Ryan R. Jensen
Jay D. Gatrell
Daniel D. McLean
Editors

Geo-Spatial Technologies in Urban Environments

Policy, Practice, and Pixels

Second Edition

Springer

R. R. Jensen
J. D. Gatrell
D. McLean
Geo-Spatial Technologies in Urban Environments
Policy, Practice, and Pixels
Second Edition

R. R. Jensen
J. D. Gatrell
D. McLean
Editors

Geo-Spatial Technologies in Urban Environments

Policy, Practice, and Pixels

Second Edition

With 39 Figures

Springer

Dr. Ryan R. Jensen
Dr. Jay D. Gatrell
Dr. Daniel McLean

Indiana State University
Department of Geography,
Geology and Anthropology
Terre Haute IN 47809
USA

E-mail: Ryan R. Jensen [rjensen@isugw.indstate.edu]

Library of Congress Control Number: 2006939007

ISBN-13 978-3-540-69416-8 Springer Berlin Heidelberg New York
ISBN (1st Edition) 3-540-22263-4 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media
springer.com

Cover design: E. Kirchner, Heidelberg
Production: Almas Schimmel
Typesetting: camera-ready by the editors

Printed on acid-free paper 30/3141/as 5 4 3 2 1 0

Foreword

I was introduced to cities as "ecosystems" by the late Professor Forest Stearns (University of Wisconsin, Milwaukee) who was an early pioneer in the studies of the urban ecology in the 1970s. I still recall the various terms: "Urban ecosystem", "Urban zones" Urban corridors" Megacities" or "Megacities' complexes" and "Megalopolis" used by various discipline experts as they grappled with the complex of all the terrestrial habitats-the city. Cities have been humanity's habitat since ancient times and one can find references to the cities even in biblical writings and other ancient texts from many parts of the world. So what is different now? It is the rate of global urbanization that has brought urban systems and urban environments into focus once again. This has captured our attention in the past several months. For example, the British Broadcasting Corporation News (BBC news) devoted a series of highly educational programs titled "Urban planet" in July 2006. I was impressed with the breadth and the depth of urban issues discussed in the series and the fact that many academic and government experts were featured to provide an assessment of the current state of global urbanization. The United Nations reports that increased urbanization has created a range of serious issues, including access to clean water, sanitation, shelter, urban poverty, HIV/AIDS and problems with urban governance, not to mention the issues related to urban environments.
What is significant is the report that sometime in 2007 and somewhere on the planet, someone migrating from a rural area to a city will tip the global urban/rural balance. United Nations agencies forecasts that majority of human population will live in urban settings very soon. These estimates indicate that about 180,000 people are being added to the urban population every day. Can the world's urban infrastructure absorb the equivalent of the population of two Toykos each year? As a news reporter put it "Homo sapiens" are fast becoming "Homo urbanis"!

In historical literature, there are fascinating examples of the use of "early forms of geospatial technologies" such as photographs of cities acquired from the hot air balloons, from cameras suspended from kites, and the very first aerial photograph of a city from Wilbur Wright's plane in 1909 and many early maps from geographic field surveys.

As we rapidly urbanize as a society, we need to usher in new and innovative technologies, social and policy initiatives, and political and economic instruments to improve the quality of life in urban environments. A set of rapidly maturing technologies termed the "Geospatial" technologies are becoming a vital part of this new urbanized world. Geospatial technologies have already had a major impact of many aspects of how we analyze, model and manage the urban systems and their environments. From population census using remote sensing and satellite technologies to modeling the impact of a contagious disease using GIS; we have come to realize the promise and potential of these rapidly maturing technologies. Despite many strides in the application of geospatial technologies for urban applications; this science is still in its infancy.

This book edited by Ryan Jensen, Jay Gatrell, and Daniel McLean is, therefore a much awaited contribution to the literature. The editors have assembled an extraordinarily talented contributors for the book. It is a pioneering effort in many ways for it uniquely brings together the physical and social aspects of urban environments via geospatial technology applications.

The book chapters address some difficult and important aspects of application of geospatial technologies in urban environments. The authors have discussed cutting-edge geospatial technologies for issues ranging from urban change detection, estimating urban population to urban health and heat wave and urban child obesity. Chapter on role of geospatial technologies in environmental justice is a good example how social and physical aspects of urban landscape can be brought together using these technologies.
Also, the issues unique to the semi-rural counties have been addressed in the book.

I believe this timely publication will spur further development of geospatial technology applications for urban systems. Many challenges remain and need to be addressed if we are to fully apply these technologies for urban environments. I am confident that this book will inspire a new generation of researchers who will apply geospatial technologies to urban systems before the impact of rapid global urbanization becomes a crisis for our societies in both the developed and the developing world.

Kamlesh P. Lulla, Ph.D; Ph.D.
NASA Johnson Space Center
Houston, Texas 77058

Dedications

To my family for their inspiration and support – RRJ

To my favorite editor, s., and our children, f. m. and e. – JDG

To my wife and soul mate for your support, compassion, and challenge to be better – DDM

Acknowledgments

The editors wish to thank the excellent chapter authors for their contributions to the second volume of this book. We also acknowledge the hard work and assistance of Ms. Barbara McNeill with the final preparation of the manuscript. Finally, we would like to thank the entire Springer team for their continued help and support of our efforts.

Contents

1 Applying Geospatial Technologies in Urban Environments 1
Ryan R. Jensen, Jay D. Gatrell, and Daniel D. McLean
1.1 About this book 1
1.2 Chapters 3
References 4

2 Remote Sensing Change Detection in Urban Environments 7
John R. Jensen and Jungho Im
2.1 Introduction 7
2.2 Remote Sensing Change Detection Process 8
2.2.1 Digital Frame Camera Remote Sensing 10
2.2.2 LIDAR Remote Sensing 10
2.3 Case Study 1 14
2.4 Case Study 2 22
2.5 Conclusion 30
References 30

3 Assesement of Risk in Urban Environments Using Geo-Spatal Analysis 33
James D. Hipple
3.1 Defining Risk 33
3.2 Methods for Hazard Identification and Risk Assessment 34
3.2.1 Photogrammetry 34
3.2.2 Remote Sensing 35
3.2.3 Results of the Analysis 41
3.2.4 How Much Development is Occurring? 42
3.2.5 Case Study Conlclusions 43
3.3 Summary 44
References 45

4 Intraurban Population Estimation Using Remotely Sensed Imagery 47
Perry J. Hardin, Mark W. Jackson, and J. Matthew Shumway

4.1 Traditional Approaches to Population Estimation 48
4.2 Population Estimation Using Remote Sensing 51
4.2.1 Dwelling Identification 52
4.2.2 Landtype Surrogates 60
4.2.3 Pixel-Based Estimation 66
4.3 Case Study 75
4.4 Concluding Comments 87
References 89

5 Using Satellite Data to Estimate Urban Leaf Area Index 93
Ryan R. Jensen and Perry J. Hardin
5.1 Introduction 93
5.1.1 Urban Remote Sensing 94
5.2 Data and Methods 95
5.2.1 Study Area 95
5.2.2 LAI Field Measurements 95
5.2.3 Satellite Sensor Data 97
5.2.4 Estimating LAI Using Regression 97
5.2.5 Estimating LAI Using a Back-Propagation Feed-Forward Network 98
5.3 Results and Discussion 100
5.3.1 Regression Results 100
5.3.2 Artificial Neural Network Results 101
5.5 Conclusion 102
References 105

6 Public Participation Geographic Information Systems as Surveillance Tools in Urban Health 109
Daniel P. Johnson
6.1 Introduction 109
6.1.1 Urban Development and Medical Geography 110
6.2 Health and the Built Environment 113
6.3 Public Participation Geographic Information Systems (PPGIS) 114
6.4 Heat Related Deaths 115
6.5 PPGIS for Heat Wave Surveillance 116
6.6 Summary 118
References 118

7 Examining Urban Environment Correlates of Childhood Physical Activity and Walkability Perception with GIS and Remote Sensing 121

Gilbert C. Liu, James T. Colbert, Jeffrey S. Wilson, Ikuho Yamada, and Shawn C. Hoch
7.1 Introduction 121
7.2 Data and Methods 123
7.3 Results 131
7.4 Discussion 134
7.5 Conclusions 137
References 137

8 Mapping, Measuring, and Modeling Urban Growth 141
Perry J. Hardin, Mark W. Jackson, and Samuel M. Otterstrom
8.1 Urban Growth and Planning Policy 142
8.2 Mapping Urban Growth 145
8.2.1 Change Detection 145
8.2.2 Comparative Studies 152
8.3 Quantification and Modeling of Urban Growth 162
8.3.1 Urban Growth Measurement with Landscape Metrics 162
8.3.2 Urban Growth Modeling with Cellular Automata 164
8.4 Future Research Directions 165
References 167

9 Deer-Vehicle Collisions Along the Suburban-Urban Fringe 177
Rusty A. Gonser and J. Scott Horn
9.1 Introduction 177
9.1.1 Background 178
9.1.2 Mitigation: Reducing the Number of DVCs 182
9.1.3 Spatial Autocorrelation and Likelihood Maps 182
9.2 Methods 184
9.2.1 Location 184
9.2.2 Data 186
9.2.3 Data Analysis 186
9.3 Results 191
9.4 Discussion 192
9.5 Conclusions and Broader Impacts 194
References 194

10 Scale and Spatial Autocorrelation from a Remote Sensing Perspective 197
J. Scott Spiker and Timothy A. Warner
10.1 Introduction 197

10.2 Measuring Spatial Autocorrelation 199
10.2.1 Global Measures of Spatial Autocorrelation 199
10.2.2 Local Measures of Spatial Autocorrelation 201
10.3 Data and Processing 202
10.4 Discussion 205
10.5 Concluding Remarks 211
References 212

11 The Spatial Imperatives of Environmental Justice 215
Trevor Fuller, Jay D. Gatrell, and E. LaFary
11.1 Study Area 215
11.2 Placing and Scaling Environmental Justice 217
11.3 GIScience: GIS, RS and GWR 218
11.4 Data and Methods 219
11.4.1 Environmental Data Sets 219
11.4.2 NDVI 221
11.4.3 Socioeconomic/Demographic Characteristics 222
11.5 Methods 222
11.5.1 Interaction Terms 223
11.6 The Models 223
11.7 Results 224
11.8 Discussion 229
References 230

12 Geotechnologies, Public Policy, and Practical Applications 233
Jay D. Gatrell, Ryan R. Jensen, and Daniel D. McLean
12.1 Technolgies and Methods 233
12.2 Risk 234
12.3 Planning 234

Subject Index 237

About the Contributors

James (Jay) T. Colbert (M.S., Indiana University – Purdue University, Indianapolis) is a GIS Analyst at The Polis Center at Indiana University Purdue University Indianapolis. His research interests include Geographic Information Science, built environment and physical activity.

Trevor K. Fuller (B.S., Purdue University) is an M.A. student in Geography at Indiana State University. His primary research interests are the use of remote sensing to assess environmental quality and examination of race, class, and economics in urban areas.

Rusty A. Gonser (Ph.D., University at Albany-State University of New York) is an assistant professor in the Department of Life Sciences at Indiana State University. His research interests include behavioral ecology, genetic variation in ubiquitous species, and spatial components of Wildlife-Vehicle Collisions and their impact on genetic structure of a population.

Perry J. Hardin (Ph.D., University of Utah) is an Associate Professor of Geography at Brigham Young University. His primary research interests are computational statistics and the estimation of urban biophysical parameters from remotely sensed data.

James D. Hipple (Ph.D., University of Utah) is currently with the United States Department of Agriculture (USDA) Risk Management Agency (RMA) serving as RMA's Deputy-CIO for Geospatial Policy & Planning and Remote Sensing & GIS Advisor.

Shawn Hoch (B.S. Indiana University) is a Graduate Research Assistant in the Department of Geography at Indiana University – Purdue University, Indianapolis where he is pursuing an MS in Geographic Information Science. His research interests include GIS applications in public health, urban planning, and linguistics

J. Scott Horn (M.A., Indiana State University) is a GIS Analyst with the Environmental sciences program at the Utah Geological Survey, Utah Department of Natural Resources. His research interests include biogeography, ground water, and human/ wildlife interaction.

Jungho Im (Ph.D., University of South Carolina) is a Post-doctoral Research Scientist at the Department of Geography, University of South Carolina. His primary research interests are model (algorithm) development for estimating urban/suburban biophysical parameters from remote sensing data, remote sensing and GIS-based modeling for environmental processes, and remote sensing data fusion.

Mark W. Jackson (Ph.D., University of South Carolina) is an Assistant Professor of Geography at Brigham Young University. His research interests center on the societal causes and environmental effects of landscape change and methods of monitoring and analyzing these changes using remotely sensed imagery and GIS.

John R. Jensen (Ph.D., University of California, Los Angeles) is a Carolina Distinguished Professor of Geography at the University of South Carolina. His interests include remote sensing of the environment to extract vegetation and water-related biophysical data and urban/suburban infrastructure information.

Daniel P. Johnson (MS, Indiana University) is an Assistant Professor of Urban Affairs and Geography at Wright State University. His primary research interests are medical geography and the relationship between the built environment and human health.

Eric W. LaFary (MA, Indiana State University) is currently undertaking research toward his Ph.D. at The University of Auckland. He is utilizing a trans-disciplinary political ecology framework incorporating applied remote sensing and statistical modeling to examine the perceptions and politics surrounding marine and estuarine ecosystems.

Gilbert C. Liu (M.D. University of Mississippi; M.S. University of North Carolina, Chapel Hill) is an Assistant Professor in the Indiana Children's Health Services Research Center Indiana University School of Medicine. His research interests include environmental factors associated with obesity risk.

Samuel M. Otterstrom (Ph.D., Louisiana State University) is an Associate Professor of Geography at Brigham Young University. His research focuses on historical and contemporary urban development and population geography issues in the United States, Central America, and Europe.

J. Matthew Shumway (Ph.D., Indiana University) is a Professor and Chair of Geography at Brigham Young University. His primary research interests are spatial demography and the development of the rural West.

J. Scott Spiker (Ph.D. West Virginia University) is a Lecturer in Geography at the University of Wisconsin-Parkside. His primary research interests are spatial analytical methods and spatial structure in geographic data.

Jeffrey S. Wilson (Ph.D., Indiana State University) is an Associate Professor and Chair of the Department of Geography at Indiana University – Purdue University, Indianapolis. His research interests include environmental remote sensing, Geographic Information Science, human health, and the environment.

Ikuho Yamada (Ph.D., University at Buffalo, the State University of New York) is an Assistant Professor in the Department of Geography at the University of Utah. His research interests include spatial statistics and space-time analysis, Geographic Information Science and Systems, transportation, and public health.

1 Applying Geospatial Technologies in Urban Environments

Ryan R. Jensen, Department of Geography, Geology & Anthropology, Indiana State University, Terre Haute, IN

Jay D. Gatrell, Department of Geography, Geology & Anthropology, Indiana State University, Terre Haute, IN

Daniel D. McLean, Department of Recreation and Sport Management, Indiana State University, Terre Haute, IN

The world has entered the urban millennium. Nearly half the world's people are now city dwellers (Annan, 2001). The city is everywhere and everything (Amin and Thrift 2002). ... Towns and cities are (the) focus of today's social and ecological problems. Urban activities are the foundation of economic prosperity. Cities are strategic places (Annan, 2001). Cities compete. Cities are going through a renaissance (Amin and Thrift, 2002).

European Science Foundation, Urban Science, 2006

1.1 About this book

As the epigraph above indicates, cities have become an important part of human existence, and they represent and support most human activity. Urban areas have been the primary locations for social movements, intellectual discoveries, and the rise and fall of nations and civilizations (Greene and Pick, 2006). It is projected that cities will only become more important as societies continue to go through the demographic transformation process. Geospatial technologies will probably play a critical role throughout this because of their ability to examine things synoptically, help manage existing infrastructure and services, and predict and model future growth.

According to the United Nations Information Service (2004), 48 per cent of the world's population lived in urban areas in 2003, and urban population is projected to exceed the 50 per cent mark by 2007. Additionally, the proportion of the world's population that is urban is expected to rise to 61 per cent by 2030. Conversely, rural population is anticipated to decline slightly from 3.3 billion in 2003 to 3.2 billion in 2030. Further, during 2000-2030, the world's urban population is projected to grow at an average annual rate of 1.8 per cent, nearly double the rate expected for the total population of the world (almost 1 per cent per year). At this rate of growth, the world's urban population will double in 38 years.

Most of the urban growth will probably occur in lesser-developed regions where the percentage of urban population is lower (42% in 2003; expected to rise to 57% by 2030). This urban growth trend in less developed regions is forecast to average 2.3 per cent per year during 2000-2030. In fact, almost all the growth of the world's total population between 2000 and 2030 is expected to be absorbed by the urban areas of less developed regions, and by 2017, the number of urban dwellers will equal the number of rural dwellers in the less developed regions (United Nations Information Service, 2004).

In contrast, the urban population of more developed regions is expected to increase very slowly, from 0.9 billion in 2003 to 1 billion in 2030, because the process of urbanization is already advanced in these regions, where 74 per cent of the population lived in 2003. The percentage of the population in more developed regions living in urban areas is expected to increase to 82 per cent by 2030 (United Nations Information Service, 2004).

As these figures and projections suggest, people will continue to migrate to urban areas – particularly in developing countries. The ability to examine and mitigate the potential negative impacts of this migration is very important today and will be even more important tomorrow. Also, the ability to adequately prepare for this migration probably will rest on the shoulders of those urban scientists currently studying the urban environment. This book presents many ways that the urban environment can be studied using geo-spatial data and techniques.

1.2 Chapters

There are many ways to classify the chapters of this book including geospatial techniques used, size of urban area studied (population, area, etc.), data sets used, spatial resolution or scale of the data and so on. Because of this breadth, classifying the chapters into specific groups was very difficult. Simply put, this book provides many examples of cutting-edge geospatial technology research in urban areas. Many chapters demonstrate the potential role of geospatial technologies in examining, mapping, and modeling urban problems. Specifically, chapter 2 describes how geospatial technologies can be used to study urban change using LIDAR and digital frame camera data. Chapter 3 shows how these technologies help to assess risk in urban areas using two case studies. Chapters 4 and 8 provide reviews the role that geospatial technologies have in measuring and modeling urban population and growth, respectively. These chapters also provide case studies to describe the concepts that are discussed. Those who wish to see how satellite remote sensing data can be used to quantify the urban forest using Artificial Neural Networks should read chapter 5. Chapter 6 describes the role of Public Participation GIS to study urban health. Chapter 7 describes how the urban environment affects childhood physical activity (and corresponding obesity). The spatial relationship of deer-vechicle collisions along the suburban fringe is presented in chapter 9, and the role that scale plays in spatial autocorrelation studies is described in chapter 10. Finally, a spatial perspective of environmental justice is presented in Chapter 11.

Table 1. . Summary of substantive chapters in this book.

	Author(s)	**Subject**
2	J. Jensen et al.	Urban change detection with digital frame and LIDAR data
3	Lawrence et al.	Geo-spatial technologies to study risk
4	Hardin et al.	Estimating urban population
5	Jensen and Hardin	Measuring urban forest canopy with remote sensing data
6	Johnson	Public Participation GIS's role in studying urban health and heat waves
7	Liu et al.	Urban physical activity and childhood obesity
8	Hardin et al.	Mapping, measuring, and modeling urban growth
9	Gonser and Horn	Deer vehicle collisions along the suburban fringe

10	Spiker and Warner	Scale considerations and spatial autocorrelation
11	Fuller et al.	Spatial imperatives of environmental justice

The authors and editors hope that the applications described in this book will serve as an impetus to better understand the complex urban environment. Indeed, the future will probably present many challenges in urban areas. These challenges may be centered on such diverse issues as environmental justice, urban quality of life, effective planning, and many others. As will be shown in this book, geospatial technologies are uniquely suited to study these and many other urban problems.

This book will help anyone concerned about the urban environment to learn about additional geospatial data and techniques to study the changing dynamics in urban areas. With so much policy discourse and concern given to many other organisms and the environments in which they live, we hope that more discourse and concern will be aimed at humans and the human environment. Further, we hope that the ideas, applications, methods, and data presented in this book will enable planners, landscape architects, urban foresters, GIS and remote sensing specialists, and many others to improve quality of life in the urban environment. Finally, we hope that the studies and methods contained within this book will be used as a point of reference for those who might imagine and re-imagine the range of potential geo-technical applications to assist urban decision making and promote the overall sustainability of social and physical systems.

As the opening epigraph of this chapter suggests, cities will continue to become very important throughout the world. Our ability to model, map, and predict changes in the urban environment will be very important as humanity becomes evermore urbanized.

References

Amin, A. and Thrift, N. 2002. Cities. Reimagining the Urban. Polity.

Annan, K.A., Secretary-General, United Nations. Foreword. Cities in Globalizing World. Global Report on Human Settlements 2001. Earthscan. London & Sterling.

European Science Foundation, Urban Science. http://www.esf.org/esf_article.php?activity=8&article=285&domain=5.

Greene, R.P. and J.B. Pick. 2006. Exploring the Urban Community – a GIS Approach. Prentice Hall, Upper Saddle River, New Jersey. 495 pp.

United Nations Information Service. 2004. "UN Report Says World Urban Population of 3 Billion Today Expected to Reach 5 Billion by 2030." http://www.unis.unvienna.org/unis/pressrels/2004/pop899.html

2 Remote Sensing Change Detection in Urban Environments

John R. Jensen, Department of Geography, University of South Carolina, Columbia, SC

Jungho Im, Department of Geography, University of South Carolina, Columbia, SC

2.1 Introduction

Timely and accurate change information in the urban environment is essential for successful planning and management. The change detection may range from 1) monitoring general land cover/land use found in multiple dates of imagery, to 2) anomaly (e.g., subsidence) detection on hazardous waste sites. Remote sensing approaches to change detection have been widely used due to its cost-effectiveness, extensibility, and temporal frequency. Since the advent of high-spatial resolution satellite imagery, it has become increasing popular to detect, analyze, and monitor detailed changes such as new buildings, roads, and even patios in the urban environment. Basically, there are two types of change detection methods: 1) detection of the change using various image enhancement methods, and 2) extraction of detailed types of land-cover change based on the use of classification techniques (Chan et al. 2001; Jensen 2005)

Traditional remote sensing change detection techniques, which are generally applicable to coarse spatial resolution optical imagery, include image algebra multi-band differencing (Coppin and Bauer 1996), image transformation such as principal components analysis (Collins and Woodcock 1996), and the widely used post-classification comparison method (Jensen et al. 1995). More recent change detection methods are based on expert

systems, artificial neural networks, fuzzy sets, and object-oriented approaches. These change detection methods are explained in Lu et al. (2004) and Jensen (2005).

This chapter provides several examples of remote sensing change detection based on new change detection techniques using the remote sensor data obtained from 1) a digital frame camera, and 2) a LIDAR (Light Detection and Ranging) sensor system. These sensors function according to the logic shown in Figure 1. The change detection techniques include neighborhood correlation image analysis and single date elevation-based subsidence detection.

2.2 Remote Sensing Change Detection Process

Jensen (2005) reviews the general steps that are used to conduct change detection using remotely sensed data. The steps include 1) specifying the nature of the change detection problem, 2) identifying the remote sensing system and environmental considerations associated with change detection, 3) processing remote sensor data to extract change information by applying appropriate change detection techniques, and 4) evaluating the change detection results. Using these steps, scientists are able to decide whether their change detection results are of value. Selecting appropriate remote sensor data and change detection techniques according to the nature of the change detection problem under investigation is critical in change detection studies.

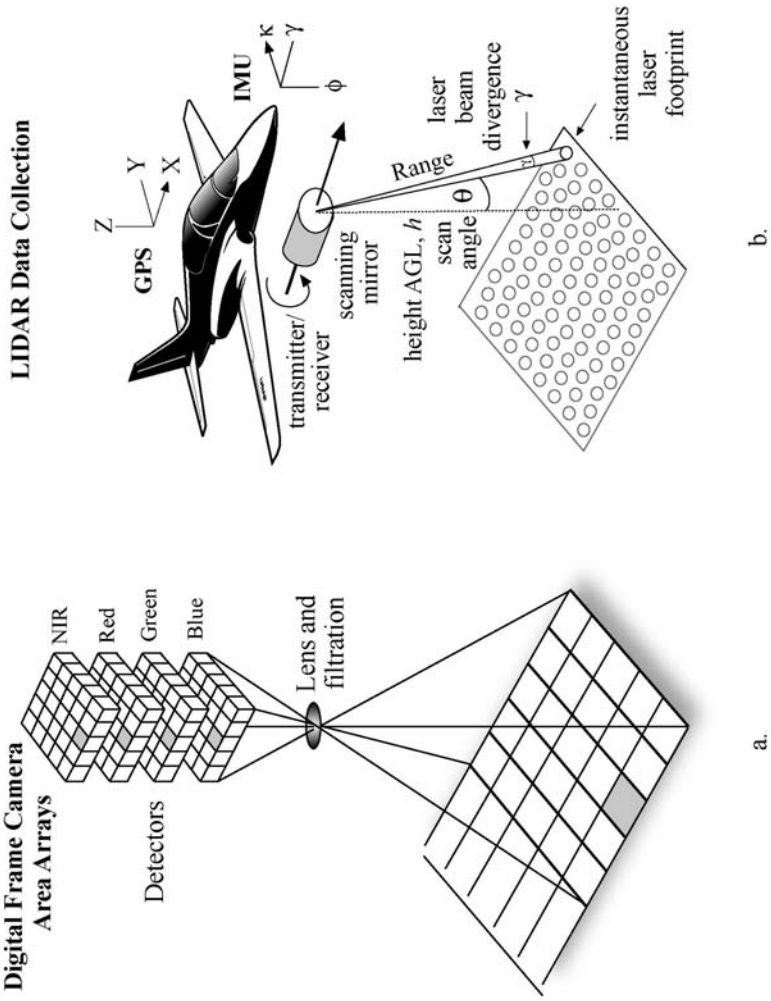

Fig. 1. Two remote sensor systems often used to collect information in the urban environment. a) Digital frame camera based on area arrays. b) LIDAR scanner.

2.2.1 Digital Frame Camera Remote Sensing

Digital frame cameras have many similarities to regular cameras. Instead of film, however, they use an area array of charge-couple-devices (CCD) detectors (Figure 1). Like a traditional camera system, the digital CCD area array records a "frame" of terrain during a single exposure. Three parameters determine the geographic area of the terrain recorded by the CCD area array, including 1) the dimension of the CCD array in rows and columns, 2) the focal length of the camera lens (the distance from the rear nodal point of the lens to the CCD array), and 3) the altitude of the aircraft above ground level (Jensen 2005). A major advantage of digital frame camera remote sensing is its timeliness. The remote sensor data are available as soon as they are collected since there is no need for an analog-to-digital (A to D) conversion.

2.2.2 LIDAR Remote Sensing

LIDAR is an optical remote sensing system that uses near-infrared laser light to measure the range from the sensor to a target on the surface of the Earth. Three fundamental technologies are used in the LIDAR system, including 1) laser range-finding, 2) differential global positioning system (DGPS), and 3) inertial measurement units (IMUs). LIDAR was initially introduced to facilitate the data collection for digital elevation models (DEM). Digital elevation information is a critical component of most geographic databases used by many agencies such as the USGS and FEMA. Digital elevation models can be subdivided into digital surface models (DSM) and digital terrain models (DTM). DSM contain elevation information about all features in the landscape, including vegetation and buildings. DTM contain elevation information solely about the bare-Earth surface (Jensen 2006). LIDAR technology can be used to generate the two types of elevation models.

