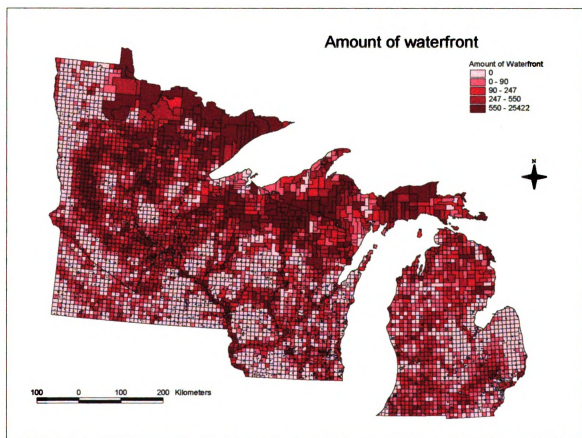
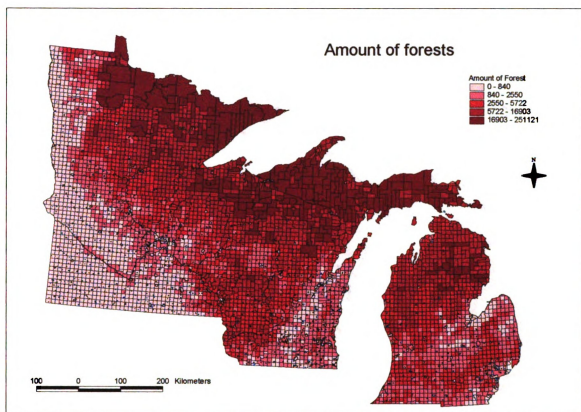
100

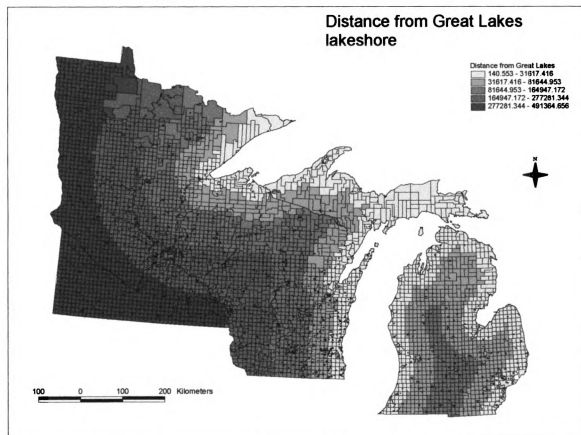
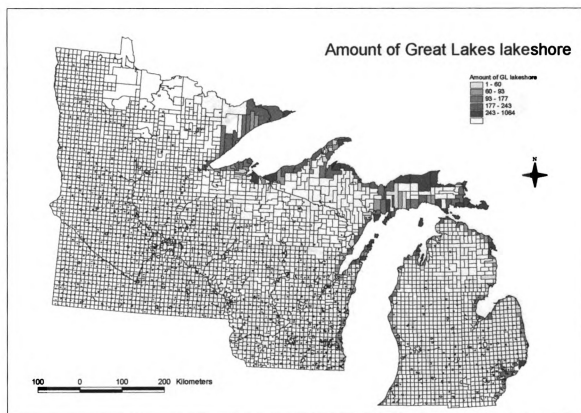














MICHIGAN STATE LIBRARIES



3 1293 02177 8273