Most LIDAR systems that are used for terrestrial topographic mapping use near-infrared light from 1040 to 1060 nm. Blue-green laser light centered at approximately 532 nm is used for bathymetric mapping due to its water penetration capability (Mikhail et al. 2001; Boland et al. 2004). Since LIDAR is an active system, it can also be used at night. The accurate measurement of the laser pulse travel time from a light transmitter to a target on the ground and back to a receiver is critical in the LIDAR systems. The range measurement process produces elevation data points, which are commonly referred to as masspoints.

One of the advantages of LIDAR remote sensing is that each LIDAR point is already georeferenced. It does not require additional geometric correction (Flood and Gutelius 1997). LIDAR systems receive multiple returns depending on the type of a target on the Earth surface. If a laser pulse hits directly on the ground, it will be recorded as a single return. If there are any materials (trees, grass) with local relief within the instantaneous footprint of a pulse, then the pulse will produce multiple returns (first, second … last returns). First returns including single returns can be used to generate a DSM, while last returns can be used to create a DTM. Additional processing is generally required to generate a DTM from last returns because some laser pulses never make it to the ground in heavily forested areas.

Most LIDAR systems provide intensity information in addition to the multiple return range data. The recorded intensity is in most cases just the maximum of the returned signals (Baltsavias 1999). The intensity values are dependent on several factors including gain setting, bidirectional effects, the size of the target, range to the target, angle of incidence and atmospheric dispersion (Leonard, 2005).

Neighborhood Correlation Image Analysis

The Neighborhood Correlation Image (NCI) analysis concept was introduced by Im and Jensen (2005). Correlation analysis can be applied to bitemporal imagery in a specified neighborhood to extract spectral *contextual* information, which contains three unique variables associated with the change in two dates of imagery. These variables include neighborhood correlation, neighborhood slope, and neighborhood intercept. The neighborhood correlation variable represents Pearson's product-moment correlation coefficient between the brightness values from bi-temporal imagery in a specified neighborhood. The neighborhood slope and intercept variables are calculated using the least squares estimates from the sets of brightness values:

$$correlation = \frac{\sum_{i=1}^{n}\sum_{j=1}^{k}(BV_{ij1} - \mu_1)(BV_{ij2} - \mu_2)}{s_1 s_2 (n \times k - 1)} \quad (1)$$

$$slope = \frac{\sum_{i=1}^{n}\sum_{j=1}^{k}(BV_{ij1} - \mu_1)(BV_{ij2} - \mu_2)}{s_1^{\,2}(n \times k - 1)} \tag{2}$$

$$intercept = \frac{\sum_{i=1}^{n}\sum_{j=1}^{k} BV_{ij2} - a\sum_{i=1}^{n}\sum_{j=1}^{k} BV_{ij1}}{n \times k} \tag{3}$$

where n is the number of pixels in a specified neighborhood, and k is the number of bands in each dataset. s_1 and s_2 are the standard deviations of the brightness values found in all bands of each dataset in a specified neighborhood, respectively. BV_{ij1} and BV_{ij2} are the ith brightness values of the pixels found in band k of the Date 1 and Date 2 images in a specified neighborhood, and μ_1 and μ_2 are the means of brightness values found in all bands of the Date 1 and Date 2 images in a specified neighborhood, respectively.

If the spectral changes of the pixels within a specified neighborhood between the two dates are significant, the correlation coefficient between the two sets of brightness values in the neighborhood will decrease to a lower value. The slope and intercept values may increase or decrease depending on the magnitude and direction of the spectral changes. Ideally, if there is no change in a certain pixel location between two dates, the pixel will have high correlation, a slope around 1, and an intercept around 0. An example of correlation analysis with two sample locations (change vs. no change) from bi-temporal ADAR digital frame camera imagery is shown in Figure 2.

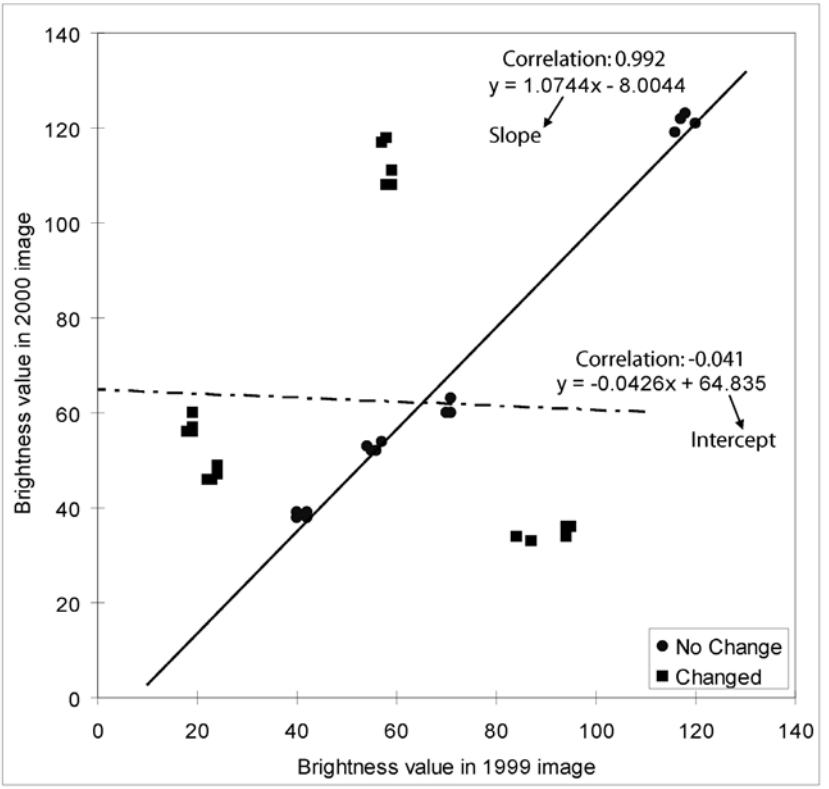

Fig. 2. An example of correlation analysis with two sample locations (change versus no change) based on the analysis of bi-temporal ADAR imagery.

Several shapes of neighborhood can be applied to neighborhood correlation image analysis within a GIS context, including rectangle, circle, annulus, wedge, and irregular. A module to create NCIs was developed as a dynamic linked library (DLL) in the ESRI ArcMap 9.1 environment using Visual Basic. Two general shapes – rectangle (square) and circle – of neighborhood were incorporated into the module. The size of neighborhood can be specified by users (e.g., 3 × 3).

2.3 Case Study 1 – Land Cover Change Detection Using NCI Analysis and Decision Tree Classification

The objectives of this study were to explore three types of neighborhood correlation image variables using several neighborhood configurations and to extract detailed "from-to" change information from bi-temporal imagery plus the NCIs using a decision tree classifier (Im and Jensen 2005). This study examined five different sizes of circular neighborhoods (i.e., 1- to 5-pixel radius). The study area, located in Edisto Beach near Charleston, SC, exhibited considerable residential development between two dates of imagery. The processing steps required to implement the change detection study based on neighborhood correlation image analysis and decision tree classification are summarized in Figure 3.

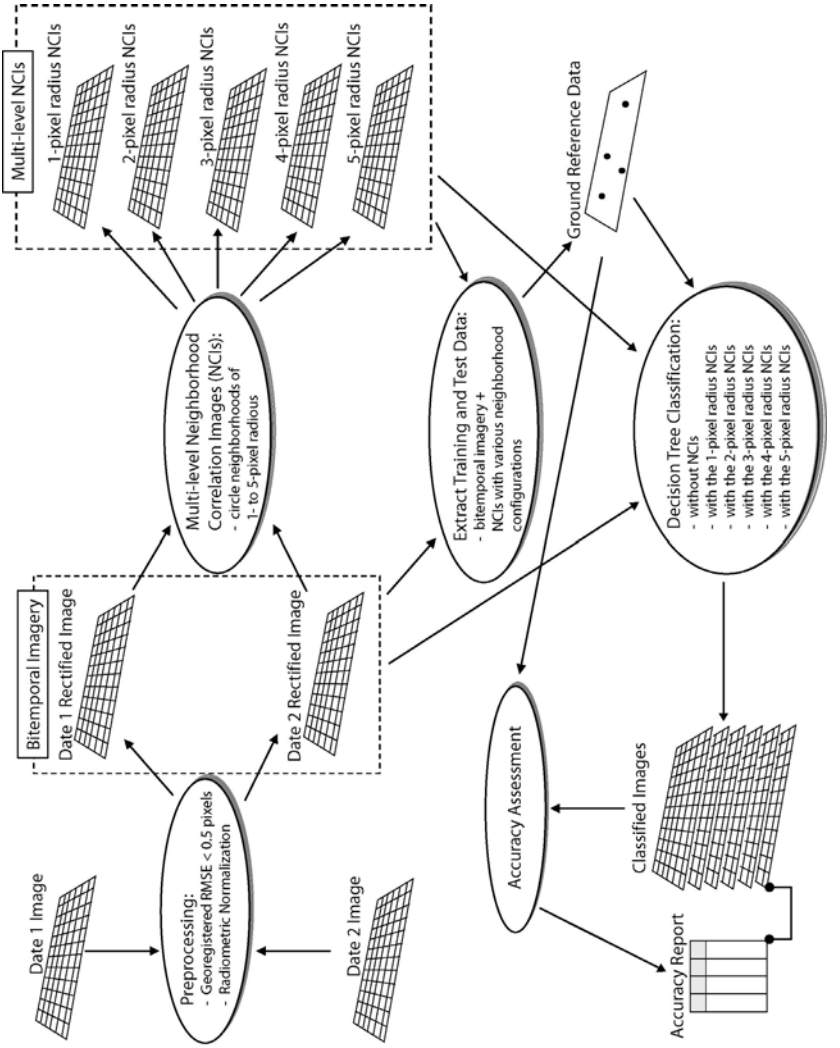

Fig. 3. Data processing flow diagram of the urban case study.

Bi-temporal remote sensing data were collected on September 23, 1999 and October 10, 2000 over the study area using an ADAR 5500 area array frame camera in four spectral bands, which included the blue, green, red, and near-infrared (Figure 4a,b). The multispectral data were preprocessed

(coregistered and radiometrically normalized) before the creation of neighborhood correlation images and subsequent change detection.

Eight hundred checkpoints were randomly generated and used as ground reference information. Each checkpoint was investigated using visual interpretation and assigned to one of thirteen land cover change classes. The classes included eight change classes (Barren to Developed, Tree to Developed, Tree to Barren, Grass to Barren, Barren to Grass, Tree to Grass, Tree to Shadow, and Change in Grass) and five unchanged classes (Developed, Barren, Tree, Grass, and Shadow). Five hundred of the samples were used to train a decision tree classifier. The remaining three hundred samples were used to evaluate the accuracy of the change classification.

Five sizes of circle neighborhoods were explored (1- to 5-pixel radius). Figures 4c-e depict the 1-pixel radius neighborhood correlation image variables. Based on visual inspection, large neighborhood sizes (e.g., 5-pixel radius) reduced noise e.g., caused by shadow difference in the NCIs, yielding a smoothing effect in the images. However, it altered change information (size) to some extent, e.g., a narrow linear change was barely distinguishable in the NCIs. Conversely, the use of a small neighborhood size helps detect change more precisely, but can introduce some noise. Two-dimensional planes between the 3-pixel neighborhood correlation image variables based on the eight hundred reference data are shown in Figure 5. In most cases, the unchanged samples resulted in high correlation values and slope values ~1 and intercept values ~ 0. Conversely, the changed samples generally exhibited low correlation values and variant slope and intercept values. Although a few changed samples yielded relatively high correlation (e.g., Barren to Developed), they were distinguishable using the other two variables (i.e., slope and intercept). These three unique change variables were very useful for the identification of change versus no change in the study area.

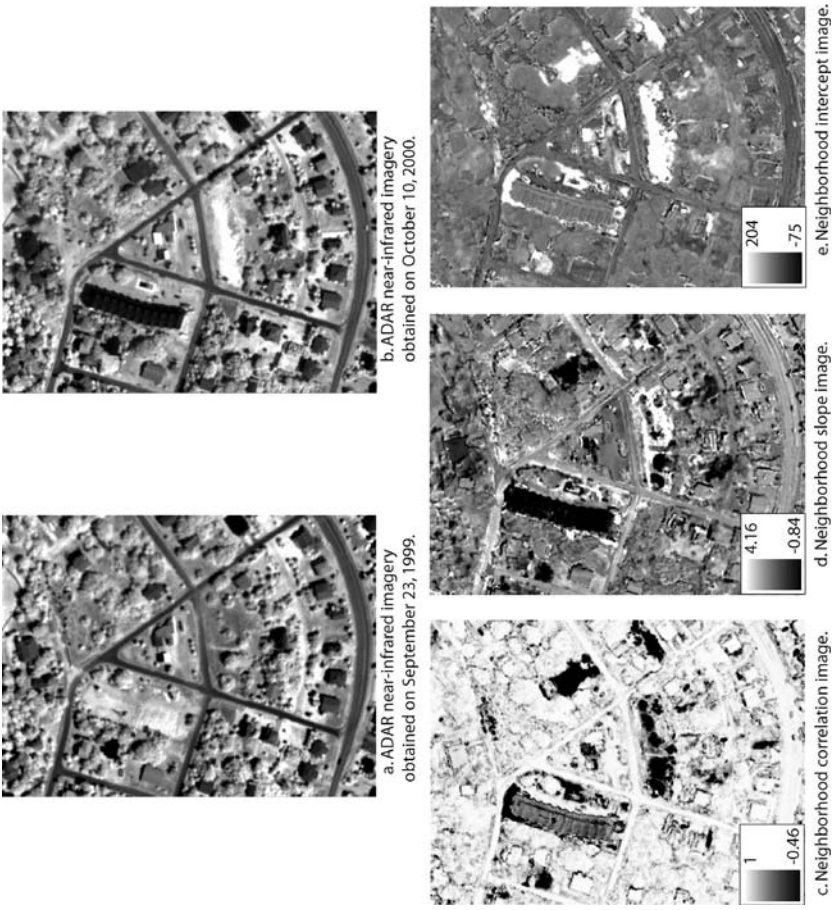

Fig. 4. a), b) Bi-temporal ADAR near-infrared imagery obtained on September 23, 1999 and on October 10, respectively. c) - e) Neighborhood correlation images (correlation, slope, and intercept) created from the bi-temporal ADAR imagery.

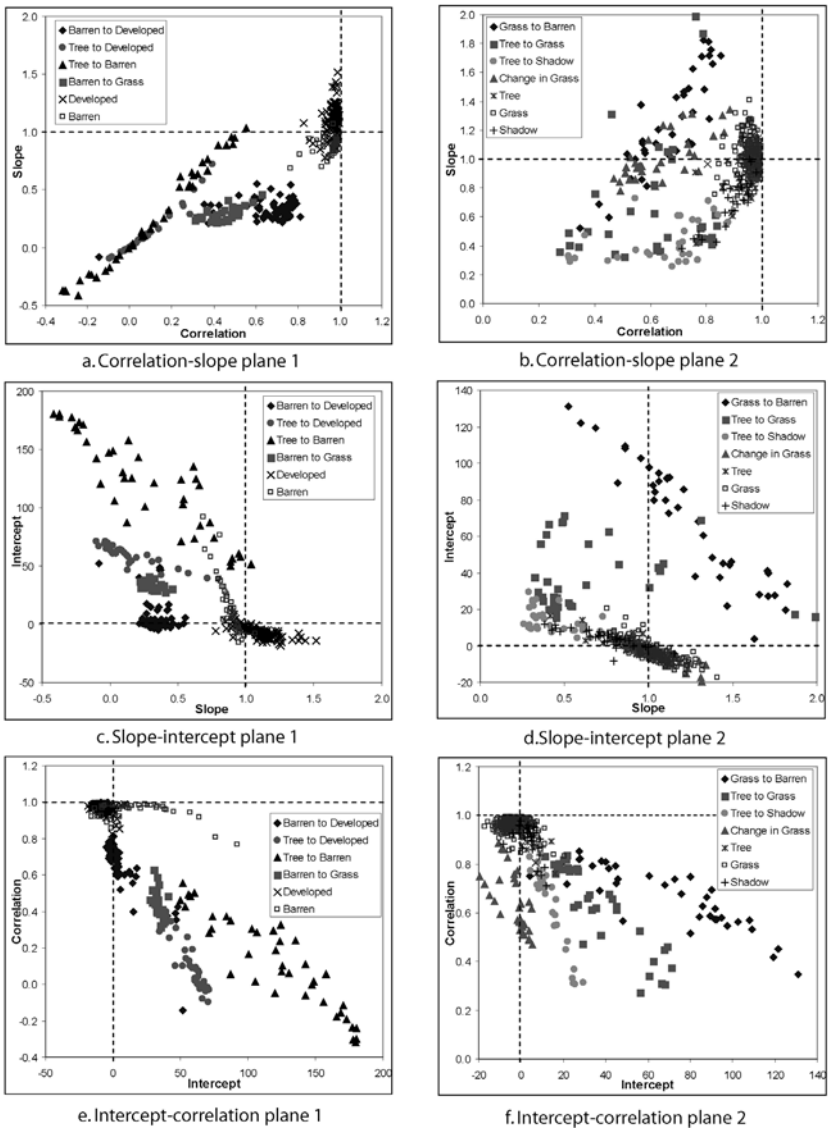

Fig. 5. Two-dimensional planes between the three NCI variables using the reference data.

The C5.0 decision tree was utilized to classify the land cover change using the bi-temporal data plus the NCI information. A detail discussion of C5.0 is found in Jensen (2005) and Quinlan (2003). Five hundred samples were

used to train a decision tree classifier and the remaining three hundred samples were used to evaluate the change classification. In order to apply the decision tree generated from C5.0 to the corresponding imagery, a C5.0 decision tree inference engine was developed and embedded in the ESRI ArcMap 9.1. All change classifications that included the NCIs resulted in significantly higher Kappa accuracies than the change classification based solely on the use of bi-temporal imagery (Table 1a). The change classification that incorporated the 3-pixel neighborhood correlation images produced the highest accuracy (overall accuracy = 94.3%; Kappa = 0.94). Figure 6 shows the change classification output image using the 3-pixel radius NCIs and the change detection matrix between the two dates.

Table 1. a) Land-cover change classification results based on the thirteen classes using a decision tree classifier.

Category	Overall accuracy	Kappa	ASE	Kappa *Z*-test (between the first case and others)
Bi-temporal data	87.3%	0.86	0.0216	N/A
Bi-temporal data plus 1-pixel radius NCIs	92.3%	0.91	0.0175	Significant (1.99)
Bi-temporal data plus 2-pixel radius NCIs	93.3%	0.92	0.0164	Significant (2.47)
Bi-temporal data plus 3-pixel radius NCIs	94.3%	0.94	0.0152	Significant (2.97)
Bi-temporal data plus 4-pixel radius NCIs	93.3%	0.92	0.0164	Significant (2.47)
Bi-temporal data plus 5-pixel radius NCIs	92.7%	0.92	0.0171	Significant (2.16)

b) Binary change classification results using a decision tree classifier.

Category	Number of rules	Overall accuracy	Kappa	Kappa *Z*-test (between the first case and others)
Bi-temporal data	10	90.7%	0.81	N/A
Bi-temporal data plus 1-pixel radius NCIs	4	98%	0.96	Significant (1.99)
Bi-temporal data plus 2-pixel radius NCIs	3	98.3%	0.96	Significant (2.47)
Bi-temporal data plus 3-pixel radius NCIs	5	99%	0.98	Significant (2.97)
Bi-temporal data plus 4-pixel radius NCIs	4	98.3%	0.96	Significant (2.47)
Bi-temporal data plus 5-pixel radius NCIs	6	97.7%	0.95	Significant (2.16)

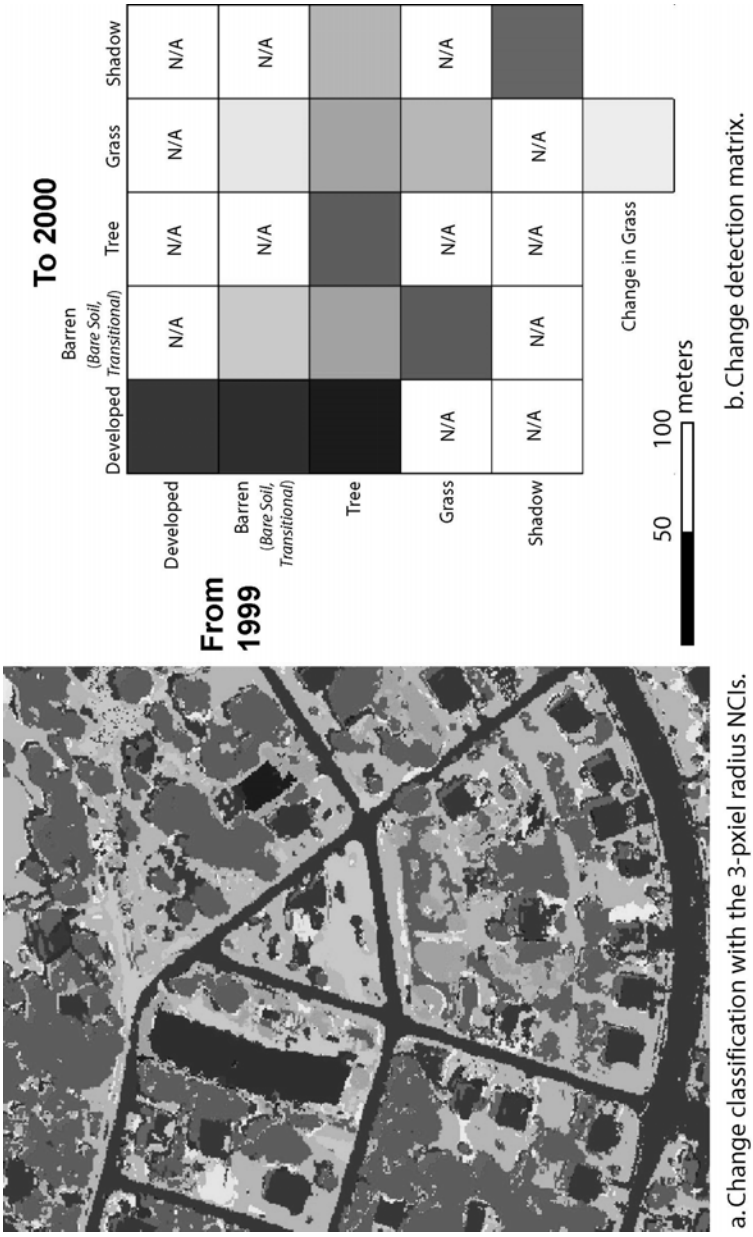

Fig. 6. a) Change classification output including the 3-pixel radius NCIs. b) Change detection matrix between two dates.

Three hundred samples were not sufficient to evaluate the change classification with the thirteen classes. For more reliable evaluation statistics, binary change detection was also conducted using the same reference data and decision tree classification. The decision tree binary classifications with the NCIs resulted in higher accuracies compared to the binary change detection without the NCIs (Table 1b). In addition, the number of rules generated from the decision trees was generally less than one-half of the number of rules from the decision tree without the NCIs. Binary change detection using the NCIs (i.e., without bi-temporal imagery) yielded very high accuracies over 97%.

These results support the use of neighborhood correlation image variables for change detection. Various levels of neighborhood correlation images have their own characteristics. The concept of neighborhood correlation images can be extended to "objects," and object correlation images (OCIs) may be incorporated into object-oriented change detection.

2.4 Case Study 2 – Subsidence Detection Using Single-date LIDAR-derived Elevation Data

Human beings have produced large amounts of hazardous waste. Hazardous waste must be stored in safe places to avoid contaminating the environment. Monitoring hazardous waste sites is also an essential safety measure. One of the possible failures on hazardous waste sites is subsidence of surface materials such as claycaps due to damage to the storage underneath. The purpose of this study was to investigate the potential of single-date LIDAR data with dense postings to detect subsidence in experimental waste sites at the Savannah River National Laboratory (SRNL) near Aiken, SC.

SRNL installed claycaps to hold nuclear-related hazardous waste products buried in shallow pits (Jensen et al. 2006). Claycap monitoring is normally conducted through *in-situ* visual inspection, which is very costly and may miss early claycap failure. Conversely, it is possible to use remote sensing techniques such as photogrammetry or lidargrammetry to identify subsidence or other direct topographic expressions of claycap failure on the order of just a few centimeters (Garcia-Quijano 2006).

This project used LIDAR data obtained by an Optech ALTM 2050 sensor mounted on a Cessna 337 aircraft flown by Sanborn, Inc. of Charlotte, NC.

The LIDAR data were collected over SRNL on November 14, 2004. The LIDAR data collection included small footprint first and last return location (*x*, *y*, and *z*) and intensity data using a 1064 nm laser at a pulse repetition frequency (PRF) of 50 kHz. The nominal post spacing was 0.4 m at an altitude of 700 m above ground level (AGL) over an area of 2.6 km^2. Last return LIDR data were processed using TerraModel's TerraScan morphological filtering software, eliminating obstructions on the ground such as trees and buildings to generate a bare-Earth elevation. An accuracy assessment of the LIDAR-derived elevation is found in Garcia-Quijano et al. (2006). The elevation and hillshaded surfaces of the first returns and bare Earth LIDAR data using the IDW interpolation method are shown Figure 7. The experimental waste sites, which were used for subsidence investigation, are located in the middle-left of the surfaces. Figure 8 depicts the LIDAR-derived digital terrain model overlaid with 25 cm contours showing two locations of potential subsidence.

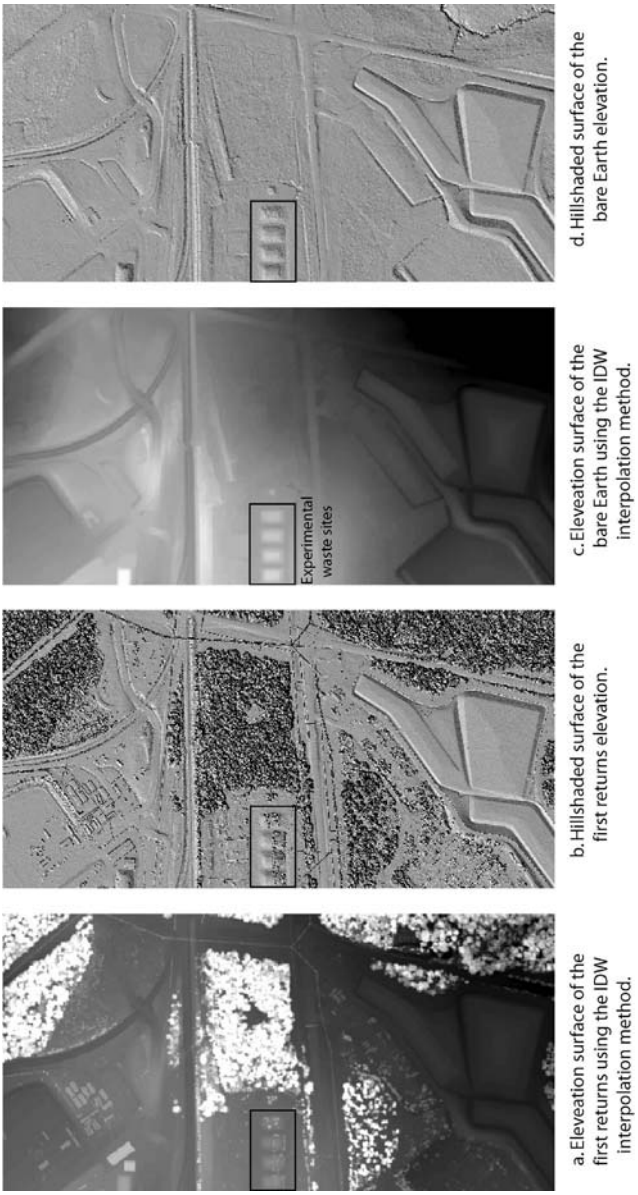

Fig. 7. Elevation surfaces of the first returns and bare-Earth LIDAR data using the IDW interpolation method (a and c) and the hillshaded surfaces of the elevation data (b and d).

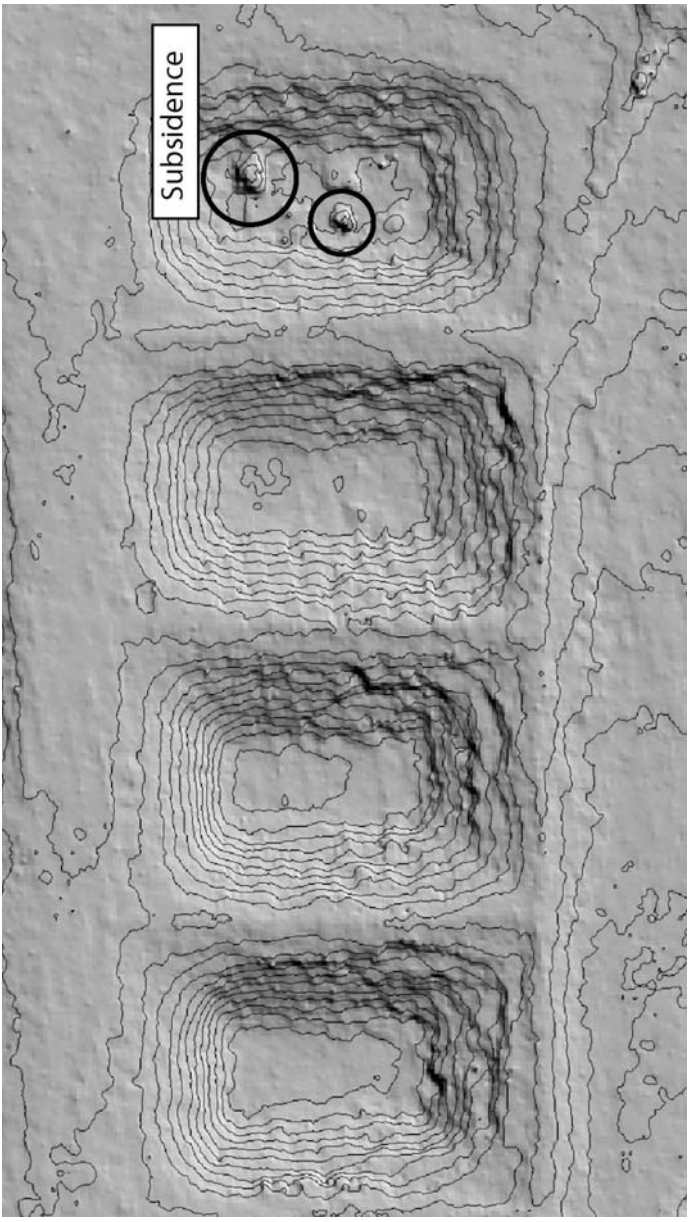

Fig. 8. LIDAR-derived digital terrain model with 25 cm contours showing two locations of potential subsidence.

The processing flow diagram of the study focusing on single date elevation-based subsidence analysis is summarized in Figure 9. The tabular LIDAR bare Earth data were imported into ESRI ArcMap 9.1 as a point shapefile. An elevation surface was generated based on the LIDAR bare Earth masspoints using an Inverse Distance Weighted (IDW) interpolation method (Figure 7). The elevation surface was used as reference elevation in the single elevation-based subsidence detection module. The logic of the single elevation-based subsidence detection include extraction of regular samples using the user-defined parameters, generation of simulated elevation surface using the regular samples and the IDW interpolation method, and creation of subsidence images from the difference between the reference and simulated elevation surfaces with a user-specified subsidence threshold.

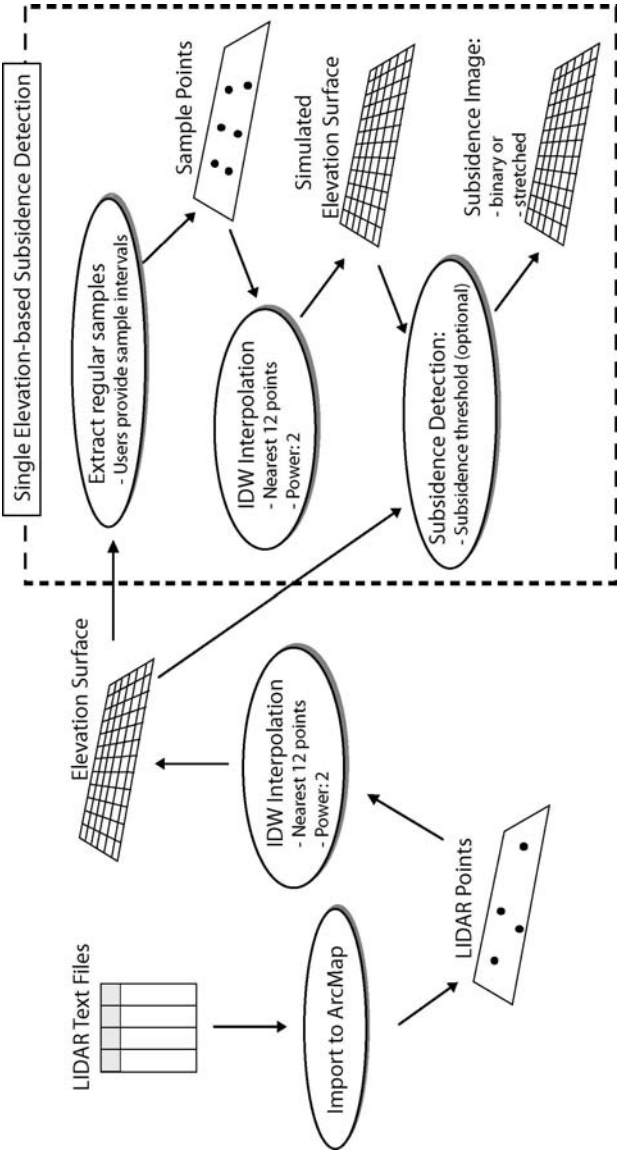

Fig. 9. Data processing flow diagram of single LIDAR-derived elevation-based subsidence detection.

Various user-specified parameters were applied to the single elevation-based subsidence detection module. The four selected subsidence detection results are shown in Figure 10. The subsidence detection based on the 3 m-interval sampling and 0.4 m threshold parameters yielded the best result (two locations inside the circle in Figure 10b). Although other combinations of parameters also detected the subsidence, most of them tended to overestimate subsidence, which resulted in false alarms on the normal claycaps. Those false alarms can be easily found in Figures 10c and 10d.

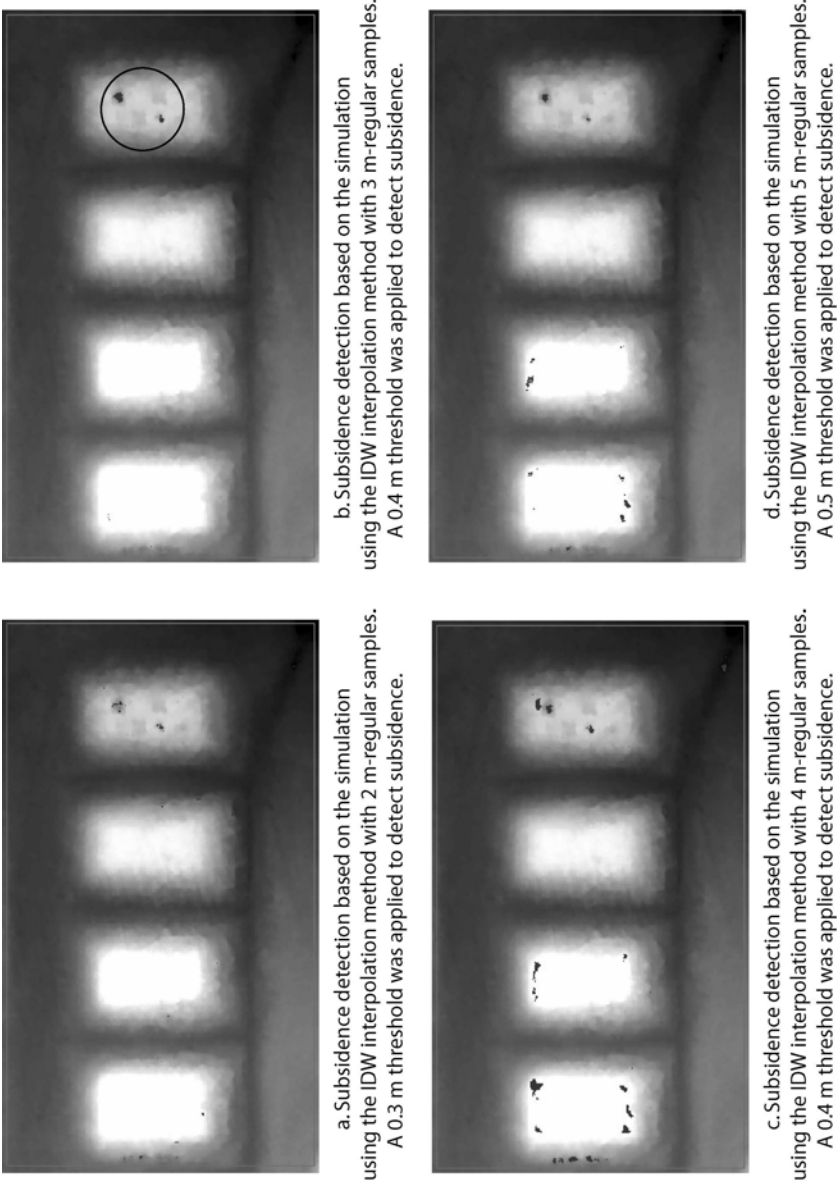

Fig. 10. Subsidence detection results associated with the experimental waste sites and using different parameters. The 3 m-interval sampling and 0.4 m threshold for subsidence resulted in the best result based on visual inspection.

Subsidence detection using multiple date (LIDAR-derived) elevation data may provide much more accurate and precise subsidence information. However, obtaining multiple date elevation data with high quality is not always possible. This study suggests that single-date quality elevation-based subsidence analysis can be an alternative to the multiple date approach in hazardous waste site monitoring.

2.5 Conclusion

As advanced remote sensors provide improved high-quality data, new and/or more sophisticated techniques are needed to extract accurate and reliable change information from the data. This chapter provided examples of the application of new digital change detection techniques using two different remote sensing data sources for change detection in urban environments.

References

Baltsavias EP (1999) Airborne Laser Scanning: Basic Relations and Formulas. *ISPRS Journal of Photogrammetry & Remote Sensing* 54:199-214

Boland J, 16 co-authors (2004) Chapter 8: Cameras and Sensing Systems. In McGlone JC (Ed.) *Manual of Photogrammetry* 5th Ed. Bethesda: ASP&RS 629-636

Chan JC, Chan K, Yeh AG (2001) Detecting the Nature of Change in an Urban Environment: a Comparison of Machine Learning Algorithms. *Photogrammetric Engineering & Remote Sensing* 67(2):213-225

Collins JB, Woodcock CE (1996) An Assessment of Several Linear Change Detection Techniques for Mapping Forest Mortality Using Multitemporal Landsat TM Data. *Remote Sensing of Environment* 56:66-77

Coppin PR, Bauer ME (1996) Digital Change Detection in Forest Ecosystems with Remote Sensing Imagery. *Remote Sensing Reviews* 13:207-234

Flood M, Gutelius B (1997) Commercial Implications of Topographic Terrain Mapping Using Scanning Airborne Laser Radar. *Photogrammetric Engineering & Remote Sensing* 63:327

Garcia-Quijano M (2006) *Claycap Anomaly Detection using Hyperspectral Remote Sensing and Lidargrammetric Techniques*. Ph.D. Dissertation, Department of Geography, University of South Carolina, 141 p.

Garcia-Quijano M, Jensen JR, Hodgson ME, Hadley BC, Gladden JB, Lapine LA (2006) Significance of Altitude and Posting-density on Lidar-derived Eleva-

tion Accuracy on Hazardous Waste Sites. *Photogrammetric Engineering & Remote Sensing* in press.

Im J, Jensen JR (2005). A Change Detection Model Based on Neighborhood Correlation Image Analysis and Decision Tree Classification. *Remote Sensing of Environment* 99:326-340

Jensen JR (2005) *Introductory Digital Image Processing*. 3rd Ed. Upper Saddle River NJ: Prentice Hall 526 p.

Jensen JR (2006) *Remote Sensing of the Environment: An Earth Resource Perspective*. 2nd Ed. Upper Saddle River NJ: Prentice Hall 592 p.

Jensen JR, Garcia-Quijano M, Im J, Hadley BC, Gladden JB (2006) Development of a Remote Sensing and GIS-assisted Spatial Decision Support System for Hazardous Waste Site Monitoring. *Photogrammetric Engineering & Remote Sensing* in review

Jensen JR, Rutchey K, Koch M, Narumalani S (1995) Inland Wetland Change Detection in the Everglades Water Conservation Area 2A Using a Time Series of Normalized Remotely Sensed Data. *Photogrammetric Engineering & Remote Sensing* 61(2):199-209

Lu D, Mausel P, Brondizio E, Moran E (2004) Change Detection Techniques. *International Journal of Remote Sensing* 25(12):2365-2407

Leonard J (2005) *Technical Approach for LIDAR Acquisition and Processing*. Frederick MD: EarthData Inc. 20 p.

Maune DF (Ed.) (2001) *Digital Elevation Model Technologies and Applications: The DEM Users Manual*. Bethesda: ASP&RS 1151 p.

Mikhail EM, Bethel JS, McGlone JC (2001) *Introduction to Modern Photogrammetry*. NY: John Wiley 479 p.

Quinlan JR (2003) *Data Mining Tools See5 and C5.0*. St. Ives NSW Australia: RuleQuest Research http://www.rulequest.com/see5-info.html

3 Assessment of Risk in Urban Environments Using Geo-Spatial Analysis

James D. Hipple, USDA Risk Management Agency, Washington, DC, USA

Geospatial technologies are focused around the acquisition, integration, analysis, visualization, management and distribution of data having an explicit spatial and temporal context (Wachter et al. 2006). These data are usually analyzed within geographic information systems (GIS). These technologies have grown to include a wide array of technologies, many of which are actively used in urban risk assessment. First, these geospatial tools and technologies are often used for the identification of "hazards" or the establishment of "risk" parameters, like height above flood stage (elevation derived through photogrammetric methods or LIDAR) and proximity to hazards (distance). Second, they can be used to actively map risk (e.g., active wildfires detected through remote sensing). Finally, these geospatial technologies can be integrative through visualization tools and models often delivered through internet-based mapping and services.

3.1 Defining Risk

Risk is defined as probability that an event will occur, whereas hazard is a qualitative term referring to the potential of an environmental element doing harm (WHO 1988). The assessment of risk in urban places broadly covers hazard identification and risk quantification resulting from a specific hazard or vulnerability (WHO 1988). Quantifying these risks and their potential impact on urban places ideally requires the establishment of measures of the magnitude of effect and response relationships of the individuals and environment within that urban place.

All too often, we see what is defined as risk compensation occurring. Risk compensation is "effect whereby individual animals" (humans) "may tend to adjust their behavior in response to perceived changes in risk (Adams

1995)." When the perception of risk increases individuals are expected to act more cautiously; conversely, when individuals feel safer or more protected they will behave less cautiously (Adams 1995). As illustrated by the 1993 flooding of the Missouri and Mississippi Rivers covered in the case study section, or the flooding associated with the 2005 Hurricane Katrina in New Orleans, Louisiana, areas taken out of the 100-year or 500-year floodplain due to flood-wall construction or levee construction can indeed flood.

3.2 Methods for Hazard Identification and Risk Assessment

The disaster risk management tools derived from geospatial technologies include an active updated information base, risk assessment methods, hazard mapping, early warning systems, or disaster response plans. This section looks briefly at two methods, photogrammetric derivation of geospatial information and remote sensing as a source of data for integration into geospatial tools.

3.2.1 Photogrammetry

A highly accurate urban base map (or image) is often the starting point for most geospatial applications. This base, sometimes referred to as a planimetric base, is a layer that presents only the horizontal positions for features represented (Hipple & Haithcoat 2005). This base establishes the "truth" in a point in time of the status of the urban area upon which all other changes to are recorded and all other layers of information are associated. The cadastre serves as a foundation for other urban infrastructure applications, such as urban infrastructure condition assessment, utility line mapping, and precisely locating utility poles, signs, fire hydrants, and other infrastructure elements.

The geospatial technologies used for the generation of these cadastral bases include aerial photography (both analog and digital), with resolutions of 0.25 to 0.62 meters, high resolution satellite imaging from the panchromatic band of the commercial high-resolution satellites (IKONOS, QuickBird, OrbView) with spatial resolutions of 0.70 to 1.0 meters (Hipple & Haithcoat 2005). These images are usually acquired by jurisdictions once every one to five years and must be orthorectified (a process in which terrain and camera displacements have been removed from the image). Often, many vector-based data layers, "including street centerlines, parcel

boundaries, and sewer lines, are routinely superimposed upon the image base, these images must have a high degree of horizontal resolution and planar accuracy (Hipple & Haithcoat 2005)."

Detailed topography is another layer of geographic data often used in hazard identification and risk assessment. Parameters derived from the elevation data (slope & aspect) play important roles in many derived hazard and risk products and serve as important inputs to risk models. Traditional methods extract elevation through stereoscopic analysis of aerial photography, or in situ methods such as surveying. New methods for deriving elevation include airborne Interferometric Synthetic Aperture Radar (IFSAR), LIDAR (LIght Detection And Ranging), and Global Positioning System (GPS)-based surveying.

3.2.2 Remote Sensing

Remote sensing is a valuable tool for the identification of hazards, potential hazards and the evaluation of risk. The benefits of remote sensing include, but are not limited to, synoptic coverage, high temporal frequency, multi spectral capabilities, and the ability to fuse different data sets for analysis. Because the image footprints of satellites are so large, ranging from tens of kilometers (with the high resolution sensor systems like IKONOS, QuickBird and OrbView) through thousands of kilometers (with sensors such as MODIS and AVHRR) large areas or synoptic scale coverage permits the analysis of hazards in single scenes or in mosaics. Typically, we see moderate resolution data sets like Landsat TM and ETM+ (30-m spatial resolution with a 16-tday repeat), Indian Resourcesat AwIFS (56-meter spatial resolution with upwards of a 5-day repeat duer to extensive scene overlaps); and SPOT (20-meter spatial resolution).

Through synoptic scale analysis entire regions can be analyzed concurrently, whereas with aerial photography or ground-based methods, numerous images must be analyzed or numerous site visits must take place. Many sensor systems, particularly those with coarser resolutions like AVHRR (1-km pixel) or MODIS (250-m pixel)) have daily revisit rates (high temporal frequency). These sensors can be used to monitor emerging risks and are essential in post- event evaluation. Figure 1 is an image from the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor system of the emerging devastation surrounding Hurricane Katrina acquired on September 7, 2005. The extent of the flooding can be seen southeast of Lake Pontchartrain which is in the center of the image. Other

systems with longer repeat cycles (temporal frequency) have less of an opportunity to collect cloud free imagery close to the date of an event.

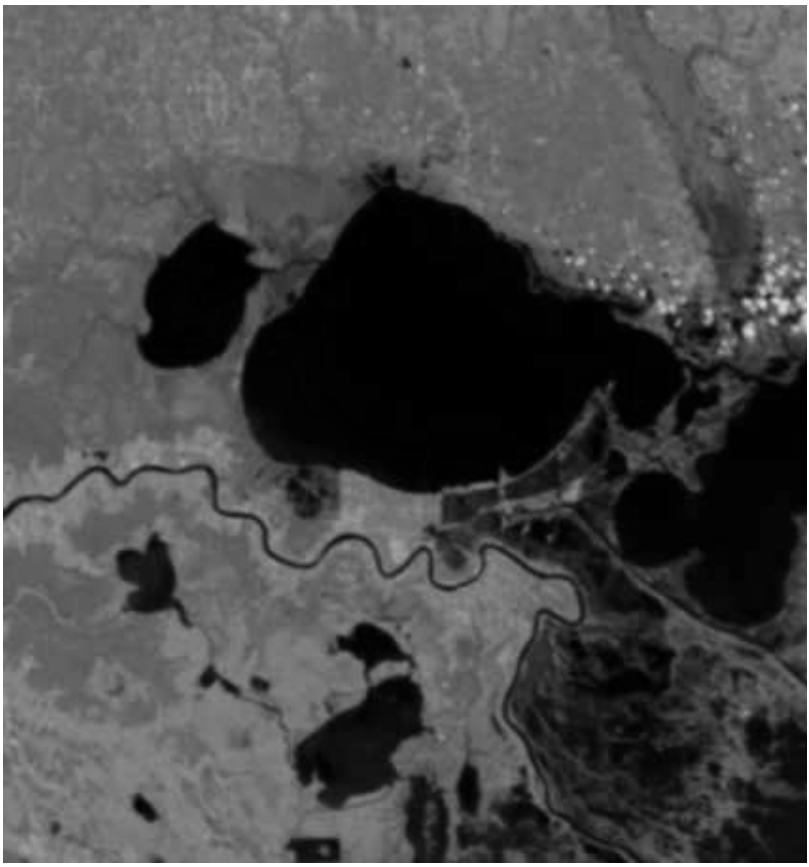

Fig. 1. September 7, 2005 MODIS image of New Orleans, LA (Courtesy: NASA).

The added ability to collect and analyze multiple spectral bands (multispectral) of information in the visible and infrared portions of the electromagnetic spectrum allows for the quantitative detection of objects in the image, classification of land covers, and ability to monitor changes in the images and allows for the automation of these for hazard detection and risk analysis. Once the images are acquired the data can be fused or integrated with different types of remote sensing products, elevation data and with other geospatial data sets.

Case Study: Wildfire Mapping

An example of combining many geospatial technologies in the evaluation of risk is the USDA Forest Service MODIS Active Fire Mapping Program. Although the primary purpose of this program was to monitor the fire activity that occurs annually in the western states, it has further been expanded to the Eastern and Southern United States. Such wild fires pose direct risk to human settlements and urban areas throughout the United States (McNamera et al. 2002, Quale 2005).

The Active Fire Mapping Program is based upon the NASA remote sensing platform MODIS TERRA and AQUA satellites. The data from these satellites provide four-times daily coverage (10:30am, 1:30pm, 10:30pm, 1:30am local times) to the mid- and high- latitudes. Data is downloaded and processed by the USDA Forest Service Remote Sensing Application Center (RSAC) in Salt Lake City for the western United States and by the NASA Goddard Space Flight Center (GSFC) for the eastern United States. The data is acquired near-real time (3-6 hour delay) (Quale 2005).

Fire locations are detected using the thermal bands of MODIS (1-km band 21 with the 3.929-3.989 micrometer wavelength and the 1-km band 31 with a10.780-11.280 micrometer wavelength) (NASA 2005). The algorithms used by the Active Fire Mapping Program are described Gigolo et al. (2003) and consist of outputs of fire locations as centroids of 1-km pixels in a gis format. Once fire locations are detected, the data is fed into an Internet mapping service (see Figure 2).

Fig. 2 MODIS Active Fire Map (http://activefiremaps.fs.fed.us)**.**

FireLine ™, a commercial product of ISO Properties, Inc., utilizes satellite imagery and other geospatial information to quantify and locate potential wildfire hazards. FireLine ™ is a geospatial technology based hazard detection program from which the data can be fed into a risk model for underwriting hazard property insurance (ISO Properties, Inc. 2004). The FireLine ™ database uses Landsat TM sensor data for determining availability of fuels, data about accessibility of sites based upon road access, and slope, a significant factor in wildfire spread and creates hazard scores for fuels, slope, and road access tied to given residential and commercial addresses.

Both the Active Fire Mapping Program and the FireLine ™ product from ISO Properties, Inc. provide geospatial technologies useful for hazard identification and risk assessment in and around urban environments. The evolution of geospatial technologies over the past decade, including the

growth and refinement of internet mapping services, makes these data sets accessible for urban risk assessment.

Case Study: Flood Impact Assessment

The Flooding in the Upper Mississippi River Basin during the summer of 1993 caused between US$ 12 and 16 billion worth of damage to urban and rural infrastructure and cropland. Since 1993, millions of dollars of new development have poured into the flood-impacted areas much of it in or around urban areas like St. Louis, Missouri, contrary to the recommendations of Interagency Floodplain Management Review Committee, among others (Hipple et al. 2005; Pinter 2005). Tracking development is difficult due to the diverse regulations and land use controls that vary widely by jurisdiction, causing varying amounts of development in the Upper Mississippi River Basin (Hipple et al. 2005). The amount of such infrastructure has dramatically increased, with approximately 28,000 new homes built, a 23% increase in population, and 26.8 km^2 (6630 acres) of commercial and industrial development added on land that was inundated during the 1993 flood (Hipple et al. 2005). In all, $2.2 billion in new development has occurred in the St. Louis area alone on land that was under water in 1993 (Carey 2005, Shipley 2005).

Immediately after the 1993 flooding, the interdisciplinary Scientific Assessment and Strategy Team (SAST) was formed which included specialists from the Soil Conservation Service (now referred to as the Natural Resource Conservation Service0, US Army Corps of Engineers (USCOE), US Fish and Wildlife USFWS), US Geological Survey (USGS), National Biological Service (NBS, now the Biological Services Division of the USGS), Environmental Protection Agency (EPA), and the Federal Emergency Management Agency (FEMA) (SAS, 1994).

The SAST developed a large spatial database for the Upper Mississippi and Missouri River Basins providing a baseline for the report of the Interagency Floodplain Management Review Committee (IFMRC), titled Sharing the Challenge: Floodplain Management into the 21st Century (commonly referred to as the Galloway Report) (IFMRC 1994), as well as the Report of the Scientific Assessment and Strategy Team (SAST) (SAST 1994). Large integrated spatial databases were created (perhaps the largest of the time) of elevation data sets, soil survey information, flood zone maps, flood extent maps, and census data, to name a few. This baseline included many key pieces of spatial information essential for the under-

standing and reconstruction of the Upper Mississippi and Missouri River Basins.

The drawback was that many of these data sets were less accessible previous to the disaster for real time evaluation and monitoring, partly due to data sharing issues at the time and partly due to the state of the technology in 1993; similarly, during Hurricanes Rita and Katrina in 2005, similar geospatial data sharing and delivery issues were experienced.

In response to events such as these, a Federal government task called The Homeland Infrastructure Foundation-Level Database (HIFLD) Working Group was formed consisting of "a coalition of federal, state, and local government organizations, Federally-funded Research and Development Centers (FFRDC), and supporting private industry partners" (http://www.hifld.org) who are involved with geospatial issues related to homeland defense and security, or emergency preparedness and response. The HFLD team created a Foundation-level data set for use by these agencies for emergency preparedness and disaster response.

Hipple et al. (2005) conducted a systematic study of development in the Upper Mississippi River Basin to document the changes in the basin affected by the 1993 floods ten years after the event by conducting an analysis to identify new development within the 500-year floodplain and in the floodwater inundated areas. The purpose of the study was to document development and analyze areas at risk should another event of similar magnitude occur. Since the completion of the SAST datasets and the report to the US Congress on the effects of the 1993 flooding by the IFMRC and SAST, no follow-up study had yet taken place.

The study conducted by Hipple et al. (2005) sought to analyze the development that occurred in the bluff-to-bluff area and floodwater inundated areas of the Upper Mississippi and Missouri River main stem that occurred during the period of 1993 to the present. This study defined the Upper Mississippi River Basin as the major Mississippi River tributaries upstream of Cairo, Illinois in the States of Illinois, Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota, and Wisconsin. The analysis focused on changes from floodplain agriculture or undeveloped land to developed floodplain. *Developed* is defined as having an artificial surface covering or replacing the natural surface that existed during the pre-1993 flood period. Artificial surfaces could include buildings, concrete, asphalt or turf, and adjacent land area. In addition, US Census Bureau data were analyzed to quantify housing unit and population changes

within each of these two areas, allowing a county-by-county and state-by-state comparison of population change and housing unit change within the flood affected areas (Hipple et al. 2005).

In order to quantify new development in the flood affected areas Landsat satellite imagery was classified looking for residential, commercial/industrial, and highway & interchange development. Fourteen Landsat scenes were analyzed for two dates. Landsat ETM+ data was acquired to analyze the most recent developments ("post-flood") and included cloud-free imagery from 2003, with some areas covered by imagery from the 2001 to 2002 time period. The "pre-flood" imagery consisted of Landsat TM imagery. The "pre-flood" Landsat data were the same images as used by the SAST team. The images were analyzed to identify and determine the aerial extent (in acres) of development in the floodplain (Hipple et al. 2005).

The U.S. Census data used in this project included housing and population statistics at the Census Block level; county level statistics were summed in ArcGIS. The 1990 and 2000 U.S. census Block boundary files were selected out based upon their inclusion in the SAST Floodplain boundary and the SAST Flood Extent boundary. Selecting criteria used was whether the block centroid was located within the selection. The U.S. Census data analysis gives some insight into the development in the flood plain, but what it fails to capture is population and housing unit changes from 1990 to 1993 (which could be either a gain or loss); and population and housing unit changes from 2000 through to the present (also, a gain or loss).

3.2.3 Results of the Analysis

In the St Louis Metropolitan Region, new development is identified in Maryland Heights, Hazelwood, Bridgeton, and county land in St Louis County and in St Charles, St Peter and Chesterfield in St Charles County in Missouri. These developments primarily occur in the floodplain and, with few exceptions, the tracks of new development in and around the St. Louis Metropolitan region are found wholly or in part in the 1993 flood inundated area. Development on the Illinois side of the St Louis Metropolitan Region, in contrast to the Missouri side, is almost entirely in areas of the floodplain not water inundated by the 1993 flooding (Hipple et al. 2005).

In the Kansas City Metropolitan Region, new development is identified in Kansas City in Clay and Jackson Counties. The development consists of

commercial or industrial development. This development is almost wholly in the 1993 flood inundated area.

Development and the addition or improvement of flood control structures, including urban and agricultural levees, in the Missouri and Mississippi River floodplain within Missouri follows no regional plan. Regional land use plans in Missouri are either nonexistent or lack sufficient regulatory power. Most land use planning takes place at the county or municipal level, where land use regulation, implemented through zoning, is locally regulated often leading to local interests being put ahead of the larger regional good.

3.2.4 How much development is occurring and of what type of development is it?

Commercial/Industrial

The types of development we observe from the satellite image analysis are not limited to any single class of development. Satellite image analysis identified residential, commercial, industrial, and recreational use development. Within the flood plain there were over 6630 acres of new commercial and industrial development, 2557 acres of residential development, and 2327 acres of highway and interchange development.

Missouri ranked first in new commercial and industrial development – 56% of the approximately 6,630 total acres of development occurred in Missouri – with 51% of the approximately 6,630 total acres of development just in St. Louis and St. Charles Counties (Hipple et al. 2005).

Residential

The majority of new residential development occurred in Wisconsin and Nebraska. Approximately 2,557 acres of new development were identified in Wisconsin mostly in La Crosse County near the city of Holmen. This development accounted for approximately 43% of the new residential development detected by satellite imagery analysis within the floodplain. Residential development in Nebraska accounted for nearly 40% of the 2557 acres of new residential development observed. All detected residential development, exclusive of one site in Wood River, Illinois occurred in floodplain areas not water inundated by the 1993 floods (Hipple et al. 2005).

Highway & Interchange

There have been a large number of highway and interchange developments within the floodplain and many have stretches built in the 1993 flood water inundated areas. Not only does this pose a risk to the highway and interchange (which estimates for construction range fro $20-million to $40-million per mile) but also to the often and inevitable pressures for development around highway interchanges (Moon 1988)

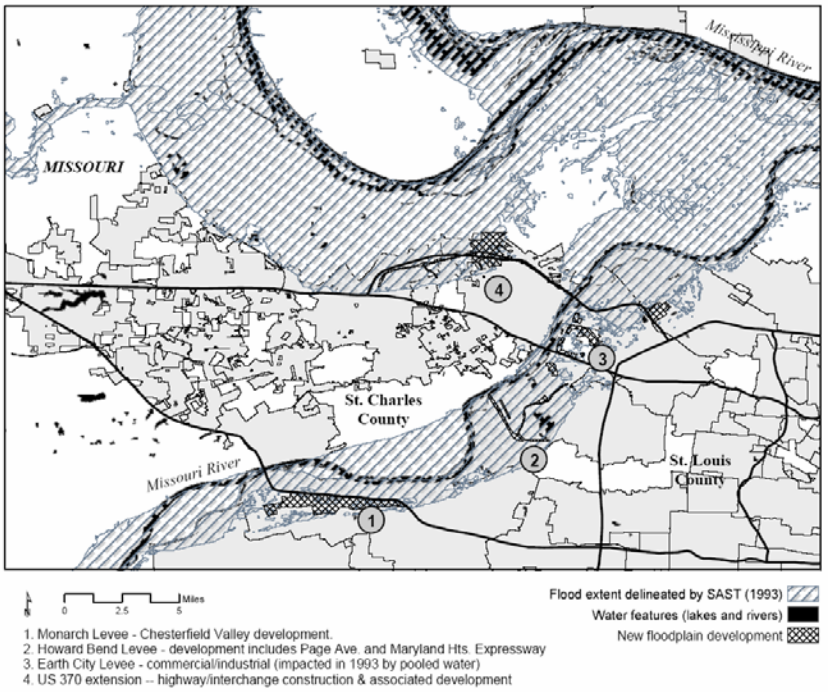

Fig. 3. The Missouri River Basin 1993 flood extent as delineated by the SAST with areas of new development within the flooded area (SAST 1994; Hipple et al. 2005).

3.2.5 Case Study Conclusions

Floodplain projects in the United States are constrained by Federal Emergency Management (FEMA) National Flood Insurance Program (NFIP), the Clean Water Act administered by the Army Corps of Engineers, and by more stringent state and local regulations. In general, guidelines limit development in the central portion of the floodplain (the "floodway"), but prohibit almost unlimited development in the rest of the floodplain so long as the development is raised above or a levee has been constructed with 100-year protection (1% chance of occurring in any given year) (Pinter

2005). Pinter (2005) points out that in some areas development has been enabled by the local governments in flood-prone areas, "in Sacramento, California, at least 60,000 new homes and billions of dollars of new infrastructure have been recently built or are planned on several floodplain tracts of the American, Feather, and Sacramento rivers (Leavenworth 2004a, 2004b)." Other municipalities "--including Denver and Boulder, Colorado; Austin, Texas; Phoenix, Arizona; and Charlotte, North Carolina -- have limited encroachment and guided development to more compatible locations and land uses (Pinter 2005)." With either more lax restrictions on floodplain development or greater control over development, the need for geospatial tools for monitoring growth in urban places prone to flooding is essential.

A valid public perception following the record 1993 flood damages are that a repeat flood would not cause similar damages; structures were removed from the floodplain; and it probably won't be that bad again, after all the Federal government had a major study and task force convened to address the flood and put mechanisms in place to prevent that scale of damage in the future. However, if private levees have been raised since the 93 flood, if the watershed has lost any storage capability, if the floodplain has been further reduced by new levees or the repair of levees that broke in 1993, then a similar event could be even more devastating and put new construction at increased risk if it is behind a levee of 1993 level of protection. All or any combination of these factors will increase the stage for the areas they occur in.

3.3 Summary

The disaster risk management tools derived from geospatial technologies are widely varied. The two examples presented here represent only a small fraction of applications of geospatial technologies for assessing risk in urban environments. The ability to provide real time monitoring of hazards, like with the US Forest Service Active Fire Mapping Program and the ISO Properties, Inc. FireLine ™ system, is becoming a reality with technologies such as direct broadcast of satellite imagery and web mapping services for the near real-time delivery of derived products. The value of archive geospatial data sets with the integration of more current data is illustrated in the post-event assessment shown in the case study dealing with the 1993 flooding of the Upper Mississippi and Missouri River Basin. *Post-hoc* analyses such as these are important to conduct to determine

where policies regarding risk to the urban environment may not have been adhered to.

References

Carey, C., et al., 2003. A flood of development: building booms in the flood plain. *Saint Louis Post-Dispatch*, 27 July 2003.

Freeman, G.E., et al., 1993. The scientific assessment and strategy team contributions assessing the 1993 flood on the Mississippi and Missouri river basins 19 (4), 177–185.

Giglio, L., Descloitres, J., Justice, C. O., and Kaufman, Y. J., 2003: An enhanced contextual fire detection algorithm for MODIS. *Remote Sensing of the Environment*, 87, 273-282.

Heisler, E., 2003.Aflood of development: $400 million in investment flows into Chesterfield Valley. *Saint Louis Post-Dispatch*, 28 July 2003.

Hipple, et al. 2005 "Characterizing and Mapping Human Settlements" in M. K. Ridd & J. D. Hipple, eds, *Manual of Remote Sensing, 3rd Edition (4 Series) Volume 5, Remote Sensing of Human Settlements* , American Society for Photogrammetry & Remote Sensing, Bethesda, MD, pp. 149-206.

Hipple, J, B. Drazkowski, and P. Thorsell 2005. Development in the Upper Mississippi Basin: 10 years after the Great Flood of 1993, *Landscape and Urban Planning*. 72:313–323.

Hipple, J. and and T. Haithcoat. 2005. Remote sensing in urban infrastructure and business geographics, in S. Aronoff, *Remote Sensing for GIS Managers*, ESRI Press; Redlands, CA, pp. 405-413.

Holway, J.M., Burby, R.J., 1993. Reducing flood losses: local planning and land use controls. *Journal of the American Planning Association* 59 (2), 205– 216.

Interagency Floodplain Management Review Committee (IMFRC), 1994. *Sharing the Challenge: Floodplain Management into the 21st Century*. US Government Printing Office, Washington.

Leavenworth, S. 2004. *Sacramento Bee*, 12 April 2004, p. A1.

Leavenworth, S. 2004. *Sacramento Bee*, 29 March 2004, p. A1.

McNamara, D., G. Stephens, B. Ramsay, E. Prins, I. Csiszar, C. Elvidge, R. Hobson, and C. Schmidt, 2002: Fire Detection and Monitoring Products at the National Oceanic and Atmospheric Administration. *Photogrammetric Engineering and Remote Sensing* 68, 774-775.

Moon Jr., H.E., 1988. Modelling land use change around non-urban interstate highway interchanges. *Land Use Policy* 5 (4), 394–407.

National Aeronautics and Space Administration, 2005: *MODIS Fire and Thermal Anomalies Guide. Online technical reference* , accessed September 16, 2006. URL: http://modis-fire.umd.edu.

Pinter, N. 2005. One Step Forward, Two Steps Back on U.S. Floodplains. *Science* 308(5719): 207 - 208

Scientific Assessment Strategy Team (SAST), 1994. *Science for Floodplain Management into the 21st Century*. US Government Printing Office, Washington (preliminary report).

Shipley, S., 2003. A flood of development: Missouri lacks rules on flood plain growth. *Saint Louis Post-Dispatch* 27 27 July 2003.

WHO (1988) Basic terminology for risk and health impact assessment and management. Internal Report of Working Group, 24-25 March, 1988, Geneva, World Health Organization (Annex 3).

Wachter, S.; L. Hirschorne; H. Sokoloff; and H. Steinberg 2006. *The Geospatial Industry: A perspective on Technology Diffusion*. The Association of American Geographers, accessed 2 October 2006; URL http://www.aag.org/roundtable/

Quale, B. 2005 Satellite based fire mapping for the Eastern United States. *EOM: Earth Observation Magazine* 14(6).

ISO Properties, Inc. 2004. *Pinpointing Insured Losses in the Southern California Wildfires of 2003*. ISO, Jersey City, NJ.

4 Intraurban Population Estimation Using Remotely Sensed Imagery

Perry J. Hardin, Department of Geography, Brigham Young University, 690 SWKT, Provo, UT 84602, USA

Mark W. Jackson, Department of Geography, Brigham Young University, 690 SWKT, Provo, UT 84602, USA

J. Matthew Shumway, Department of Geography, Brigham Young University, 690 SWKT, Provo, UT 84602, USA

Of the Earth's 6.5 billion human inhabitants, nearly three billion live in urban settlements (UNCHS, 2001). Natural increase, land tenure practices, political policy, environmental degradation, and the dynamics of regional / global economics are largely responsible for the ongoing population shift from rural agrarian regions to cities. This increased urbanization is not just a developing country phenomenon. Urban areas of North America in 1900 were home to only 50% of the continent's population. In 2000, the percentage of North American urban inhabitants rose to 75%.

Given the importance of urban regions as human habitat, there is an established need for accurate intraurban population counts to support decision making. In comparison to population estimates or projections, an exhaustive per-dwelling enumeration acquired through fieldwork is the accepted gold standard for counting people and determining their sociodemographic characteristics. A census is a complex undertaking; it requires significant human, technological, and fiscal resources to plan and execute. Because of high cost, industrialized nations conduct enumerations only periodically. The American decennial census mandated by the U.S. constitution is an example. The United Kingdom also conducts census surveys every ten years. The Australian Bureau of Statistics conducts a census every half decade. In contrast, developing nations experience almost insurmountable obstacles to obtain accurate and regular enumerations of their national population. These include vast rural areas, nomadic populations, func-

tional sovereignty limitations, cultural mistrust, and a lack of technical and financial wherewithal.

The Decennial Census of the US is intended as a temporal snapshot; a depiction of the national population on April 1 of the census year. In areas of rapid urban growth, the population counts recorded in a decennial census become progressively less representative as the decade progresses. Recognizing this, intercensal population estimations are commonly required. Additionally, small area population estimates provide a key source of data for local planning agencies and businesses. Often the data are required at a geographic level smaller than what is easily found in census data. In these cases, population estimates are the most cost-efficient way to generate the required small area data.

In this chapter we will first briefly review the traditional methods for population estimation. Three broad methodologies for estimating intraurban population totals and densities using overhead imagery will then be discussed. Our focus is on the developed urban world rather than developing or rural areas. Remote sensing may provide a vital role for population estimation in developing countries with significant rural regions, but it is not our expertise.

Once the three broad methodologies have been appraised, a short case study will be presented. In this case study, we use rudimentary image processing techniques to estimate the population of the Wasatch Front urban corridor in Utah, U.S. After the case study, some concluding comments are then offered about future research directions.

4.1 Traditional Approaches to Population Estimation

Population estimates should not be confused with population projections. Although data and methods may differ, the primary difference is one of time period. Population *estimates* are used for the present and the past, whereas population *projections* are used to guesstimate future population size. In this chapter, our focus is on population estimates.

Estimating population of small areas at various scales of space and time is a difficult demographic task. However, because small area population estimates are often necessary for local planning departments and businesses (billions of dollars in federal funds are allocated to states and local entities

based on the estimates), there have been several estimation methodologies developed. Regardless of the method, four preliminary factors must be considered; 1) the purpose of the estimation, 2) the spatial scale of the estimation, 3) the target temporal period for the estimation, and 4) data availability. Once those factors have been considered, three other issues must likewise be addressed; 1) the collection of any necessary data, 2) the selection of correct statistical methods, and 3) the method for judging the goodness of the estimates.

The collection of necessary data is framed by the purpose and geographic scale of the required estimate. At smaller geographic scales certain administrative records are aggregated at a county or metropolitan scale and may not be available at sub-county levels (e.g. tax returns). Deciding the scale and purpose of the estimate often determines what data can be used, which in-turn constricts the choice of appropriate methods. However, almost all traditional population estimation techniques use various types of administrative records that are correlated with population change. Predictors derived from these records are called *symptomatic variables* (Plane and Rogerson 1994). Ideally, symptomatic variables should be updated regularly. They should also temporally co-vary with population change in a predictable fashion. Exemplar symptomatic variables include residential building permits, utility connections, school enrollments, tax returns, and Medicare enrollments. A second important data source is vital records – particularly birth and death certificates. These data are used as major inputs into the cohort-component method of population estimation described below.

The U.S. Census Bureau, in cooperation with state partners, is legally required to provide intercensal population estimates to support federal fund allocations. To comply, the Census Bureau has developed three principle methods:

1. *Ratio-correlation procedures*. As the name implies, ratio-correlation procedures use the ratio of symptomatic variable values for adjacent time periods as independent and dependent variables to estimate population. Changing ratios of symptomatic variables within a geographic region are assumed to be a function of the region's changing population ratio (Plane and Rogerson 1994). The ratio of a subregion's population to the larger region population for two time periods is regressed on similarly formed symptomatic variables (ratio of ratios). These models will generally use vital records (i.e., births

and deaths) and administrative records (e.g.,elementary school enrollment, vehicle registration, voter registration).

2. *Component-method II procedures*. All component methods are predicated on demographic accounting, i.e. population change = births - deaths + net migration. The greatest difficulty with this approach is correctly specifying the migration factor. The component-method II procedure utilizes registration data on births and deaths, and tries to estimate net migration using other administrative information. The U.S. Census Bureau splits the use of different administrative data genres based on age. For populations less than 65 years of age individual tax returns are used.[1] For populations older than 65 years of age, medicare enrollment is used. In this method, migration – based on different tax return addresses or medicare payment addresses – is estimated. Both sources also have information on household size. For entities without access to individual tax returns, school enrollment is often used and assumed to be indicative of migration in the total population – with adjustments being made for the historical differences between the school-age migration rate and the total population's rate of migration.

3. The *housing-unit method* is based on change in the housing stock of an area from the base date to the estimate date. Data on the housing stock and flow are generally derived from; 1) U.S. Bureau of the Census survey of building permits and demolitions, and 2) State Data Center survey of counties and cities issuing permits for residential buildings and demolitions. The housing unit method requires the specification (assumption) of vacancy rates and average household size. Once specified, housing unit count change between base and estimate dates is multiplied by the occupancy rate and average household size to estimate population change. Individuals in group quarters (prisons, college dormitories, nursing homes, and military barracks) are included in the total. As a refinement, separate estimates are constructed by housing structure type (e.g., single-family dwellings, 2-to-4 unit, 5+ units, mobile homes). This refinement permits different vacancy and household size factors to be more precisely tailored to the structure types within the housing stock.

1 Only the Census Bureau has access, from the IRS, to the individual tax returns. Other governmental or private businesses will use school enrollment data in lieu of individual tax returns.

Perhaps the most important aspect of population estimation is validation. Population estimates are typically based on assumptions of temporally stable relationships between population change (i.e., births, deaths, migration) and their symptomatic variables. The temporal stability of the symptomatic variables themselves is likewise assumed. These are generally safe assumptions, but the estimate will always contain error. The only way to assess the error is to do an actual count. This would of course obviate the need for the estimate in the first place.

4.2 Population Estimation Using Remote Sensing

There are four primary approaches to estimating population with remotely sensed data:

1. The use of allometric population growth models based on place size. Typically the area of cities, towns and villages is measured from small scale air photography or satellite imagery and submitted to a calibrated allometric model to estimate population. Central place theory and road connectivity are sometimes employed to improve accuracy. The allometric technique is very useful in developing countries where ground enumeration is impossible and a single population total for each city, village, or region is acceptable. It is less useful when population estimates are required for small enumeration districts such as US census tracts. This method is beyond the scope of this paper but Lo (2006) provides an excellent review.

2. The use of dwelling unit type as a surrogate for family size.[2] This technique requires an interpreter to identify, classify, and count dwelling units manually from large scale imagery. A simple model relates dwelling type to resident family size.

3. The use of landtype zones[3] as a surrogate for population density. Different landtype zones are identified on medium-scale imagery and

[2] With some loss of precision, we call this approach *dwelling identification.*

[3] Every student of remote sensing quickly learns the difference between land use and land cover. Nonetheless, to simplify phraseology by avoiding the repetition of the phrase “land use and/or land cover” throughout this chapter, the term

a model is employed that links landtype with population density. This approach has historically been used with medium scale air photography and satellite imagery.

4. Pixel based approaches that seek to model population or housing unit density directly as a function of spectral reflectance or spectral texture on a medium-scale satellite images.

4.2.1 Dwelling Identification

The process of estimating population density via dwelling identification is conceptually simple and requires the following general steps. A schema of dwelling unit types based on family size is initially itemized. For example, the schema may include designations such as duplexes, single family residential homes, and apartments. Using information acquired from census data, interviews, or rental agencies, average resident counts for each dwelling unit type in the schema are determined. Each dwelling unit in the study area is then placed into one of the *a priori* schema classes by its appearance on large scale photography. Total estimated population is the sum of the dwelling units of each type weighted by their corresponding average resident population.

The success of the procedure outlined above depends on the successful identification of various dwelling types from high-resolution imagery. This was established early by Green (1956, 1957) who postulated that the social structure of a city could be determined through the analysis of aerial photography. Green suggested that this identification of dwelling type is the first step to the use of air photography for demographic, sociological, and urban ecological applications. In this pioneering research, Green (1956) examined 17 residential neighborhoods in Birmingham, AL to ascertain whether stereo air photography (1:8,000 scale) facilitated the identification of urban dwelling structure type. Although the research focused on measuring; 1) the percentage of detached single-unit homes, and 2) the dwelling unit density per block, several other residential structural types were discriminated as part of the study (e.g., duplexes, multiunit). Green utilized the following characteristics in his photographic key to housing identification:

"landtype" will be used instead. Where differentiating between land cover and land use is important to the discussion, the two separate terms will be used.

1. Roof structure and form (i.e., gables, dormers, porches) including the existence and number of chimney stacks and rooftop plumbing fixtures.
2. The overall shape and size of the building.
3. The situation of the building i.e., "the location of the building with respect to the street, the building line, and other structures" (p. 143).
4. Vehicle accommodations, including carports, parking areas, garages, and driveways.
5. Pedestrian accommodations such as footpaths, sidewalks, and entryways.
6. The shape and size of yards, courts, etc.

Not surprisingly, Green found that the error rates in classification were dependent on structure type rather than uniformly distributed across all structure classes. Nearly all the problems involved multiunit residences.[4] Specifically, Green had trouble with universally distinguishing multiunit complexes from duplexes, and differentiating between duplexes and single-unit dwellings. Summarizing, the total study area housing unit count and multiunit structure count were slightly underestimated whereas the single-unit dwellings were overestimated. Overall however, "the results [showed] 1) that 99.8 percent of the 3,629 existing residential structures in the 228 city blocks observed were correctly identified as such, and 2) that 89 percent of these structures were correctly classified by categories of numbers of dwelling units." For other related dwelling unit studies akin to Green (1956), see Hadfield (1963) and Binsell (1967).

Extending Green's groundbreaking research, the objective of Lindgren's (1971) study was to determine; 1) whether the same dwelling unit identification success reported in foregoing research could be obtained with medium-scale imagery (1:20,000), and 2) whether the use of color infrared (CIR) photography improved dwelling type identification success rates obtainable from natural color or panchromatic imagery. Lindgren's operating assumption was that "in high-density areas, CIR imagery would allow for easier identification of urban signatures" (p. 374). Although originating with Binsell (1967), Lindgren's final list of dwelling identification keys deviates little from Green (1956).

After developing the clues using three blocks of high-density housing in the metropolitan Boston area (i.e., East Boston, Chelsea, Charlestown), the

[4] The identification and treatment of multiunit structures is a reoccurring theme in population estimation with overhead imagery.

indicators were tested in 15 additional blocks containing 655 residential structures and 1744 dwellings. Lindgren's total residential structure count from the photography underestimated the actual total by only three. However, the ability to count infrastructure dwelling units was more difficult because of the prevalence of multiunit buildings within the study area – the interpretation underestimated the total number of dwelling units by 54. These summary figures obscure the seriousness of some interpretation problems – in both residential structure and dwelling unit counts, significant overestimates were offset by substantial underestimates. Overall, only 59% of the residential structures were assigned the correct number of dwelling units.

In concluding, Lindgren offered two observations. First, any personal familiarity of the study area enjoyed by the interpreter would dramatically increase chances for a successful outcome. For example, structures that Lindgren found in Charlestown with a particular roof-type were consistently mis-categorized. A single visit to Charlestown before the interpretation began would have prevented the mistake. Second, the high quality of the CIR transparencies used (i.e., their sharpness and contrast), in conjunction with the infrared distinction between built-up and nonbuilt-up urban areas more than counteracted any disadvantage of the small CIR image scale.

In coincident research, Collins and El-Beik (1971) used dwelling identification methods to estimate the population of the City of Leeds. The goal of the study was to determine whether population estimates made from air photography agreed with census estimates. Like researchers before them, the operational hypothesis was that dwelling type was strongly correlated with resident population count. Given earlier work by El-Beik (1967) demonstrating that housing types in Leeds could be identified classified from air photography, that hypothesis was reasonable.

The schema for the population estimation study required the discrimination of semidetached, terraced, and back-to-back dwelling types. Based on the interpretation of 1963 photographs of 1:10,000 scale, all the housing structures within the study area were classified into one of those three categories. Multiplication factors linking dwelling type to inhabitant number were derived from 1961 census enumeration maps and data. Only half of the enumeration districts were used to derive these factors whereas the other half was cloistered for validation purposes. For semidetached dwellings, it was found that 3.03 people lived in each house. The linear nature of semidetached and terraced dwellings in Leeds suggested a different tack

for designating a multiplier. For these structures the multiplier was the total housefront linear feet per person with 4.62 and 4.27 ft / person being adopted for terraced and back-to-back houses respectively.

For validation, these factors were applied to the housing within the sequestered enumeration districts and population totals were calculated. When compared against 1961 census data, Collins and El-Beik found that the average error for enumeration districts dominated by terraced houses, back-to-back houses, and semidetached dwellings was +0.87%, +0.32%, and -6.4% respectively. The largest errors among individual enumeration districts were underestimates among semidetached homes with unexpectedly large families.

According to Collins and El-Beik, the accuracy of this approach depended primarily on two variables. The first was the ability of the photointerpreter to properly identify housing type. The second was how well the calibrated multiplier correctly represented target areas of the same dwelling category. The authors also observed that refinements in the method were possible. Multipliers could be adjusted according to structure age and proximity to schools. Social and economic variables could likewise be used to create a more sophisticated multiplier set.

Watkins (1984) focused his research on the problem of correctly counting the number of dwelling units in a multiunit structure. Error resulting from this prevalent problem was also investigated – Watkins observed that "no studies to date have explicitly investigated the nature of multiple dwelling unit counting errors with respect to the ways in which they relate to different structure types, nor have they considered the actual impact that multiple unit structures as a whole have on the accuracy of enumerations of all dwelling units within a residential area" (p. 1599).

Watkins subdivided multiunit structures into two groups; 1) those originally designed to house multiple families, and 2) structures originally built as single-family dwellings. Watkins observed that the diagnostic photographic elements needed to estimate the number of households were different in each group. The photographic key developed included not only guidelines for differentiating between residential and nonresidential but instructions for discriminating between converted single-family structures and archetypical apartment structures. Telltale features of apartment buildings included roof divisions, outside fire escapes and porches, entrance location and number, parking, and apparent socioeconomic level. Converted structures were distinguished by structure symmetry, quality

and amount of vegetation, number of sidewalks from the structure to the roadway, walkways between sides of the same structure, site context, and the apparent method of property subdivision. Structure size, shape and height, as well as vents and chimneys were important to identifying both types.

After developing the photographic key employing these factors, Watkins conducted tests in Boulder, Colorado for three study areas. Dwelling count estimates made from 1:20,000 panchromatic imagery acquired in 1970 were compared to 1970 census block dwelling counts. Similarly, dwelling counts from the 1980 census were compared to estimates derived from 1979 1:6000 scale panchromatic photography. The photographic key was highly successful. For the 1970 data involving 695 buildings, errors in multiunit counts within the three study areas ranged from an underestimation of 1.61% to 0.37%. The error rates for 1980 (2545 buildings) were significantly higher, ranging from 1.64% to 4.91%. As hypothesized, error rates differed by multiunit structure type. Converted single dwelling units were overestimated by 8.45% whereas dwelling units within traditional apartment buildings were underestimated by 5.51%. Single units (unconverted) were underestimated by 4.80%.

Lo (1986) stands alone as the only researcher who has actually applied the dwelling unit identification method to an entire city. The goal of Lo's research was to estimate population in 93 traffic zones in Athens, Georgia from 1:20,000 aerial photography. Like preceding researchers, Lo used a simplified residential structure schema that included only a few structural types; 1) small single family structures, 2) large single family structures, and 3) multifamily structures. Estimated resident counts used in the population calculations were 3.0 and 4.0 for small and large family dwellings respectively, and 2.0 per dwelling unit within multifamily structures.

Comparison of the photointerpretation results with 1980 census data revealed an average population count underestimate of 1.7% per traffic zone with considerable variation from zone to zone. Counting errors were attributed to the following factors; 1) family sizes different from those assumed in the estimation process, 2) photointerpreter skill, 3) the number of multiunit structures, 4) the area of the traffic zone, and 5) the quality of the photographic source. Lo demonstrated that urban population estimation for an entire city was feasible, and resulting accuracy could be high. We consider Lo (1986) to represent the state-of-the-art in the dwelling unit identification / counting approach to intraurban population estimation.

Comments

Although the identification and counting of dwelling units showed significant promise for intraurban population estimation, there is little literature mentioning it after Lo (1986). Of all three methods for intraurban population estimation discussed in this chapter, it remains the most accurate, particularly when enumeration districts are small. In addition, there is increasing availability of high resolution digital orthophoto and high altitude CIR coverage for many urban areas in the U.S. that can be analyzed with the method. These data can generally be downloaded online from state or university GIS repositories without cost. In our experience, these data range around 30 to 15 cm in pixel resolution, are geocorrected to a map base, and are excellent quality. These characteristics make the interpretation of most dwelling diagnostic features a straightforward task.

These high resolution digital image data have some limitations. One limitation is the inability to view the photographs in stereo. Because of this, other clues such as shadows must be used to measure building heights. In addition, the end user has no control over the date of the photography. This not only includes year, but season as well. Figure 1 is a medium-density residential block near downtown Salt Lake City, Utah. The photography (originally in color) was acquired in late summer 2003 as part of a USGS program to acquire high resolution imagery of the most populated urban areas in the United States. The data are available to the public from the U.S. Geological Survey, EROS Data Center, Sioux Falls, SD (http://www.usgs.gov/).

Figure 1. Leaf-on 30cm aerial photo of Salt Lake City, UT.

The pixel resolution is 30 cm. Figure 2 is a similar block in Pittsburgh, Pennsylvania, taken in leaf-off conditions. Classifying the buildings in the Pittsburgh imagery would prove a simpler task because the deciduous trees are not obscuring dwelling yards, roofs, etc. In contrast, the interpretation of the leaf-on Salt Lake City imagery would be more challenging. Of course if tree cover were an important diagnostic feature in a landtype schema, leaf-on imagery may be preferred.

Adeniyi (1983) summarizes our viewpoint on the dwelling unit identification approach when writing "The results have revealed, in general, that remotely sensed data have the capability to provide timely, verifiable, and relatively accurate intercensal population data, based on uniform criteria at local, metropolitan, and regional levels" (p. 546). Adeniyi is equally per-

ceptive regarding the method's greatest limitation when stating, "[Each]

Figure 2. Leaf-off 30cm aerial photo of suburban Pittsburgh, PA.

be valid only for the particular area under consideration. Consequently, there seems to be a need to formulate for each cultural area a suitable model based on relevant attributes of the area" (p. 546).

4.2.2 Landtype Surrogates

Along with the lack of generality in interpretation keys, another drawback to counting dwelling units is the intensive labor required to do the photo interpretation. The method also requires the use of large scale photography and is not suitable for use with satellite imagery having a pixel resolution coarser than about 1 meter. Because of these shortcomings, the use of landtype zone as a population surrogate was developed. The premise of the approach is that landtype is related to housing densities which are instage coupled to population density.

Two variants are common. In the first variant, a schema of landtypes is developed and a characteristic population density for each type is assumed, measured, or estimated. The landtype zones are delineated on overhead imagery. Given that a district can contain different landtypes, the population for a district is next estimated in three steps. First, the area of each landtype in the district is determined. Landtype area is then multiplied by the population density assumed for the category. The resulting product is the total population for that landtype within the district. Summing those totals across all the landtypes in the district produces the total district population. In the second variant, the imagery of the enumeration districts is investigated to determine the different landtype percentages constituting the district. Those percentages become carrier variables in transfer equations relating the landtype constituent amounts to the district population density.

The need for accurate population data with which to calibrate the landtype → population transfer functions is the logical equivalent of calculating the average number of people per dwelling unit type in the structure counting approach. It is generally required for both variants.

Kraus et al. (1974) were some of the first researchers to advance the landtype surrogate approach for population estimation. In their experiment, four cities in California (Fresno, Bakersfield, Santa Barbara, and Salinas) were chosen for study. The goal of the research was to estimate the cities' population from high-altitude aerial photography with scales of 1:600,000 (panchromatic), 1:120,000 (CIR), and 1:60,000 (CIR). The interpretation schema utilized only four land use types; single family residential, multifamily residential, trailer park residential, and commercial / industrial. The entire built-up area of the four cities was placed into one of those four classes based on landscape appearance in the photographs. The area of each land use type was measured with a polar planimeter. "In order to ob-

tain characteristic spatial population densities for the three residential land use categories, 1970 U.S. Census Block Data was used. Areas of a single residential land use were identified from the land use maps and located on Census Block data maps. Random samples of blocks within each residential land use category were then obtained to determine population densities per square kilometer for that land use. The spatial population density figures derived from each random sampling were then averaged to obtain characteristic spatial population densities for each residential land use category within each city" (p. 39). The total population for one of the mapped zones was the product of the characteristic population density and zone area. Summation of the zonal population across the city produced the total population estimate.

The results of the zonal procedure were a 7.0% population overestimation in Santa Barbara, and an average 7.2% underestimation in the other three cities. Two causes of the underestimation in Fresno, Salinas, and Bakersfield were given. The first was the inability to identify residences in older built-up business districts. Secondly, the enlargement of the original photography for easier interpretation unexpectedly increased the difficulty of identifying isolated individual apartments, causing an underestimation of area in the multifamily residential class. The overestimation in Santa Barbara was primarily due to the large lot sizes in many single-family residential zones – the characteristic population density applied to those zones in Santa Barbara was too large. To ameliorate these problems, Kraus et al. recommended; 1) a correction factor for "hidden" residential uses in commercial districts, 2) the use of larger scale air photography, and 3) a refined residential land classification system that permitted fine tuning of characteristic zonal population density factors.

In research reported by Adeniyi (1983), the objective was to examine the feasibility of systematically estimating Nigerian population with aerial photography. The research was warranted by the historic failure to accurately estimate Nigerian population using traditional methods, a failure attributed to the paucity of accurate social and administrative data. Additionally, urban planning demands of Nigeria required population estimation for small areas (e.g., voting districts) undergoing rapid urbanization. Based on significant preceding research (Adeniyi 1976, 1980) the project began with the simple hypothesis that population estimation based on land use zones would be appropriate for Nigeria. Two reasons were given for the use of the landtype zone method in preference to the dwelling unit method described above, both related to the communal housing structures used in Nigeria. First, the individual dwelling units were not readily

countable on the available air photographs. Second, because each single storied structure might house between three and ten families of considerable and variable size, the utility of an "average family size factor" was dubious and would have produced significant estimation error.

Adeniyi chose the Federal Capital of Nigeria (Lagos) as the study site. The methodology was complex, required several steps, and will only be summarized below. Initially, different residential areas were delineated and classified on the 1:20,000 panchromatic photographs acquired in 1974. Housing quality and other sociocultural information was next collected from the air photographs. These photographically measured variables included building density, plot size, layout, garden existence, number of stories, dwelling type (e.g., apartment, communal), and building usage.

Examination of these variables suggested a landtype schema with nine species. Once the schema was completed, the analysis required that the population densities of each residential landuse category be gauged. Using a random sampling scheme stratified by landuse category, information on family size and number of families for each residential structure type was determined by limited field survey. The total number of field samples (i.e., residential blocks) was 58 with the number of samples per land use ranging from 1 to 20. A total of 3,479 buildings were included in the 58 blocks. A cluster analysis of these 58 field samples using the field data alone regrouped the field samples into nine temporary subsets for the purpose of exploring intraresidential and interresidential class differences and similarities. The clustering exposed two broad divisions in the land zones. The first division was the planned land zones – planned residential areas with apartment housing of moderate density. The second division consisted of the higher density communal dwelling structures considered traditional and unplanned. It was also observed that many of the residential landtypes could be distinguished almost completely by the density of residential buildings typifying them.

Average population density figures were calculated from the field data for each residential land use. Regression analyses were then used to optimally model the population density of each land use using all of the survey and air photograph variables as candidate independents. For the complete sample of 58 blocks, three variables were able to explain 90% of the variation in the population density; 1) density of communal type buildings, 2) average population per building, and the 3) density of all buildings. Different residential land use classes had different models; however the two

variables appearing the most frequently in the models were density of communal type buildings and average population per building.

Because the data gathered had been exhausted in the model building phase, validation of the models was impossible, but Adeniyi made the following observations.

1. The Lagos population estimates in zones of planned residential development were quite accurate, primarily because the average population density value for the planned residential land uses was uniformly applicable to all samples of that type. If air photos were used operationally to estimate Nigerian population, fieldwork to establish the average building population for planned zones would be minimal.

2. In contrast, fieldwork to support estimates for the unplanned residential landuse categories would be necessarily extensive. The highly variable average population density factor would require tailoring for different regional areas.

Another example of the residential landtype approach with a simpler methodology is Olorunfemi (1984). This study was conducted in the city of Ilorin, Nigeria, a city of ½ million population that serves as the capital of Kwara State. The goal of the study was to "define a mathematical model which may be used in conjunction with data on housing land area measured from aerial photography to obtain urban population estimates for Nigerian cities" (p. 221). The photographs for the study were acquired in 1950 and 1963 at scales of 1:2400 and 1:12000 respectively. Census data to support the research was taken by survey method in 1952 and 1963.

A total of 74 square sample sites of 4 hectares each were randomly selected from topographic maps and their location transferred to the two sets of aerial photography. The area (percentage) of each major landtype within each sample site was measured using a dot grid. The landtype classification schema used six categories; 1) indigenous residential type housing, 2) barrack / flat housing, 3) flat housing, 4) uncompleted housing, 5) bare ground / grass / agricultural land, and 6) trees. The population of each sample site was determined from the census data. Multiple regression analysis was conducted to model population within the 4 hectare sample sites as a function of the area in each of the six categories.

For the 1950 case, population was significantly correlated with the percentage area of uncompleted housing ($r = 0.66$)[5], the percentage of trees ($r = -0.66$) and percentage area devoted to barrack / flat housing ($r = 0.57$). A linear regression model using all six variables explained 72.6% of the variation in population density. For the 1963 data set, the only significant explanatory variable was the percentage of area devoted to flat housing ($r = - 0.51$).

Olorunfemi concluded there was a functional relationship between landtype and population density which justified the use of landtype as a population surrogate. The goodness of the regression models was deemed sufficient to warrant an examination of its utility for wider application in Nigeria. "It should be stressed, however, that, for this method to be useful in generating nationwide data, there is need for further research aimed at testing the applicability of the model in cities with similar and/or different characteristics" (p. 227). Olorunfemi also considers the method particularly appropriate in communities where housing land area is known beforehand or population data is unavailable because of "remoteness, political obfuscation, or insufficient resources to conduct frequent census enumerations" (p.227).

Since the launch of Landsat-1 in 1972, the mapping of landcover from medium resolution satellites has become operational in many disciplines such as range management and agriculture. The potential for adapting the landtype zone population estimation method from aerial photography to satellite imagery was natural. The work of Langford et al. (1991) serves as an excellent example of this adaptation. One objective of this study was to model the 1981 population of 49 wards in four districts of northern Leicestershire, England using the landtype surrogate approach. The methodology was straightforward. First, using automated image processing methods, a satellite image dated July 1984 was classified into various landtype categories on a pixel-by-pixel basis. The single satellite image covered all 49 wards of the study area. Creation of the landtype map proved challenging, primarily because the unsupervised classification highlighted land cover differences in the rural hinterlands of the area but did not sufficiently discriminate between important urban landtypes. After some trial and error, principal components analysis of the original seven band TM image resulted in three principal component bands that revealed urban land cover differences necessary for accurate discrimination. The resulting map consisted of 12 landtypes collapsed into five broader categories; 1) commer-

[5] Unless otherwise noted, correlation coefficients (r) refer to Pearson's r.

cial and industrial, 2) high-density residential, 3) ordinary residential, 4) quarries, woods, water bodies, and 5) agriculture. Once the landtype map was complete, the 1981 census ward boundaries were digitized, rasterized, and imposed on the digital landtype map. This permitted the area of each type (listed above) to be tallied for each of the 49 wards. Except for agriculture, the greatest difference in cover amounts between the various wards was in the ordinary residential category – ordinary residential land cover within a ward varied from 10 to 430 hectares with a mean and standard deviation of 158 and 95 hectares respectively.

Simple correlation analysis revealed that total ward population was most highly correlated with the ordinary residential (r = 0.75) and commercial / industrial (r = 0.60) land area in the ward. Total population was more weakly correlated with agriculture (r = -0.28) and high-density residential (r = 0.33) area. Encouraged by the results of the correlation analysis, several regression models were created that explained total ward population as a function of the five independent variables listed above. The models differed primarily on the number of variables used, the use (or not) of a Poisson error term, the fitting (or not) of an intercept, and whether negative coefficients were permitted. The last requirement (no negative coefficients) was designed to preclude models that might generate negative population counts. Regression equations that had a non-zero intercept were also considered logically flawed.[6] In summary, it was "argued that that any statistical model linking pixel counts of land cover to population should be simple, linear, additive, and without any intercept constant" (p. 67). The most effective model produced by Langford et al. included only two variables, ordinary residential landcover and high density landcover area. This ordinary least squares model with only additive coefficients produced an R^2 of 0.82. Areas of underestimation and overestimation, sometimes severe, were noted by spatially analyzing the residuals from the regression. The residual patterns from the different models were similar.

Comments

Although the required methodology differs slightly, the use of landtypes as population surrogates is equally applicable to air photos and satellite data alike. The use of satellite data has some particular advantages. Unlike analog air photography, satellite imagery lends itself to automated interpretation, classification, and georeferencing. For large areas with no pre-

[6] Harvey (2002a; p. 2086) discusses the issues of negative coefficients and zero intercepts in some detail and provides alternative opinions.

existing air photo coverage, purchasing satellite data is much less expensive than contracting for an aerial photography acquisition mission. For example, one Landsat image covering about 34,000 km^2 can cost less than $1000 U.S. Because of its lower expense, satellite imagery can be acquired more often (or on demand) and can thus facilitate more timely population estimates than air photography. Although high resolution satellite data are available, the landtype surrogate approach only requires less expensive moderate resolution imagery (i.e. 30m - 50m pixel resolution). Satellite data can also provide a fine degree of spatial granularity over large urban areas.

The landtype surrogate approach does have disadvantages. It is less accurate than counting dwelling units. Although a landtype classification schema may only contain five or six categories, the task of creating the landtype schema is critical for success. As illustrated by Langford et al. (1991), the task of generating an urban landtype map with sufficient detail to support population estimation may require substantial trial and error. Nonetheless, the use of landtype surrogates continues to be an important tool in geographic urban analysis.

4.2.3 Pixel-based Estimation

Pixel-based estimation is an approach designed entirely for moderate resolution satellite imagery. In its basic form, the goal of pixel-based estimation is to model population or population density directly as either a function of multiband satellite sensor reflectance values or some mathematical derivative thereof.[7] Adopting the logic of Iisaka and Hegedus (1982), the justification for this approach lies in the nature of the pixel itself. In an urban area, a single satellite pixel will contain a variety of land cover types that contribute to the spectral reflectance of the pixel. Figure 3 shows an area of several pixels covered by a Landsat MSS image compared to an aerial photo with 6 inch resolution to demonstrate the variety of objects within a single 79 x 79 meter pixel. Exemplar cover types are rooftop shingle, road surface concrete, lawn grass, and parking lot asphalt. As population density varies, the relative percentage of these cover types covaries. This variation accordingly modulates the spectral signature of the pixels having an urban footprint (Iisaka and Hegedus, 1982). While the causal relationship may be population → housing → landtype → spectral

[7] The term “spectral features” is given to these mathematical derivations. It should not be confused with “spatial features.” Example spectral features include vegetation indices, principle components, and texture measures.

signature, the pixel based estimation usually proceeds spectral signature → population with the landtype and housing treated implicitly.

Figure 3. Landsat MSS image compared to aerial photo of Provo, Utah.

The goal of Iisaka and Hegedus (1982) was to model the relationship between spectral reflectance and the population of metropolitan Tokyo. The satellite images acquired from Landsat 1 and Landsat 3 were dated November 1972 and January 1979 respectively. Census data required to calibrate and validate the modeling were acquired in 1970 and 1975. One of the assumptions required in modeling population as a function of spectral signature "is that the environmental alteration…should share similar characteristics in different areas, in both quantitative and qualitative respects" (p.261). This permits measurements made in sample areas to be logically applied to other locales. The authors claimed that the homogeneity of housing materials, dwelling size, land use systems, and housing density in residential Tokyo satisfied this assumption.

A total of 88 sample sites (25 hectares each) outside the Tokyo central business district were initially identified. The size of the sample sites (500 m × 500 m) corresponded to the resolution of the government census maps of the Tokyo ward area. Once identified, the population for each sample site was extracted from the 1970 and 1975 census maps. The satellite imagery was resampled and georegistered to the census maps and the mean

spectral values for the four multispectral scanner (MSS) bands determined for each of the 88 sample sites. This created a data set of 88 records having five fields each. Population was the dependent variable whereas the carrier variables were the four MSS bands.

Exploration of the data indicated that the green and infrared bands were strongly correlated (linearly) with population. Regression analyses were conducted to determine whether population for the sample sites could be explained as a function of the four-band spectral signatures. As expected, regression equations utilizing the green and infrared bands were most capable of predicting population density. The signs for the coefficients remained the same for the 1972 and 1979 data although the magnitudes of the coefficients were different. Examination of the regression equations showed reflectance in the green band increasing and reflection in the two infrared bands decreasing with increasing population density. Iisaka and Hegedus do not explain the physical basis for the signs of the coefficients or the reason why they are different between the two years. Based on our own research in North America, we cautiously suggest that denser urban build-up associated with greater population densities co-occur with less urban vegetation, hence the inverse relationship with infrared reflectance.[8] At first blush the same argument might also suggest the same inverse relationship between population and green reflectance. However the increased green reflectance from concrete and other lightly colored inert materials more than compensates for the loss of green reflectance from sparse urban vegetation in such situations. Likewise, we suggest that the difference in the reported regression coefficient magnitude between the two years *might* have been reduced by employing radiometric / atmospheric correction and standardization methods which have become common in Landsat data processing since that time (e.g. Singh, 1989).

Overall, the models generated multiple R values of 0.84 and 0.77 for the 1972 and 1979 studies respectively. Judicious removal of a few atypical sites improved the multiple R values to 0.94 (n = 60) for 1972 and 0.90 (n = 62) for 1979. Nonconformant sites were residential sample sites containing train stations, schools, churches and other features not prevalent in the Tokyo residential area.

The next important milestone of intraurban population estimation research was reached by Lo (1995) who rigorously evaluated the use of SPOT im-

[8] An increase in shadow from increased urban “canyonization” is an alternative explanation.

agery for population estimation[9] using methods predicated on Iisaka and Hegedus' (1982) earlier success. Whereas the Landsat MSS used by Iisaka and Hegedus (1982) utilized four spectral bands with a pixel size of 79 m, the SPOT imagery used by Lo had three bands (i.e., green, red, infrared) and a pixel resolution of 20 m. When the project was initiated, it was thought that the smaller SPOT pixel size would be a decided advantage, "Because of the low spatial resolution of the Landsat-MSS data, the spectral radiance is the average of reflectance of different cover types over an area of 79 m by 79 m in a pixel. The spectral reflectance of the residential cover, on which population estimation has to be based, is therefore highly diluted. This dilution will likely affect the accuracy of the population estimates. … Because of its better resolution, each SPOT image pixel covers a much smaller area on the ground, and hence the spectra radiance is more representative of its ground cover characteristics than the 79-m Landsat-MSS counterpart" (p. 18). In the exploratory work of Iisaka and Hegedus (1982) cited above no attempt was made by the authors to incorporate any *a priori* knowledge about the study area that might permit different regression models to be used in different neighborhoods. In contrast, the objective of Lo was to model population density in a metropolitan Hong Kong study area (i.e., Kowloon) as a function of SPOT spectral reflectance while employing GIS technology to permit different regression transfer equations to be used with different landtypes.

As mentioned by Lo, the mixed landuse in Kowloon was a significant challenge. Landuse was complex with residential areas intricately mixed with non residential areas. Transitions between low-density residential areas and high-density overcrowded areas were abrupt. Multistoried buildings in Kowloon not only housed multiple dwelling units, but the buildings themselves were multiple-use, serving commercial and industrial functions too.

Population data for Kowloon used to support the research was collected by the Hong Kong Census and Statistics Department in 1986 for 60 planning units via complete enumeration. Because of computer storage limitations, only 44 of the planning units were examined in the study. These ranged in area from 6 to 291 hectares. The SPOT data used for the population estimation was acquired in January, 1987. After significant preprocessing, the

[9] Lo (1995) also treated the counting of dwelling units with SPOT imagery. To simplify the review, we cite only the results for the population estimation component of the research. The dwelling unit estimation results closely paralleled those of the population estimation.

SPOT data was georegistered to the boundaries of the planning units. This permitted mean spectral reflectance values to be calculated for each unit. By calculating the area of each planning unit with the GIS, population densities for each were also determined.

Exploratory analysis indicated a moderate negative correlation (r = -0.62) between mean infrared reflectance and population density. After these initial findings, several regression analyses were performed to model planning unit population as a function of mean planning unit spectral reflectance. Twelve of the planning units were manually selected to calibrate the regression equations and the remaining 32 were reserved for testing regression goodness. Regression models were built using all three bands as well as simplified regression models using only the infrared band. The dependent variable in the regression model was population density rather than population counts.[10] The model using three bands was capable of estimating population of the whole study area with a relative error of 1.7%, whereas the infrared band model resulted in an error of 15.0%. When population estimation was attempted at the smaller scale of the planning unit, serious estimation errors sometimes exceeding 500% were encountered, primarily in commercial and industrial planning units. The mean relative error for the micro-scale planning unit estimation was about 75%.

Because each planning unit in both Kowloon and the greater Hong Kong metropolitan area was a mixture of both residential and nonresidential land uses, Lo sought to refine the population estimation process. The refinement required that the pixels actually representing residential land use be identified within each planning unit. This would permit the regression equations to be applied to those residential pixels alone and avoided the errors associated with attempting prediction for those pixels known to be predominantly commercial or industrial.

This refinement was completed by classifying the SPOT image into eight landtypes which included both low-density residential and high-density residential categories. An average per-pixel population density for high and low density residential zones was calculated using the census data for the 12 calibration planning units. The refinement produced a modest decrease in the population estimation errors of the smaller planning units, and the absolute mean relative error dropped to 67%.

[10] When multiplied by the area of a single SPOT pixel (0.04 hectares), the population per pixel could be easily estimated from the density.

For its rigor in methodology and actual success in modeling population density of small districts, Harvey (2002a) represents a landmark work of significant dimensions that should be read in its entirety. Harvey's goal was to model small-area population densities for Australian census collection districts (CDs) using spectral features measured from Landsat TM imagery.

Imagery of Ballarat Statistical District (west of Melbourne Australia) containing 138 CDs was used to build models of urban population density. Thematic Mapper imagery of Geelong Statistical District (225 CDs), nearly 100 km southeast of Ballarat was then used for model validation. Population data supporting the study was collected in 1986 by the Australian Bureau of Statistics and was preprocessed to correspond more closely to the February 1988 imagery date.[11]

Recognizing that "one obvious problem was that the values of the dependent variable ranged over three orders of magnitude" (p. 2079), both logarithmic and square root transformations of CD population density were calculated before multiple regression modeling began. A set of 80 predictor variables was submitted to a host of stepwise regression analyses in an effort to find the best predictive variable subset. These included the following classes of transformations, calculated for each CD by using the pixels captured within the digitized boundaries of the CD:

- Mean TM band reflectance
- The square of the mean of TM band reflectance
- The cross product of the mean TM band reflectance
- The ratio of mean reflectance for two TM bands
- The difference-to-sum ratio for two TM bands
- The TM band variance
- The TM band standard deviation
- The TM band coefficient of variation
- Normalized bands

The following spectral transforms were calculated on a per-pixel basis and then summarized for each CD by calculating means and measures of variation. These were numbered among the 80 multiple regression variables.

- Selected normalized bands
- Selected band ratios
- Difference / sum ratios of selected bands

[11] See Harvey (1999) for a discussion of this preprocessing.

- Hue transforms of selected bands using both rectangular and cylindrical coordinates.[12]

Summarizing Harvey's voluminous results, model R^2 values of about 0.90 were obtained with model subsets containing between four and nine variables. The best predictors were the mean and standard deviation of per-pixel spectral features listed above. The dependent variable transform associated with the highest R^2 regressions was the square root rather than logarithmic transform.

For validation, six multiple regression models predicated on the Ballarat study area were applied to the TM imagery of the Geelong Statistical District. Population counts and densities for the 225 Geelong Statistical District CDs were thus estimated and then compared against the true CD population counts. The two models based on band means alone produced very poor total population estimates. The model based on the per pixel measures listed above produced the best total population estimate in the validation CDs and had an urban total underestimation of only 3%. The following observations were made:

1. Models of high complexity performed better than simpler models.

2. Populations of lower density in rural areas were consistently and seriously overestimated. The errors were not large in terms of population numbers, but rather in percentages of the true population counts. Regarding this rural overestimation, Harvey comments, "It is concluded that the potential of this methodology is limited by heterogeneity of both land cover and population density within the individual CDs, and that are, in principle, unlikely using this approach. In particular, the sacrifice of detailed spatial information leaves no way to respond to the problem of over-estimation of population in large areas of low density" (p. 2093).

3. Given that models driven by per-pixel spectral indicators were superior to those calculated for the CDs (e.g., CD band means), Harvey conjectures that "models formulated [at a lower] spatial level can produce relatively accurate…population estimates for larger spatial aggregates, but not for spatial units at the same level of aggregation" (p. 2093). Harvey concludes that future modeling

[12] See Jensen (2005; pp. 164-167) for a discussion of hue transforms.

should be logically done at a lower level of aggregation than the CD to minimize errors in population prediction. He also points out the difficulty operationalizing this within his Australia study area – while spectral data may be available at the smaller pixel level, the census calibration data is available only at the larger level of the CD, and thus precludes lower level modeling. Methods to overcome this obstacle by exploiting an expectation-maximization statistical algorithm are presented in Harvey (2002b).

The primary objective of the research conducted by Li and Weng (2005) was to develop and compare methodologies for estimating the population density of Indianapolis, Indiana using Landsat ETM+ data. As a justification for their research, Li and Weng claimed that previous research "rarely [had] explored the integration of spectral, textural, temperature data, and advanced transformed remote sensing variables to estimate population. [13] Such incorporation may provide new insights for population density estimation" (p. 948). The ETM+ satellite data were acquired on 22 June 2000. The population data, based on census blocks, were obtained from a GIS vendor and aggregated into census block groups (CBGs).

Li and Weng's research objective required that several spectral features be examined to determine their correlation strength with population density. These spectral features included; 1) the first principle components of the ETM+ visible and optical infrared bands, 2) six different vegetation indices, 3) variance images (with various local window sizes) calculated from ETM+ red and middle-infrared bands, 4) surface temperature from ETM+ Band 6[14] and 5) impervious surface and green vegetation fraction images produced from decomposition of the six ETM+ visible and optical infrared bands.[15]

The study area consisted of 658 CBGs. An initial investigation required 162 samples for model building and the remainder for model validation. Like their predecessors, the samples used for creating the models were not entirely chosen at random. In this Indianapolis study, all the block groups with low and high population densities were included among the 162, whereas the medium population density CBGs were sampled randomly.

[13] Harvey (2002a) is an obvious exception to this generalization.

[14] See Weng et al. (2004) for detailed information on how surface temperature was derived.

[15] See Lu and Weng (2004) for more explanation about the fraction images.

Simple correlation analysis was used to explore the relationship between population density and the spectral features mentioned above. The simple Pearson's *r* showed only a weak correlation between population density and the spectral bands, principle components, vegetation indices, fraction data, and texture – *r* never exceeded 0.4 in absolute value. The correlation between temperature and population density was moderate ($r = 0.519$). It is notable that while the correlation between population density and the middle-infrared band was nearly zero, the correlation between the texture calculated in that band with a 7 x 7 window was substantive ($r = -0.402$). It was also found that transformations of the dependent variable (i.e., natural logarithm, square root) improved correlations moderately. The best multiple regression equation using a subset of the spectral features generated an R^2 of 0.83. A residual error map showed the greatest misestimation occurring in CBGs of extremely high and low population density.

According to Li and Weng, "In order to improve population estimation results, separating the population density into sub-categories such as low, medium and high densities, and developing models for each category" (p. 952) was deemed a necessity. However, the results of using separate regression models for the three strata of population density were mixed. In general, stratification improved the results, nonetheless, with R^2 values for low and high density population density CBGs never even reaching 0.2, Li and Weng questioned whether Landsat ETM+ data was suitable for modeling extremes in population. For medium density population, the results were much better with R^2 approaching 0.90. The best predictor variables for the medium density models included red band texture (7 x 7 window), thermal temperature, the simple ratio of the near infrared and red bands, the transformed normalized difference vegetation index,[16] the soil adjusted vegetation index,[17] infrared reflectance, and the value of the first principle component image. Although misestimations for low and high density population CBGs were still significant the total population estimate error for the Indianapolis study area was only 3.2%.

Comments

We consider pixel-based estimation insufficiently explored to provide a generalized judgment of its worth. Because of differences in urban physiognomy, the results of the urban Tokyo, Ballarat, Hong Kong, and Indianapolis studies are difficult to apply to other worldwide cites. However, we

[16] See Deering et al. (1975)
[17] See Huete (1988)

agree with Li and Weng (2005) that "using remote sensing techniques to estimate population density is still a challenging task both in terms of theory and methodology, due to remotely sensed data, the complexity of urban landscapes, and the complexity of population distribution" (p. 955).

The foregoing review suggests several things. First, practitioners using per-pixel estimation should expect estimation problems in areas of extremely low or high population density. Second, per-pixel estimation using spectral radiance can benefit from stratification of the landscape that permits different regression equations to be built for specific landtypes. Third, because population errors of underestimation and overestimation tend to be compensating, accuracy of the approach will increase proportionally with the area of the enumeration units. Fourth, efforts to define spectral features more related to population density than reflected spectral radiance are warranted. The use of texture by Li and Weng (2005) and the spectral measures of Harvey (2002a) have already been mentioned. Webster (1996) presented several other texture measures appropriate to urban population modeling, all of which can be automatically derived from satellite imagery. In addition, indices derived from spectral reflectance commonly used in vegetation analysis[18] may likewise be more closely related to population density than spectral reflectance values used alone. Other indices specifically related to population density may require development. Finally, as illustrated by Li and Weng (2005), urban temperature measured by satellites is modulated by the amount of inert or built-up land within the thermal sensor footprint. Further exploring the use of thermal temperature as a surrogate for population density might likewise prove fruitful. If so, then the use of imagery from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) with its five thermal bands may provide improved estimates of population density than those possible with ETM+.

4.3 Case Study

The goal of this preliminary case study was to determine whether the pixel-based approach to population estimation developed by Iisaka and Hegedus (1982) and extended by Lo (1995) would be successful in modeling population density in the Wasatch Front of Utah. Three specific objectives were outlined.

[18] See Jensen (2007; pp. 382-393) for a review of popular vegetation indices.

1. To determine whether the population density among 1990 census block groups (CBGs) could be modeled as a function of the spectral characteristics of the census blocks as measured by Landsat Thematic Mapper data acquired in 1990.

2. To determine whether other spectral measures such as texture and the Normalized Difference Vegetation Index could model population density more accurately than simple spectral reflectance used alone.

3. To determine whether thermal data collected by Landsat Thematic Mapper could be effectively used to model population density in the study area.

This case study is a small part of a larger project to determine whether the pixel-based approach can be used to create accurate intercensal population estimates of the same Wasatch Front region.

Study Area

Demographics. As shown in Figure 4, the study region includes the major metropolitan areas of Utah. Over a dozen incorporated entities are part of the study area, including Bountiful, Salt Lake City, Taylorsville, Sandy City, Orem, Provo, American Fork, and Spanish Fork. It is an area of steady population increase. The average population growth along the Wasatch Front between 1970 and 2005 was about 2.5% per year.[19] The population growth has been steady primarily because of Utah's historically large family size and relatively high fertility levels. These factors make Utah somewhat unique in the U.S.; the majority of its population growth (80%) is from natural increase (GOPB 2005). Even when net migration is low or negative, Utah still experiences population growth driven by natural increase. Nonetheless, migration will continue to be an important factor in Utah population growth. Over the last fifteen years the Wasatch Front has averaged about 25,000 new residents a year and the GOPB predicts that about 42,000 new residents a year will make the Wasatch Front region their home between 2005 and 2030. Given the rapid and steady population growth, as well as its concomitant affect on the housing and construction sector and land use change, the Wasatch Front region is an excellent case study site.

[19] This is slightly below the 2.6% average annual population growth experienced by other states in the Mountain West census region, but significantly higher than the U.S. annual average population growth of 1.1%

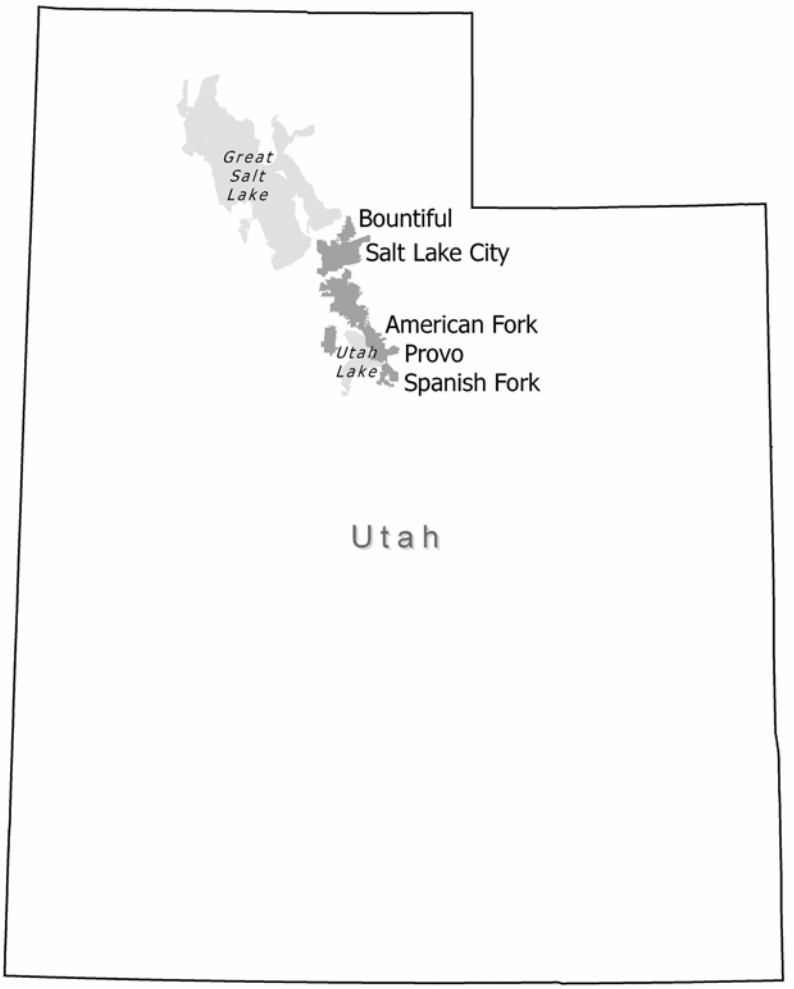

Figure 4. Municipalities in the study area along Utah's Wasatch Front.

Landtypes. Based on research conducted primarily in Salt Lake City, Utah, Ridd (1995) categorized urban fabric into three components; 1) vegetation, 2) impervious surfaces, and 3) bare soil. The common abbreviation for this triad is VIS. Impervious surfaces include concrete, asphalt streets, asphalt roofing, shingles, and metal roofing. Vegetation includes grass, tree, and shrub categories (after Ridd, 1995).

Using Ridd's model, Hung (2002) extensively studied the VIS components in Salt Lake City, findings we consider generally applicable to the whole study area. Hung found that urban commercial districts in Salt Lake City are strongly dominated by impervious surface (86%). Low density residential (i.e., single-family residential) zones are a mix of vegetation and impervious surface with vegetation representing nearly 70% of the land cover. Change from low through moderate to high density residential zones showed an increase in impervious surface area (24% $\rightarrow$ 34% $\rightarrow$ 46%) at the expense of vegetation (69% $\rightarrow$ 56% $\rightarrow$ 43%). Soil was a large component only in industrial zones (24%) and was about equal to vegetation coverage (26%).

Data and Methods

Two data sources were required to conduct this study. The first source was the 1990 Decennial Census of the United States. The specific variable extracted from the census was total population aggregated by census block groups (CBGs). A total of 812 CBGs were initially part of the study area. The average CBG size was 3053 hectares with a range from 11.3 to 520,000 hectares. Figure 5 is a map of population density in 1990 for the study area. The average population density among CBGs was 16.8 people / hectare (s = 12.6) and the average density of housing units was 6.3 units / hectare (s = 5.8). The densest population and housing recorded among the CBGs was 110.4 people / hectare and 48.5 units / hectare respectively. These high densities are found in areas of student housing adjacent to Brigham Young University in Provo, Utah.

The second data source was a Landsat ETM+ image dated May 28, 2000 (Figure 6). Table 1 shows the fundamental characteristics of the Landsat TM sensor. The quality of the image was good, but cloud cover masking was required to avoid problems relating spectral signature to population density -- portions of CBGs containing clouds were removed from the analysis. Unfortunately, thin smoke also partially obscured a few of the CBGs. Since this smoke was found at the interface between the suburban and rural areas, the affected image areas were not eliminated from the study. Instead, the areas were manually delimited and the contrast and brightness adjusted until they matched the surrounding area. This manual approach was only partially successful. The image data was radiometrically corrected, and standardized to reflectance/emittance as measured at the sensor.

Recognizing from the reviewed literature that derived spectral features may be more predictive of population density than simple spectral reflectance alone, several derived variables were generated from the seven TM bands for each CBG.

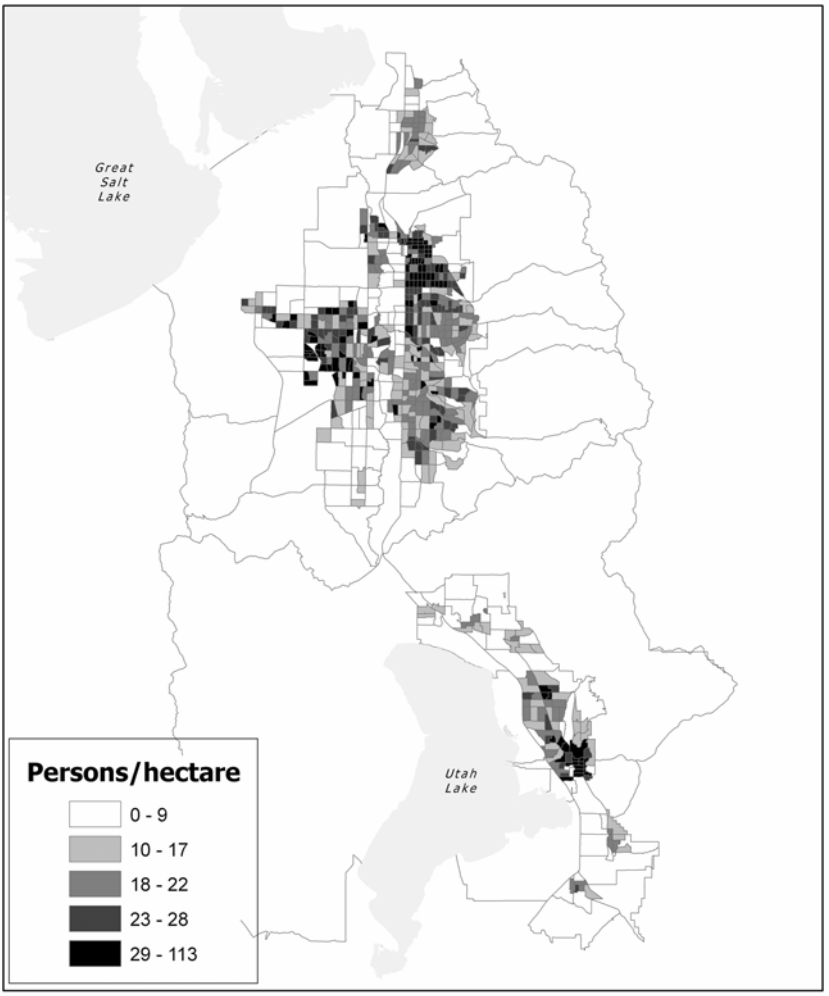

Figure 5. Population density of census block groups in the study area.

Figure 6. Band 4 (IR-1) of a Landsat ETM+ image acquired of the study area May 28, 2000 (2σ contrast stretch).

Table 1. Fundamental characteristics of the Landsat TM sensor. Columns 1, 3, and 4 from Landsat 4 Data Users Handbook (1984).

Band Number	Band Name	Band Wavelength (micrometers)	Nominal Resolution (meters)
1	Blue	0.45-0.52	30
2	Green	0.52-0.60	30
3	Red	0.63-0.69	30
4	IR-1	0.76-0.90	30
5	IR-2	1.55-1.75	30
6	Thermal	10.40-12.50	120
7	Mid-IR	2.08-2.35	30

The final variable set used in the correlation and regression analyses included the following:

- Mean Blue reflectance (B_μ)
- Mean Green reflectance (G_μ)
- Mean Red reflectance (R_μ)
- Mean IR-1 reflectance ($I1_\mu$)
- Mean IR-2 reflectance ($I2_\mu$)
- Mean thermal brightness temperature (T_μ)
- Mean Mid-IR reflectance ($I3_\mu$)
- Standard deviation of Blue reflectance (B_σ)
- Standard deviation of Green reflectance (G_σ)
- Standard deviation of Red reflectance (R_σ)
- Standard deviation of IR-1 reflectance ($I1_\sigma$)
- Standard deviation of IR-2 reflectance ($I2_\sigma$)
- Standard deviation of thermal brightness temperature (T_σ)
- Minimum of thermal brightness temperature (T_n)
- Maximum of thermal brightness temperature (T_x)
- Range of thermal brightness temperature (T_r)
- Standard deviation of Mid-IR reflectance ($I3_\sigma$)

The standard deviation, range variables, maximums and minimums were designed to be rough measure of spectral texture[20] in the CBG.

After exploratory correlation analysis, stepwise multiple regression analysis was utilized to build regression models explaining population density in

[20] Several popular texture measures, including fractal measures, were tried, but all proved less effective than these simple statistics.

the CBGs as a function of the variables listed above. Following the pattern of Harvey (2000a), several issues were considered in judging the goodness of a regression. The magnitude of the adjusted R^2 was considered. If all else were equal, a regression equation producing a higher R^2 was preferred. Simplicity was a second consideration. When regression equations producing similar R^2 values were compared, the equation using the fewest predictor variables was deemed best. Logical consistency was also necessary. Equations likely to produce negative population densities were discarded. Equations with variable combinations that appeared illogical were discarded. For example, it made little sense to include both $I1_\mu$ and $I2_\mu$ in the same equation unless we could reasonably explain why the addition of the second band added explanatory power not contained in the first. Equations exhibiting multicollinearity symptoms in the stepwise regression process were also discarded. Although a logarithmic transform of population density was required to improve linearity, we otherwise avoided variable transformations and higher order terms.

Results

Table 2 shows the simple correlation between the natural logarithm of population density (P_{ln}) and the independent variables listed above. The correlation between population density and spectral reflectance follows the same pattern reported by Li and Weng (2005, Table 3). Correlations are low but explainable patterns emerge. Blue, Green, Red, and IR-1 reflectance increase with increasing population density. Thermal temperature also increased with increasing density. In contrast, Mid-IR reflectance and IR-2 reflectance are inversely related to population density.

In the context of Hung (2002) summarized above, we consider these relationships largely a function of the relative amounts of vegetation and impervious material within the CBG. As housing unit density increases, the amount of concrete, asphalt, shingle, and other nonporous surfaces increases with the concurrent loss of grass, trees, and shrubbery. Since impervious surfaces reflect a larger proportion of incident visible light than vegetation does, the increase in reflectance with increased housing density is logical. Our interpretation of the similar increase in IR-1 reflectance is less obvious. Vegetation has a high IR-1 reflectance. Our tentative explanation for the sign in Table 2 is that the IR-1 reflectance increase from impervious surfaces more than offsets the loss of IR-1 reflectance due to vegetation loss.

Table 2. Pearson's *r* correlation between CBG spectral features and P_{ln}. Given the sample size of 807 CBGs, all are significant at the 0.01 level.

Spectral feature		*r*
Blue reflectance	Mean	0.202
	S.D.	-0.184
Green reflectance	Mean	0.185
	S.D.	-0.190
Red reflectance	Mean	0.117
	S.D.	-0.219
IR-1 reflectance	Mean	0.210
	S.D.	-0.203
IR-2 reflectance	Mean	-0.176
	S.D.	-0.427
Thermal temperature	Mean	0.238
	S.D.	-0.671
	Minimum	0.653
	Maximum	-0.488
	Range	-0.773
Mid-IR reflectance	Mean	-0.127
	S.D.	-0.352

Table 2 also demonstrates that spectral reflectance variability within a CBG is frequently a better predictor of population density than mean spectral reflectance. This agrees with the results of Harvey (2002) among others. This is clearly the case for the IR-2, Mid-IR and Thermal bands. In all cases, population density was inversely related to the variability features. We tentatively suggest that the increase in impervious surface associated with increased housing density is betrayed as an increase in urban surface homogeneity as vegetation amounts become more limited.

Given the weak correlations generally throughout the table, the strength of the thermal band features as predictors of population density is striking. This agrees with findings by Li and Weng (2005) for the Indianapolis, Indiana metropolitan study site. The reason for thermal temperature and population density covariance is well reported. Thermal temperature increases with the increased proportion of inert material associated with increased dwelling structure density. The high negative correlation between population density and temperature indicates that CBGs with higher popu-

lation densities also have less temperature variability because the impervious material is ubiquitous throughout the CBG. Areas of lower population density will have built-up zones of housing (warmer) interspersed with parks, grass, empty vegetated lots, pastures, and agricultural areas (cooler).

Given that the relative amounts of vegetation and impervious material within a CBG were fundamentally responsible for the spectral reflectance → population density relationship, we generated some ratios to better represent the inverse relationship between vegetation and impervious material within the study area. These included:

- Mean CBG red band reflectance / Mean CBG IR-2 reflectance ($R_{3/5}$)
- Mean CBG red band reflectance / Mean CBG Mid-IR reflectance ($R_{3/7}$)

Multiple Regression Models. Table 3 contains the best regression models as judged by the criteria discussed previously. In all the models, the natural logarithm of the CBG population density (P_{ln}) was the dependent variable. All of the regression equations, constants and variables are significant at a 0.05 level minimum.

From the perspective of parsimony, it appears that the three-variable model would be preferred for practical application. This model produced a multiple R of 0.80 and an adjusted R^2 of 0.64. Models with more predictor variables produced models with R^2 values exceeding 0.70, but were difficult to interpret. As the three variable model shows, the best predictive combination includes temperature range, the first TM infrared band, and the ratio of the red band to IR-2 reflectance.

Table 3. Best regression equations to model P_{ln}.

Number of predictors	Equation	R^2 (adjusted)
1	$4.390 - 0.176\ T_r$	0.60
2	$2.389 - 0.173\ T_r + 6.65\ I1_\mu$	0.62
3*	$-0.164\ T_r + 8.685\ I1_\mu + 3.107\ R_{3/5}$	0.64
4*	$-0.161\ T_r + 9.362\ I1_\mu + 3.528\ R_{3/5} - 3.003\ I3_\mu$	0.64

* = the constant of the regression equation was not significant and is not included in the equation shown.

Residual Analysis. Figure 7 is a map of standardized residuals from the regression. Generally speaking, the best prediction was obtained in CBGs with moderate to small area in the central corridor of the study area. The

relationship between CBG area and total population error[21] is linear and strong (r = +0.853, n = 812). These results are different from Lo (1996) but correspond to the results of Harvey (2002a). In the Wasatch Front, the large CBGs have boundaries which cross mountain slopes perpendicularly and capture areas of rural desert, agriculture, woodland and forest that form the hinterland of the core city area. Obviously the mean spectral values of these CBGs do not fairly represent the spectral character of the urbanized proportion of the CBG.

Any error in the population density produced by the regression equation also resulted in enormous estimation errors when multiplied by the area of large CBGs. For example, the largest CBG in the study exceeded 520,000 hectares, and had a true population of only 192. With very sparse vegetation, it has an average spectral signature similar to high density residential zones. Using that spectral signature information, the regression equation estimated a population density of 0.83 people per hectare and a total CBG population of over 431,000 inhabitants. It is an understatement to conclude that some manual adjustment of CBG boundaries to better fit the actual urban / suburban area within a CBG is warranted.

[21] Total population error was calculated as the absolute value of (true CBG population - modeled CBG population).

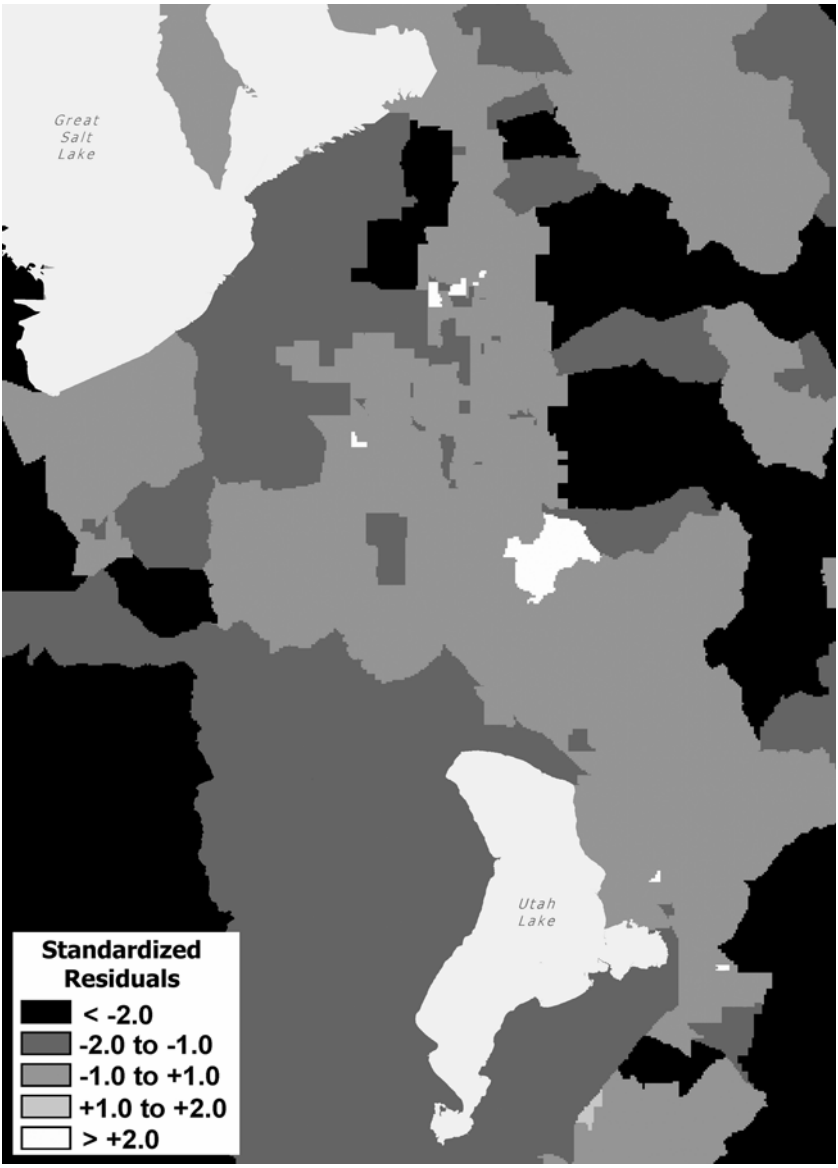

Figure 7. Map of residuals showing areas of over and underestimation by the regression equations

4.4 Concluding Comments

Research into population estimation using overhead imagery naturally leads to the question of whether the Decennial Census of the United States could be built upon remotely sensed data. In the early 1980's, there was a flurry of discussion regarding that question. The dialog was formally started by Brugioni (1983) who sided with the Secretary of the Commerce in his belief that the accuracy of the 1980 census could have been increased while concurrently decreasing the cost. At the time of Brugioni's writing, politicians in many large U.S. cities were unwilling to accept the enumeration as accurate, generally claiming the undercounting of urban subgroups. The primary issue was not apportioned representation in congress but rather the disbursement of federal funds based on population. City mayors observed that minorities and illegal immigrants used federally funded services and an accurate count of them was essential. Brugioni proposed, "It is time to stop and assess the prospects for better performance of the 1990 census. By using overhead reconnaissance systems which carry sophisticated cameras and remote sensing equipment, and by employing modern interpretation methods backed up by the latest computer technology, I am convinced that the census can be done more accurately, cheaper, faster, and better than by previous methods" (p. 1337). Given that the article was a commentary designed to elicit comments from the journal readership, Brugioni can be forgiven for such enthusiasm. The stated reasons for his certainty are hazy. Brugioni discusses the wealth of technology available for the task and recites a paragraph of remote sensing success stories such as weather prediction, forest inventory, and strategic intelligence. He then wonders why the Bureau of the Census has "not attempted to use these same technologies to determine the number of people living in a given area of a U.S. city" (p. 1338). Brugioni also cites the ongoing military use of overhead imagery to estimate urban populations abroad. Brugioni fails to mention that such foreign estimations seldom have an accuracy check. In conclusion, Brugioni appeals to his own authority, "From my nearly 40 years of experience in all phases of reconnaissance and analysis activity, I am thoroughly convince [sic] that a census using space-age technology is not only feasible but can be performed better and cheaper, and be more responsive to the needs of modern-day America."

Academics were quick to highlight the logical flaws of Brugioni's arguments as well as other problems with conducting a census with space-age technology. Morrow-Jones and Watkins (1984), both human / population

geographers, believed that overhead imagery could never become the primary data source for the census. Several problems were articulated. They cited the practical hurdle of imaging the whole nation in a timely manner that could provide the same April 1st national snapshot as the census. Operational definitions (e.g., the distinction between *de jure* and *de facto* counting) long used by the census bureau and long valued by social scientists because of their constancy through several decades would change, "interrupting an exceptionally valuable source of evenly spaced historical data" (p. 230). While some census variables (e.g. house size) might be amenable to collection on overhead imagery, "their meanings would be changed to the detriment of long term comparative research" (p. 230). Morrow-Jones and Watkins also claimed that too little was known about using image characteristics as surrogates for social characteristics measured in the census. "Can these methods tell us the change in age structure, household composition, family income, race, sex ratio, or other *characteristics* of the people? This is a crucial part of the census and the largest drawback to the suggested method for improving it" (p. 231). Ethical considerations of privacy and image data use also troubled Morrow-Jones and Watkins. Sinclair (1984) was similarly troubled and wrote, "In America, is there not a right to ignore the Census and the Census taker? Or is it so important that we be counted that our own spy networks must be trained on us?" (p. 80). In conclusion, Morrow-Jones and Watkins admitted that a decennial census might be possible using remotely sensed data, but "the tradeoff would mean a great deal less information" (p. 232).

We suspect that Paul (1984) probably states the present convergence of opinion on this matter. First is the opinion that the use of remotely sensed data for human disciplines such as social science and public health has been historically undervalued and inadequately studied. The studies collected by Liverman et al. (1988) demonstrated the kind of progress that might be possible. However, as regarding population geography, while "remote sensing technology can be useful *in conjunction* with traditional demographic enumeration techniques [it] cannot be used as a replacement" (Paul 1984, p.1611). Expanding Paul's thought, we likewise do not consider remote sensing a replacement for traditional Census enumeration, but as an important source of urban data nonetheless. As reviewed by Lo (2006), many researchers have found that remotely sensed data is the only source of population information available in many developing countries. The intelligent use of imagery to augment traditional approaches may likewise be an efficient and accurate way of completing intercensal population estimates in areas of rapid urban growth. Furthermore, as scientists seek ways to improve the human condition, it may be profitable to con-

sider *gathering new kinds of urban data* via remote sensing rather than developing new high-tech ways to capture traditional measures. See Jensen and Cowan (1999) for a review. For example, Weber and Hirsch (1992) demonstrated the calculation of spatial urban quality of life indices from imagery that are not amenable to a ground survey. Lo and Faber (1997) likewise demonstrated the generation of quality of life variables from overhead imagery. Unfortunately, the works of Weber and Hirsch as well as Lo and Faber have not been widely studied and have certainly not penetrated mainstream urban social science. Thus, although remote sensing will not likely supplant the typical Census enumeration, it nonetheless may provide critical measures of import to social scientists and policy makers in the pursuit of both knowledge and social justice.

References

Adeniyi, P.O. 1976. Applications of Aerial Photography to the Estimation of the Characteristics of Residential Buildings. *The Nigerian Geographical Journal*, 18:189-200.

Adeniyi, P.O. 1980. Land Use Change Analysis Using Sequential Aerial Photographs and Computer Techniques. *Photogrammetric Engineering and Remote Sensing*, 46:1447-1464.

Adeniyi, P.O. 1983. An Aerial Photographic Method for Estimating Urban Population. *Photogrammetric Engineering and Remote Sensing*, 49: 545-560.

Binsell, R. 1967. *Dwelling Unit Estimation from Aerial Photography*. Northwestern University Department of Geography.

Brugioni, 1983. The Census: It Can Be Done More Accurately With Space-Age Technology. *Photogrammetric Engineering and Remote Sensing*, 49:1337-1339.

Collins, W.G. and A.H.A. El-Beik. 1971. Population Census with the Aid of Aerial Photographs: An Experiment in the City of Leeds. *Photogrammetric Record*, 7:16-26.

Deering, D.W., J.W. Rouse, R.H. Haas, and J.A. Schell. 1975. Measuring Forage Production of Grazing Units from Landsat MSS Data. *Proceedings of the Tenth International Symposium on Remote Sensing of Environment.* Ann Arbor, MI: Environmental Research Institute of Michigan. Vol. 2. pp. 1169-1178.

El-Beik, A.H.A. 1967. *Air photo-interpretation applied to the study of urban land use and the determination of population densities.* Ph.D. thesis, University of Leeds, 214 pp.

GOPB (Governor's Office of Planning and Budget). 2005. *Economic Report to the Governor*. Salt Lake City, Utah: Government Printing Office. Available online at http://www.governor.utah.gov/dea. Last accessed October 15, 2006.

Green, N.A. 1956. Aerial Photographic Analysis of Residential Neighborhoods: An Evaluation of Data Accuracy. *Social Forces*, 35: 142-147.

Green, N.A. 1957. Aerial Photographic Interpretation and the Social Structure of the City. *Photogrammetric Engineering* 23:89-99.

Hadfield, S.M. 1963. *Evaluation of Land Use and Dwelling Unit Data Derived from Aerial Photography*. Chicago: Urban Research Section, Chicago Area Transportation Study.

Harvey, J.T. 1999. Estimation of Population Using Satellite Imagery. PhD Thesis, University of Ballarat.

Harvey, J.T. 2002a. Estimating Census District Populations from Satellite Imagery: Some Approaches and Limitations. *International Journal of Remote Sensing*, 23:2071-2095.

Harvey, J.T. 2002b. Population Estimation Models Based on Individual TM Pixels. *Photogrammetric Engineering and Remote Sensing*, 68:1181-1192.

Huete, A.R. 1988. A Soil-Adjusted Vegetation Index (SAVI). *Remote Sensing of Environment*, 25:295-309.

Hung, Ming-Chih. 2002. Urban Land Cover Analysis from Satellite Images, *Proceedings of the Pecora 15 Conference*, Nov. 10-15, 2002, Denver, CO, USA. Available online at http://www.isprs.org/commission1/proceedings02/paper/00099.pdf . Last Accessed September 1, 2006.

Iisaka, J. and E. Hegedus. 1982. Population Estimation from Landsat Imagery. *Remote Sensing of Environment*, 12:259-272.

Jensen, J.R. 2005. *Introductory Digital Image Processing: A Remote Sensing Perspective.* 3rd Ed. Upper Saddle River, N.J.: Prentice-Hall. 526 p.

Jensen, J.R. 2007. *Remote Sensing of the Environment: An Earth Resource Perspective.* 2nd Ed. Upper Saddle River, N.J.: Prentice-Hall. 592 p.

Jensen, J.R. and Cowan, D.R. 1999. Remote Sensing of Urban / Suburban Infrastructure and Socio-Economic Attributes. *Photogrammetric Engineering and Remote Sensing*, 65:611-622.

Kraus, S.P., L.W. Senger, and J.M. Ryerson. 1974. Estimating Population from Photographically Determined Residential Land Use Types. *Remote Sensing of Environment*, 3:35-42.

Langford, M., D.J. Maguire, and D.J. Unwin. 1991. The Areal Interpolation Problem: Estimating Population Using Remote Sensing in a GIS Framework. Pages 55-77 in *Handling Geographical Information: Methodology and Potential Applications.* I. Masser and M. Blakemore (Eds.). New York: Longman Scientific and Technical.

Li, G. and Q. Weng. 2005. Using Landsat ETM+ Imagery to Measure Population Density in Indianapolis, Indiana, USA. *Photogrammetric Engineering and Remote Sensing*, 71:947-958.

Lindgren, D.T. 1971. Dwelling unit estimation with color-IR photos. *Photogrammetric Engineering and Remote Sensing*, 37:373-377.

Liverman, D., E.F. Moran, R.R. Rindfuss, and PC. Stern. 1988. *People and Pixels: Linking Remote Sensing and Social Science*. Washington, DC: National Academy Press. 256 pp.

Lo, C.P. 1995. Automated Population and Dwelling Unit Estimation from High-Resolution Satellite Images: a GIS Approach. *International Journal of Remote Sensing*, 16:17-34.

Lo, C.P. 1986. Accuracy of population estimation from medium-scale aerial photography. *Photogrammetric Engineering and Remote Sensing*, 52 (12): 1859-1869

Lo, C.P. 2006. Estimating Population and Census Data. *Remote Sensing of Human Settlements.* Manual of Remote Sensing. 3rd Ed. Vol. 5. M.K. Ridd and J.D. Hipple (Eds.). Falls Church, VA: American Society for Photogrammetry and Remote Sensing, pp. 337-377.

Lo, C.P. and B.J. Faber. 1997. Integration of Landsat Thematic Mapper and Census Data for Quality of Life Assessment. *Remote Sensing of Environment*, 62:143-157.

Lu, D. and Q. Weng. 2004. Spectral Mixture Analysis of the Urban Landscape in Indianapolis with Landsat ETM+ Imagery. *Photogrammetric Engineering and Remote Sensing*, 70:1053-1062.

Morrow-Jones, H. and J.F. Watkins. 1984. Remote Sensing Technology and the U.S. Census. *Photogrammetric Engineering and Remote Sensing*, 50:229-232.

Olorunfemi, J.F. 1984. Land Use and Population: A Linking Model. *Photogrammetric Engineering and Remote Sensing*, 50:221-227.

Paul, C.K. 1984. More on the Census. *Photogrammetric Engineering and Remote Sensing*, 50: 1610-1612.

Plane, David A. and Peter A. Rogerson. 1994. *The Geographical Analysis of Population*. New York: John Wiley & Sons, Inc.

Ridd, M.K. 1995. Exploring a V-I-S (Vegetation - Impervious surface – Soil) Model for Urban Ecosystems Analysis through Remote Sensing: Comparative Anatomy for Cities. *International Journal of Remote Sensing*, 16:2165-2185.

Sinclair, D. 1984. The Census: It Can be Done More Accurately with Space-Age Technology." *Photogrammetric Engineering and Remote Sensing*, 50:80.

Singh, A. 1989. Digital change detection techniques using remotely sensed data, *International Journal of Remote Sensing*, 10:989-1003.

UNCHS 2001. *The State of the World's Cities*. Nairobi, Kenya: United Nations Centre for Human Settlements.

Watkins, J.F. 1984. The Effect of Residential Structure Variation on Dwelling Unit Enumeration from Aerial Photographs. *Photogrammetric Engineering and Remote Sensing*, 50:1599-1607.

Weber, C. and J. Hirsch. 1992. Some Urban Measurements from SPOT Data: Urban Life Quality Indices. *International Journal of Remote Sensing*, 13:3251-3261.

Webster, C.J. 1996. The potential of urban texture measures in monitoring urbanization form space. *GIS in Asia*, T. Fung, P.C. Lai, H. Lin and A.G.O. Yeh, eds. Singapore: GIS Asia Pacific, pp. 309-321.

Weng, Q., D. Lu, and J. Schubring. 2004. Estimation of Land Surface Temperature-Vegetation Abundance Relationships for Urban Heat Island Studies. *Remote Sensing of Environment*, 89:467-483.

5 Using Satellite Data to Estimate Urban Leaf Area Index[1]

Ryan R. Jensen, Department of Geography, Geology, and Anthropology, Indiana State University, Terre Haute, Indiana State University

Perry J. Hardin, Department of Geography, Brigham Young University, Provo, Utah

5.1 Introduction

The social value of the urban forest to local urban populations has long been recognized. In contrast, the impact of the urban forest on global and local environments is not clearly understood, and the impact of urban trees on carbon sequestration, mitigation of urban heat, and removal of pollution remain topics of contemporary scientific study. Land cover conversion in urban areas is typically faster than in wildland areas, thus there is a need for rapid measurement methods of urban biophysical variables that are repeatable and economically efficient.

Leaf area index (LAI) has been identified as one of the core biophysical variables for landscape monitoring at all scales (Pierce and Running 1988; Lymburner et al. 2000). LAI has three definitions in the literature but is usually standardized to represent the green area (m^2) of flat horizontal leaves per unit of ground area (m^2) (Chen and Black 1992; Chen et al. 1997; Barclay 1998). Many scenarios of season and landscape allow LAI measurement by earth resource satellites, and LAI is a derivative data product of many remote sensing initiatives. However, few studies have ex-

[1] Originally published in the Journal Arboriculture, Volume 31, Issue 1, Pages 21-27 under the title, "Estimating urban leaf area using field measurements and satellite remote sensing data."

amined methods of combining satellite LAI estimates with those made using ceptometers in the urban forest to estimate LAI over large urban areas.

This research extends the work of Peper and McPherson (2003) that compared the accuracy of various nondestructive field measurement devices to accurately measure urban tree LAI. In the context of that previous work, algorithmically manipulated satellite data used in this study become an additional nondestructive method of measuring urban LAI.

The objective of this research is to develop transfer equations that can be used to convert satellite LAI measurements to their gap-fraction equivalents. Our hypothesis proposes that satellite and ground LAI measurements are related and that statistical and neural network approaches can be used to interconvert between the two methods of measurement.

5.1.1 Urban Remote Sensing

Instruments aboard remote sensing satellites measure the electromagnetic energy emitted or reflected from Earth or its atmosphere, allowing terrestrial objects to be distinguished and characterized. For example, when illuminated by the noonday sun, grass on an irrigated golf course is not only visibly green but also reflects intercepted infrared solar energy in proportion to the amount of its spongy mesophyll. Grass receiving insufficient moisture to maintain mesophyll turgidity may appear equally as green as adjacent well watered grass but would decrease significantly in infrared reflectance. If spatially extensive, this stress would be detectable from spaceborne instrumentation and would allow researchers to accurately map the affected area. Using similar logic, land cover types are mapped, and vegetation biophysical variables are measured from spaceborne instruments.

Historically, remote sensing in urban areas has been constrained by the spatial complexity of urban scenes. The problem is related to the spatial resolution of the satellite sensor. A single image resolution element (pixel) may be measuring the spectral response of a land cover mixture rather than a single land cover type. For example, a suburban pixel may represent a mixture of grass, asphalt, concrete, and roof shingles. This kind of spectral mixing makes urban remote sensing less amenable to statistical methods that assume normal distributions and no measurement error. Newer spaceborne instruments, having finer spatial resolutions, reduce the constraint and provide better data for urban remote sensing (Jensen et al. 2003). The

improvement in resolution is fortunate, because governments (e.g., state, county, city) and private companies annually invest hundreds of millions of dollars acquiring remotely sensed data that detail the urban landscape more effectively than through traditional "windshield surveys" (Jensen 1996).

5.2 Data and Methods

5.2.1 Study Area

The city of Terre Haute is located in Vigo County along the banks of the Wabash River in west central Indiana, U.S. (39° 25′ N, 87° 25′ W). Terre Haute government officials have made a conscious effort to maintain the urban tree canopy through a comprehensive tree ordinance that governs both tree removal and planting. The ordinance is administered by a tree advisory board consisting of city residents appointed by the city officials to make suggestions and recommendations to the mayor, city forester, city engineer, and city council.

5.2.2 LAI Field Measurements

Traditional field measurement of LAI has taken two approaches. The first approach requires the destructive harvesting of leaves within a vertical column passing upward through the entire tree canopy. The second involves collection of leaf litterfall. These direct methods are similar — they are time intensive and require many replicates to account for spatial variability in the canopy (Green et al. 1997). However, these direct LAI measurements are accurate for a very specific geographic location, are relatively easy to perform by untrained personnel, and are well understood by ecologists. Gap-fraction analysis is a nondestructive field method that has been developed to estimate LAI.

Gap-fraction analysis is predicated on the theory that the decrease in light intensity (light attenuation) with increasing depth in vegetative canopies can be described by the relationship:

$$IL / IO = e^{-kLAI(L)}, \tag{5.1}$$

where *IL/IO* is the fraction of incident light at the top of the canopy (*IO*) reaching depth *L* in the canopy, *LAI*(*L*) is the cumulative LAI from the top of the canopy to point *L*, *k* is a stand or species specific constant, and *e* is the natural logarithm base (Larcher 1975; Aber and Melillo 1991). Different types of vegetation have different *k* values, causing different rates of light attenuation for the same leaf area. The principal factor causing this is "twig angles and the angles that the foliage subtends with the twig" (Barclay 1998; see also Larcher 1975). Field-measured LAI using gap-fraction analysis assumes that leaf area can be calculated from the fraction of direct solar energy that penetrates the canopy (canopy transmittance). By applying gap-fraction techniques to study LAI in many different forest settings, standard operating procedures have been developed (Pierce and Running 1988; Chason et al. 1991; Ellsworth and Reich 1993; Nel and Wessman 1993; Green et al. 1997).

In this study, LAI was measured using the gap-fraction approach in 145 random locations (sampling sites) throughout the study area during July and August 2001. Like most urban areas, land cover in Terre Haute consists of a wide variety of vegetated and nonvegetated patches. Vegetated areas sampled included trees, shrubs, grasses, and agricultural fields growing different varieties of corn and soybeans. Unvegetated areas included buildings, streets, parking lots, ponds, lakes, and the Wabash River. The randomly selected sampling sites represented all major land cover types in Terre Haute.

Each of the 145 sampling sites was defined as a 20 × 20 m (65.6 × 65.6 ft) quadrat identified by the global positioning system coordinates of its center. At each sampling point, 16 below-canopy, photosynthetically active radiation (PAR) measurements were collected, one in each cardinal direction at each corner of the 20 m quadrat. The PAR measurements were collected using a Decagon AccuPar Ceptometer™ held approximately 1 m (3.3 ft) above the ground beneath the tree cover. The AccuPar Ceptometer consists of a linear array of 80 adjacent, 1 cm^2 (0.16 in^2) PAR sensors mounted rigidly along a bar and oriented so that when the operator holds the ceptometer horizontally, the PAR passing downward through the canopy can be measured. The ceptometer stored the 16 PAR samples taken at each sampling site and calculated the LAI average automatically. This sitewide LAI average was then recorded along with general operator notes regarding the sampling site character.

5.2.3 Satellite Sensor Data

Data from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) sensor were used for comparison to the field LAI measurements. ASTER data are collected in several wavelengths, often referred to as bands. This study employed ASTER bands 1, 2, and 3 measuring the green, red, and near-infrared segments of the electromagnetic spectrum (520–600 nm, 630–690 nm, and 790–860 nm), respectively. These wavelengths are used in vegetation studies because of their correlation to the quantity and health of green vegetation (Jensen 2000). Remote sensing data are commonly used to calculate vegetation indices — dimensionless, radiometric measures of the relative abundance of green vegetation, including LAI (Jensen 2000). One of the most common vegetation indices is the Normalized Difference Vegetation Index (NDVI). The NDVI is calculated using the equation (Rouse et al. 1974):

$$NDVI = \frac{NIR - RED}{NIR + RED}, \tag{5.2}$$

where NIR is the near-infrared reflected radiant flux, and RED is the red reflected radiant flux.

An ASTER image of the study area acquired in July 2001 was used for this investigation. The image had a spatial resolution of 15 m (49.5 ft). Using a United States Geological Survey digital raster graphic image, the ASTER scene was geometrically adjusted to the same coordinate system used for the field data collection. This adjustment ensured that the 145 sample sites could be accurately registered to the ASTER data.

5.2.4 Estimating LAI Using Regression

As mentioned above, the principal objective of this research was to create transfer equations that could be used to convert satellite LAI measurements to their gap-fraction equivalents. Because correlation and regression are common methods used to model forest biophysical characteristics with remotely sensed data (e.g., Jensen et al. 2000), their use was suggested. In this instance, multiple regression analysis was performed using brightness values from the three ASTER bands as the independent variables (Table 1). Because previous remote sensing research has shown that ratios and vegetation indexes derived from brightness values (e.g., NDVI) frequently measure vegetation differences better than the direct brightness values alone (Fassnacht et al. 1997; Jensen 2000), five derived independent variables were also explored in the regression process.

Table 1. ASTER bands and derived variables used in the study

Variable	Meaning	Variable	Meaning
GREEN	520 – 600 nm Band (Green band)	RDIR_RAT	Red / Infrared ratio
RED	630 – 690 nm band (Red band)	NDVI	(See Formula 5.2)
IR	790 – 860 nm band (Infrared band)	GRRD_DIF	Green – Red band difference
GRRD_RAT	Green / Red ratio	RDIR_DIF	Red – Infrared band difference

These variables are described in Table 1. In all regressions, the average field site LAI value (LAI_{obs}) was the dependent variable.

The goodness of the regression models were measured in two ways. The first is the standard error of the estimate (SEE). The standard of the estimate is synonymous for root mean square error; the former term is preferred in regression, whereas the latter term is preferred in neural network studies. The standard error of the estimate is defined as:

$$SEE = \sqrt{\frac{\sum_{i=1}^{n} (LAI_{pred} - LAI_{obs})^2}{n}}, \tag{5.3}$$

where LAI_{pred} is the LAI for a given fieldsite predicted by the regression. The summation is iterated over all the observations in the dataset (i = 1 to n). Smaller values of *SEE* indicate better fit between model and observed data and can be interpreted as the best estimate of the standard deviation of the observations around the regression line. The second method was the common multiple correlation coefficient (R) as described in Marascuilo and Levin (1983). The minimum acceptable level of significance in all the statistical analyses was 0.05.

5.2.5 Estimating LAI Using a Back-Propagation Feed-Forward Network

Artificial neural networks (ANNs) grew out of research in artificial intelligence, specifically attempts to mimic the fault tolerance and learning capacity of biological neural systems by modeling the low-level structure of

the brain. Research on ANNs has been motivated from their inception by the recognition that the brain computes in a very different way than digital computers (Haykin 1994).

A neuron is the fundamental processing unit of an ANN. Artificial neurons are analogous to biological neurons in the human brain. ANN behavior resembles that of the brain in two respects. First, knowledge is acquired by the network through a learning process. Secondly, interneuron connection strengths known as synaptic weights are used to store knowledge (Haykin 1994). ANNs do not rely on statistical relationships for function fitting but adaptively estimate continuous functions from data without mathematically describing how outputs depend on inputs (e.g., adaptive model-free function estimation using a nonalgorithmic strategy) (Gopal and Woodcock 1996).

ANNs have been used in remote sensing applications to classify images (Bischof et al. 1992; Hardin 2000) and incorporate multisource data (Benediktsson et al. 1990). ANN classifiers have been successfully used with remote sensing data because they take advantage of the ability to incorporate non-normally distributed numerical and categorical GIS data and image spatial information (Jensen et al. 2000).

Several forest studies have demonstrated the utility of coupling ANN approaches with satellite data. For example, Jensen et al. (2000) used an ANN to discriminate conifer stand age in southern Brazil using remotely sensed imagery. That study demonstrated that ANNs were; (1) competent to model the complex nonlinearity of biophysical forest processes, (2) better at estimating conifer stand age than traditional image processing techniques, (3) ideal for modeling the latent complexity of plant biophysical characteristics during the plant life cycle, and (4) able to explain more variance in forest biophysical parameters than their traditional statistical counterparts. In another study, Jensen and Binford (2004) found that ANNs were more accurate than traditional statistical techniques to estimate LAI in forested ecosystems throughout north central Florida.

For this research, a back-propagation feed-forward ANN was created and trained using the variables shown in Table 1 as inputs, and the field site LAI (i.e., LAI_{obs}) as the output. This procedure is directly analogous to the multiple regression approach described previously in this article, in which ASTER variables and LAI_{obs} were the independent and dependent variables, respectively.

The procedure used to build the ANN models generally followed Hardin (2000). The calibration of several candidate networks required trial and error. The networks were trained with different variable combinations, different numbers of hidden neurons, different learning rates, and different momentum rates until an acceptable error rate was obtained or further improvement was unlikely. Parsimony was also sought in the neural network solutions. Given equal predictive value from alternative network configurations, the network with the fewest hidden neurons was considered superior to more complex networks.

Like the regression approach described above, the *SEE* was also used to measure the accuracy of the network predictions by comparing LAI_{obs} values against LAI_{pred} values across all 145 fieldsites. *R* was also calculated for neural networks by regressing predicted LAI values against their observed counterparts. Use of the same accuracy metrics allowed the regression results to be compared to the ANN outcome.

5.3 Results and Discussion

The field LAI measurements were made at 145 Terre Haute area locations (n = 145). The maximum and minimum LAI recorded were 7.7 and 0.0, respectively. The mean LAI measured was 1.2 (s = 1.9).

5.3.1 Regression Results

All possible single variable regression models were tested. Several provided statistically significant predictive ability. The regression model providing the highest correlation coefficient (R = 0.60) and lowest error (*SEE* = 1.54) was created from the ratio of the ASTER green and ASTER red bands (*GRRD_RAT*). In unstandardized form, the predictive model was

$$LAI_{pred} = 4.79 \times GRRD\_RAT - 5.81. \quad (5.4)$$

The best two-variable model included the same ratio as Equation 4 but added the infrared ASTER band. The predictive equation using these two variables was

$$LAI_{pred} = 3.99 \times GRRD\_RAT + 0.02 \times IR - 7.10. \quad (5.5)$$

This model lowered the standard error of the single variable model by only 3% (SEE = 1.51) and improved the simple correlation by only 5% (R =

0.62). The addition of further variables did not improve the predictive ability of the model.

For all the regression equations previously cited, their coefficients, and constants were significant at the 0.05 level. Use of the single variable model (Equation 4) is suggested because of its simplicity. The direction of the coefficient signs for the regression variables is logical. As the amount of green reflectance increases and red reflectance decreases, the ratio GRRD_RD increases mathematically and LAI_{pred} also increases. In addition, as infrared reflectance increases, so does LAI_{pred}. These results suggest that both the ratio and the infrared band are measuring the same physical phenomena; they are measuring the increase in green reflectance as leaf area increases with the co-occurring loss of ground reflectance.

5.3.2 Artificial Neural Network Results

All the variables in Table 1 were submitted separately to ANN analysis to create a single variable model explaining LAI. Three single-variable networks produced nearly equivalent LAI predictive accuracy ($R \approx 0.69$, $SEE \approx 1.40$). These three models employed, respectively, the ASTER green band, *NDVI*, and the ratio between the ASTER red and infrared variables (*RDIR_RAT*). In all three cases, networks with two neurons in a single hidden layer were sufficient for fitting the network. The best two variable models tested included the ASTER green band (*GREEN*) and the ratio between the ASTER red and infrared bands (*RDIR_RAT*) trained on single hidden layer of three neurons. This network produced a moderately high R value ($R = 0.71$) with an *SEE* of 1.35. A Visual Basic function that reproduces the network output is shown in Figure 1. This function returns LAI_{pred} when passed *GREEN*, *RED*, and *INFRARED* brightness values. The variable *RDIR_RAT* is calculated inside the function from *RED* and *INFRARED* and then used with *GREEN* in the network calculations. No three-variable network models significantly exceeded the predictability of this two-variable network model.

The interpretation of the neural network results follows the same logic used in discussing the regression results. LAI_{pred} increases with increased reflectance in the ASTER green band. The ratio of the red to infrared reflectance (*RDIR_RAT*) assumes the role that *GRRD_RAT* played in the regression analysis; it is a measure of the ratio of background to photosynthetically active vegetation or healthy, spongy leaf mesophyll. With an increase in vegetation at the expense of impervious material, infrared re-

flectance in the ratio increases while red reflectance decreases. This causes a corresponding increase in LAI_{pred}.

5.4 Conclusion

As shown in Table 2, the ANN technique was superior to the multiple regression approach. In all cases, the ANN produced higher values of *R* and lower values of *SEE* than did regression. These results provide another case study demonstrating that a biophysical variable critical to urban study (i.e., LAI) can be predicted from remotely sensed satellite data and be more accurately predicted using a feed-forward back-propagation neural network than multiple linear regression.

Using ANNs to estimate LAI could enhance the accuracy of some studies that have relied on traditional regression techniques in the past. To improve such studies, ANNs could be created and trained using representative ecosystem *in situ* LAI samples and then used to estimate LAI in other image areas. For example, after measuring *in situ* urban LAI using one of the methods described by Peper and McPherson (2003), an ANN could be created and trained that is unique to that specific urban area. A program such as that shown in Figure 1 could then be used to estimate LAI in the

```
'Input ASTER green, red and infrared brightness values
'Function will produce a predicted LAI
Function PredictLAI(ByVal GREEN As Double, _
                    ByVal RED As Double, _
                    ByVal INFRARED As Double) As Double
    Dim NetSum As Double
    Dim LAI As Double
    Static HiddenNeuron(3) As Double

    'Handle water and outside image
    If INFRARED = 0.0 then Return 0.0

    'Calculate ratio of red and infrared
    Dim RDIR_RAT As Double = RED / INFRARED

    'Prescale
    If (GREEN < 48.9) Then GREEN = 48.9  'Brightness
    If (GREEN > 206) Then GREEN = 206
    GREEN = (GREEN - 48.9) / 157.1
    If (RDIR_RAT < 0.18) Then RDIR_RAT = 0.18 'Ratio
    If (RDIR_RAT > 1.7) Then RDIR_RAT = 1.7
    RDIR_RAT = (RDIR_RAT - 0.18) / 1.52

    'Function for the hidden neurons
    NetSum = 0.03620234
    NetSum = NetSum + GREEN * 22.74368
    NetSum = NetSum + RDIR_RAT * 6.679393
    HiddenNeuron(1) = 1.0 / (1.0 + Math.Exp(-NetSum))
    NetSum = 0.7550668
    NetSum = NetSum + GREEN * -3.907263
    NetSum = NetSum + RDIR_RAT * 10.81035
    HiddenNeuron(2) = 1.0 / (1.0 + Math.Exp(-NetSum))
    NetSum = -1.42035
    NetSum = NetSum + GREEN * -2.62758
    NetSum = NetSum + RDIR_RAT * 5.134643
    HiddenNeuron(3) = 1.0 / (1.0 + Math.Exp(-NetSum))

    'Accumulate results across all hidden neurons.
    NetSum = 5.523126
    NetSum = NetSum + HiddenNeuron(1) * -5.202937
    NetSum = NetSum + HiddenNeuron(2) * -1.52756
    NetSum = NetSum + HiddenNeuron(3) * -1.493804
    LAI = 1.0 / (1.0 + Math.Exp(-NetSum))

    'Final scaling of result and return
    LAI = 7.71 * (LAI - 0.1) / 0.8
    If (LAI < 0) Then LAI = 0
    If (LAI > 7.71) Then LAI = 7.71
    Return LAI

End Function
```

Fig. 1. Visual Basic function of neural network to predict LAI from ASTER band data values

Table 2. Comparison of regression and neural network models. In all cases, the neural network models were superior to the regression-based models for predicting LAI.

Model Reference	Building Method	Variables	*R*	*SEE*
Equation 3	Regression	GRRD_RAT	0.60	1.54
Equation 4	Regression	GRRD_RAT, IR	0.62	1.51
*	ANN	GREEN	0.69	1.39
*	ANN	NDVI	0.68	1.39
*	ANN	RDIR_RAT	0.69	1.39
Fig. 1	ANN	GREEN, RDIR_RAT	0.71	1.35

unsampled remainder of the urban area. This is demonstrated in Figure 2. In this example, LAI has been estimated for the Terre Haute region using the ANN represented in Figure 1. This kind of map may be useful when urban planners and others examine the distribution of LAI in urban and suburban areas.

While the ANN method proved most accurate in Terre Haute, this may not be the case in other urban areas under different environmental conditions. Future research could focus on these issues and determine whether ANNs provide the most accurate method to estimate LAI elsewhere. Also, care should be taken to ensure that the network algorithms and regression equations developed in this research are only applied in areas having similar solar zenith angles and vegetation types. While this study was completed at the landscape level, it suggests that artificial neural networks may be created and trained in other areas throughout the world to provide an accurate method to remotely estimate LAI. Further, these models can be used to answer important geographic questions by describing temporal and spatial LAI dynamics at landscape to regional scales (e.g., Jensen 2002). Of equal importance, this methodology can help land managers, conservationists, and urban foresters formulate urban environmental policy that is empirically supported by inexpensive remotely sensed biophysical data.

Fig. 2. Estimated LAI map of the Terre Haute area computed using the artificial neural network and three ASTER bands. Lighter areas represent higher LAI values. Note the city center in the middle left of the image and Terre Haute International Airport in the middle right of the image.

References

Aber, J.D. and J.M. Melillo. 1991. *Terrestrial Ecosystems*. Saunders College Publishing, Chicago.

Barclay, H. J. 1998. Conversion of total leaf area to projected leaf area in lodge pole pine and Douglas-fir. *Tree Physiology.* 18:185-193.

Benediktsson, J.A., Swain, P.H., and O.K. Ersoy. 1990. Neural network approaches versus statistical methods in classification of multisource remote sensing data. *IEEE Transactions in Geoscience and Remote Sensing.* 28:540-551.

Bischof, H., Shneider, W., and A.J. Pinz. 1992. Multispectral classification of Landsat-images using neural networks. *IEEE Transactions in Geoscience and Remote Sensing.* 30:482-490.

Chason, J.W., Baldocchi, D.D., and M.A. Huston. 1991. A comparison of direct and indirect methods for estimating forest canopy leaf area. *Agriculture Forestry and Meteorology*. 57:107-128.

Chen, J. M. and T. A. Black. 1992. Defining leaf-area index for non-flat leaves. *Plant Cell and Environment.* 15:421-429.

Chen, J. M., Rich, P.M., Gower, S. T., Norman, J.M., and S. Plummer. 1997. Leaf area index of boreal forests: theory, techniques, and measurements. *Journal of Geophysical Research-Atmospheres.* 102:29429-29443.

Ellsworth, D.S. and P.B. Reich. 1993. Canopy structure and vertical patterns of photosynthesis and related leaf traits in a deciduous forest. *Oceologia.* 96:169-178.

Fassnacht, K.S., Gower, S.T., MacKenzie, M.D., Nordheim, E.V., and T.M. Lillesand. 1997. Estimating the leaf area index of north central Wisconsin forests using the Landsat Thematic Mapper. *Remote Sensing of Environment.* 61:229-245.

Green, E.P., P.J. Mumby, A.J. Edwards, C.D. Clark, and A.C. Ellis. 1997. Estimating leaf area index of mangroves from satellite data. *Aquatic Botany* 58:11-19.

Gopal, S. and C.E. Woodcock. 1996. Remote sensing of forest change using artificial neural networks. *IEEE Transactions in Geoscience and Remote Sensing.* 34:398-404.

Hardin, P.J. 2000. Neural networks versus nonparametric neighbor-based classifiers for semisupervised classification of Landsat Thematic Mapper imagery. Optical Engineering. 39:1898-1908.

Haykin, S. 1994. *Neural Networks: a Comprehensive Foundation*. Upper Saddle River, NJ: Prentice-Hall.

Jensen, J.R. 2000. *Remote Sensing of the Environment: An Earth Resources Perspective*. NJ: Prentice-Hall.

Jensen, J.R. 2005. *Introductory Digital Image Processing: A Remote Sensing Perspective*, 3rd Ed. Upper Saddle River, NJ: Prentice Hall.

Jensen, J.R., Qui, F., and M. Ji. 2000. Predictive modeling of coniferous forest age using statistical and artificial neural network approaches applied to remote sensor data. *International Journal of Remote Sensing*. 20:2805-2822.

Jensen, R.R. 2002. Spatial and temporal leaf area index dynamics in a north central Florida, USA preserve. *Geocarto International*. 17 (4): 45-52.

Jensen, R.R., Boulton, J.R., and B.T. Harper. 2003. The relationship between urban leaf area and household energy usage in Terre Haute, Indiana, USA. *Journal of Arboriculture.* 29(4): 226-230.

Jensen, R.R. and M.W. Binford. 2004. Measurement and comparison of leaf area index estimators derived from satellite remote sensing techniques. *International Journal of Remote Sensing*. 25(20): 4251-4265.

Larcher, W. 1975. *Physiological Plant Ecology*. NY: Springer-Verlag.

Lymburner, L., Beggs, P.J, and C.R. Jacobson. 2000. Estimation of canopy-average surface-specific leaf area using Landsat TM data. *Photogrammetric Engineering and Remote Sensing.* 66:183-191.

Marascuilo, L.A. and J.R. Levin. 1983. *Multivariate Statistics in the Social Sciences*. Monterey, CA: Brooks/Cole.

Nel, E.M. and C.A. Wessman. 1993. Canopy transmittance models for estimating forest leaf area index. *Canadian Journal of Forest Resources*. 23:2579-2586.

Peper, P.J. and E.G. McPherson. 2003. Evaluation of four methods for estimating leaf area of isolated trees. *Urban Forestry and Urban Greening*. 2:19-29.

Pierce, L.L. and S.W. Running. 1988. Rapid estimation of coniferous forest leaf area index using a portable integrating radiometer. *Ecology*. 69:1762-1767.

Rouse, J.W., R.H. Haas, J.A. Schell, and D.W. Deering. 1974. "Monitoring vegetation systems in the Great Plains with ERTS." *Proceedings of the Third Earth Resources Technology Satellite-1 Symposium*, Greenbelt, MD: NASA SP-351, 3010-3017.

6 Public Participation Geographic Information Systems as Surveillance Tools in Urban Health

Daniel P. Johnson, Wright State University, Department of Urban Affairs and Geography

6.1 Introduction

Urban environments offer unique opportunities for researchers in the spatial sciences. Urban systems create their own microclimate, are the realization of intensive human activity and are strongly associated with high levels of human population (Steemers et al 1997). These aspects among others afford the opportunity for geospatial technologies to monitor and enhance humankind's relationship with the environment. However, far too often is this monitoring done by researchers or interested groups which are empowered to do so (Elwood 2006). Marginalized groups within urban settings rarely are offered the opportunity to participate in the development of a monitoring system, which could include geospatial technologies. Often these marginalized groups are not just lacking in the sense of policy decisions but their quality of health is often times inferior to more prominent groups within the city. This level of poor health is not defined and in order for it to be understood sufficiently, the level of spatial differentiation in health needs to be measured (Galea et al. 2005).

Public Participation Geographic Information Systems (PPGIS) have successfully been used to enhance the participation or "voice" of marginalized members of society (Elwood 2006). However, their use as a tool to enhance the quality of health and or health care has not fully been explored. PPGIS in this setting could be used effectively as a surveillance tool and as a public awareness program geared toward the health status of a community. Participation in such a program would need to be voluntary as many are reluctant to admit certain health conditions. However, the proposed

program could be used to provide health insight into the community without privacy infringement.

This chapter will propose a PPGIS intended to be useful in the management and surveillance of heat related illnesses in major urban centers across the United States. Heat related death is the number one cause of death due to extreme weather conditions; outnumbering lightening, hurricanes and tornados combined (NOAA, 2005). More often than not the deaths associated with such an event are manifested in marginalized groups within the city (Klinenberg, 2002; Browning et al., 2006). Enhancing the surveillance of these populations is of primary importance in understanding urban heat events and the death they cause (Bernard & McGeehin, 2004).

6.1.1 Urban Development & Medical Geography

Medical geography has historically been involved in the analysis of disease as it is associated with humankind. Central to this investigation is an understanding of the human-environment interaction, which includes heat related illness. Humans, as we know, have certain power over their environment as they alter the complex web of ecological interactions through urbanization and other activities. Additionally the environment has impacts on humankind such as Hurricane destruction of urban centers. This interplay has had significant impact on one developmental trend in medical geography which could include disease ecology, landscape epidemiology, and environmental health. This is more collectively associated with the geography of disease. Recently, Howard Frumkin published a text, Urban Sprawl and Public Health, whosemain premise was that one of humankind's greatest impacts on the environment, our cities and the sprawl associated with them, are a major culprit in current health problems in the United States. This underscores the relationship between human-environment interaction and medical geography.

It is known that specific infectious diseases can be associated with certain environmental stressors that humans may come into contact with. If these areas are known and are understood to be detrimental to health, then the disease association itself could be either eradicated or a least minimized. The *Culex pipiens* mosquito is a domesticated species which spreads WNV and EEE (Fonseca et al. 2004). The breeding sites of this mosquito are arguably increasing due to anthropogenically induced landscape alteration. Many other infectious diseases fit into this description. Additionally, John

Snow's cholera map and his subsequent spatial analysis leading to the Broad Street pump as the culprit of a cholera epidemic in the late 1800's could fit into this category. Both of these examples are related to the built-environment and its negative effect on health.

Not just infectious disease but chronic (non-infectious) maladies can also be traced to human environmental interactions. Current research suggests that chronic diseases such as cirrhosis or certain cancers have an environmental component. This can be understood with tracheo-esophageal cancer and its known relationship with proximity to coal burning electrical power generators in the UK (Openshaw et al. 1987).

Behavioral aspects of human activities are now beginning to be traced to the human-environment component as well. Recent studies are suggesting that drug addiction behavior may be influenced by components of the built environment or the loss of social capital as a result of suburbanization (Hembree et al. 2005). Depression also seems to be linked to poor-quality built environment. Depression can lead to a number of behaviors which can be detrimental to health.

This discussion would not be complete without recent developments on the nature of human health with regard to healthful environments. Health is not defined as just the absence of disease but is the entire quality of the healthy individual. It has been known for some time that some environments are more healthful than others. This plays into the human-environment interaction paradigm in medical geography. These healthy places are clearly just as important to delineate as the areas of poor health.

As mentioned, the main component of the human-environmental interaction paradigm on medical geography has been to define areas which equate to bad health. The emphasis on this component is also being viewed in the context of human induced global environmental change (Khasnis & Nettleman 2005). One component to this, understood in landscape simplification, one aspect of which is the decrease in biodiversity, can have a significant detrimental impact on human health. This in some examples is from bacterial or viral amplification due to a loss of natural predators or a simplified transmission cycle. Additionally, as global environmental change continues we can expect to see a change in the extent of certain disease amplification, biological, and mechanical vectors as well as possibly new disease components which have evolved in response to the change. We could clearly see tropical diseases extend further northward.

Geospatial technologies themselves have been studied as they relate to the surveillance of disease and global environmental change (Linthicum et al., 1987; Manguin & Boussinesq, 1999; Thomas & Lindsay, 2000; Ruiz et al., 2004). The efficacy of remote sensing with particular regard to the determination of habitats associated with insect-borne illness has been substantiated in numerous studies (Pope et al., 1994; Kitron et al., 1996; Hay et al., 1997). It is well known that the specific insect, or arthropod, which is most closely associated with enzootic disease, is the mosquito. The specific illnesses associated with this class of insect are numerous and in many cases are endemic to tropical areas. Clearly, an environmental change event could alter these extents.

Epidemiological characteristics have been studied using geospatial technologies for malaria (Thomas et al., 1997, 2000; Beck et al., 1997; Hay et al., 1996), rift valley fever (Linthicum et al., 1987, 1990, 1991), tryponosomiasis (Rogers et al., 1996; Kitron et al., 1996), schistosomiasis (Zhou et al., 2002; Seto et al., 2002; Xu et al., 2004), and West Nile Virus encephalitis (WNVE) (Theophilides et al., 2003; Ruiz et al., 2004). The primary vector for such encephalatides in the United States are the mosquitoes based in the *Culex* complex containing, among others, the species *Culex pipiens pipiens* and *Culex pipiens quinquefaciatus* (Fonseca et al., 2004). *Culex* mosquitoes demonstrate propensity for breeding in water that is highly polluted with organic content. Areas containing such contaminated water are typically in proximity to human areas of settlement, thus allowing *Culex* mosquitoes to share space with humans. This characteristic substantially increases the risk of human infection if viremic mosquitoes are existent. As suburbanization increases and urban decay continues the habitat of these particular mosquitoes will undoubtedly increase.

Clearly geospatial technologies have a place in the epidemiological investigation of certain disease complexes. However, the human component is much more difficult to investigate. Local knowledge is often necessary in such investigations and remote sensing and GIS have limitations in such endeavors. Public Participation Geographic Information Systems (PPGIS) have the potential to add the human component to such health surveillance in the form of narratives and high local spatial knowledge.

6.2 Health and the Built Environment

Recent published work in relationship to GIS and Remote Sensing in public health has concentrated on the location of disease vectors or areas of the environment which have a negative impact on human health. Some researchers are asserting that there are beneficial aspects to the environment which contribute to the enhancement of human health (Frumkin 2001). Many of these components are tied to us through the biological evolution of the human species.

Frumkin's main premise is that there are healthful environments that can just as easily be examined and delineated as unhealthful environments. Some of these healthful environments we can see recreated in our suburbanization settings and the planning associated with them. The location of nature close to human habitation and the development of parks with trees that spatially mimic the savannah of sub-Saharan Africa are two of the examples provided. The location of open water sources close to human habitation also reveals a sense of the return to nature. The development of walking paths to enhance exercise and mimic the hunter-gatherer renditions of the human psyche seems to play as well. These all seem to contribute to the betterment of human health. As noted older asylums understood the importance of nature as a location for the enhancement of health. Many of these asylums were developed for tuberculosis patients as it was thought movement to a healthy location would speed the recovery. As cited by Frumkin numerous studies suggest post-operative patients recover more quickly in natural surroundings than they do in closed hospital rooms.

Another component he mentions as a basis to good health is pet ownership; specifically canine as felines do not seem to hold the same association. Owners of dogs seemed to live longer and suffer from less depression than their non-pet owning or feline owing counterparts. This once again may be the formation in the human psyche that a dog offers protection as well as companionship. Pets may also allow the companionship bond to remain intact for elderly persons who have lost much of their social network.

These examples link well into the Biophyllia hypothesis. One component is that humans are drawn to nature especially that which is linked to our evolutionary development. We can see this by the location of our urbanization patterns as well as our mimicking of the environment in parks and

suburban settings. In this example humankind loves life (Bio Phyllia) and their exposure to it enhances health.

Geography and GIScience in particular could prove to be very fruitful in the examination of these types of settings. Clearly one could delineate areas of enhanced health and create an atlas of healthy environments. However, the true strength of this may be in exacting the specific spatial arrangement of nature and how that arrangement enhances health. This may lead to greater exemplars in sustainable development and urban planning, leading to healthier communities.

6.3 Public Participation Geographic Information Systems (PPGIS)

PPGIS systems have typically been promoted as a means of addressing issues for marginalized populations within a particular society. However, the use of such a system to monitor the health of a particular population is starting to emerge from its limited discussion in the PPGIS literature. A majority of these studies have dealt with planning mitigation or acquiring a more complete picture of the environment (Hassan, 2005; Cinderby and Forrester, 2005). Its use specifically as an epidemiological surveillance instrument is absent from the literature.

Of health related note in the PPGIS research is an article by Hassan in 2005 dealing with the effects of arsenic poisoning in Bangladesh. Arsenic is present in the groundwater drinking system of the country where there are over 300,000 reported cases of arsenic-induced cancer in the neighboring province of India. The method outlined by Hassan used local data as a method of identifying arsenic well locations. Persons intimately familiar with the surroundings are more knowledgeable than those from different areas. Based on this premise the identification of arsenic contaminated wells was fairly easy to establish. Hassan's development of the PPGIS system involved his own interaction with the local population and his guidance in focus groups. Thus leading to a conclusion that PPGIS can "bridge the information gap" between the community and the decision maker (Hassan, 2005).

Cinderby and Forrester describe a GIS for participation (GIS-P); differing from PPGIS is its level of community involvement. GIS-P differs from PPGIS in that the GIS is used to "facilitate participation" and whereby the

users are not responsible for the creation or maintenance of the GIS. Additionally, gathered from their research, there is a continuum of PPGIS-GIS-P implementation. They demonstrate that there exists a level of use and implementation ranging from users creating the data and not analyzing it to users analyzing the data but not creating it. Therefore GIS-P is more interested in the collection of local data rather than analysis where PPGIS is concerned with the inverse. For purposes of health relationships the principle concern should be with local data collection. Many studies have shown the added value of such local information and knowledge generation and it is not a significant stretch to suggest this is true with local health knowledge as well.

6.4 Heat Related Deaths

Heat related illness is the number one cause of human death in relation to extreme weather events in the United States, resulting in an average of 400 deaths per year over the past few decades (NOAA, 2005). Predominately these deaths have occurred in primary residencies in highly urbanized areas (NOAA, 2005; Naughton et al., 2002). Response and surveillance of this epidemic is highly lacking and most municipalities lack any response or planned intervention program. One of the main sources of information necessary to conduct a response plan is to identify the population which is at risk (Rothman & Greenland, 1998). The population at risk consists of elderly and very young persons; chronically ill and isolated individuals also are at increased risk of environmental hyperthermia (Klinenberg 2002). Apparently having a strong exacerbating effect on the heat wave within an urbanized center is the phenomena known as the Urban Heat Island (UHI). Spatial delineation of these phenomena should assist in the assessment of those areas within a city which are at increased risk of heat related morbidity.

The summer of 2006 saw a major increase in the publicity of heat related illness in the United States with heat warnings for St. Louis, Oklahoma City, Chicago, Dayton, Cincinnati, and Philadelphia. This is a growing problem with the prospect of global environmental change, increases in urbanization and an aging population. However, since the heat wave in Chicago of 1995, in which over 700 people died from hyperthermia, there has been action taken by local policy makers to relieve some of the situation. Chicago, for example, has built cooling centers in locations where they are believed to be needed. However, delineation of the at-risk groups

within the city needs to be achieved. A major component of such a research endeavor should include the incorporation of a PPGIS which will add to the local knowledge about particular locations within the city which are affected by the heat. Incorporating this local knowledge into a policy plan and surveillance tool could reasonably assist in intervention.

6.5 PPGIS for Heat Wave Surveillance

PPGIS is promoted as a means to increase public participation in the decision making process especially as it relates to spatial awareness. This aspect of PPGIS has the potential to decrease marginalization in groups in large scale social situations (Elwood 2005). Another research thread in PPGIS is the view that such tools could increase the marginalization and add to the isolation of certain groups. Central to PPGIS is the level of participation the public is involved in. Clearly, without significant participation the voices of those marginalized groups are not heard. PPGIS systems can be developed in a number of ways and system development depends on the goal of the system (Leitner et al., 1999). The system outlined here is a web-based PPGIS which is targeted at groups within a city which are at risk during urban heat events. The level of participation and the logistical development of participation parameters are central to discussions of PPGIS. This description will discuss the technology necessary and the public outreach model to make the system effective.

As discussed, the spectrum of PPGIS as a tool encompasses several different levels of community involvement. Ideally, the local community should create their own data for the PPGIS and then proceed to analyze it. Public health officials, especially in urban settings, do not have a solid grasp on the location of marginalized groups or at-risk groups or individuals in relationship to extreme weather events. Many times there is an understanding of these populations as related to infectious disease but not for chronic or environmental stresses. PPGIS could assist public health and other officials in determining locations of these at-risk individuals and indeed monitor the development of an urban heat event. Undoubtedly, such a system would be highly dependent on how it was developed and the level of community participation.

Developing such a system would be a very time consuming task and would in itself require community participation. Initially such development would require outreach to groups within the setting. This could follow

with focus groups using analog spatial information which would, at a later date, be digitally replicated, giving the community a voice in the development process and providing them with initial empowerment. GIS base layers already exist for a majority of the major cities within the United States. These layers could be used as the analog maps during the focus groups, thereby allowing for easier digitization. These focus groups however would focus on the usability of the system, before its development, and would give the community an indication about the reasons for the PPGIS.

Such focus groups would require usability specialists that would concentrate on making the system user friendly for each community involved. The interface could be different for each community or group that is participating. The component of the system which is used for monitoring, by the health officials, would be centrally located and each monitoring node would be connected for immediate update.

Before development it would be necessary to identify groups or individuals that will be using the system. Identification of areas where there is a high concentration of elderly individuals or areas of low socio-economic status, two of the most studied at-risk groups during an urban heat wave ,would be the initial phase. However, in such a monitoring system it would not be absolutely necessary to have a member of the at-risk group putting information into the system. An individual who lives in the setting with the at-risk population, but is not part of it, could be just as effective at monitoring those individuals of concern. Such an individual could provide relative point locations to the PPGIS of at-risk cases and give qualitative indicators of the stress to the community in an urban heat event. In the case of monitoring such events the analysis of qualitative, perhaps narrative information, during the event can provide significant insight into the dynamics of the group.

Such a PPGIS could be implemented in a number of ways. It is important that such a system be developed so that it may be tested for its efficacy in relation to a significant health event. Garnering information from the local population, which perhaps is the most effected, is central to such an undertaking. Additionally, such collaboration with the community could lead to increased resistance when intervening.

6.6 Summary

GIS and remote sensing is increasingly being used in public health circles as a means to track and monitor disease. The ability of GIS to visualize and display spatial information is unchallenged in public health. However, PPGIS is not thoroughly being examined in such a context. These systems have the ability to efficiently collect local data and allow an analyst to examine spatial interactions at that local level. Clearly such undertakings would assist in the monitoring of disease in many settings.

Urban areas offer high levels of complexity in many different dimensions. Spatial and cultural dimensions are just two that add problems to disease diffusion studies. Identification of many of these factors at the local level is very important in order to gain a solid understanding of disease in urban settings. PPGIS allows for such data collection and a solid push in this direction should address many of the current concerns in PPGIS research.

References

Beck, L.R. et al. (1997) Assessment of a remote sensing-based model for predicting Malaria transmission risk in villages of Chiapas, Mexico. American Journal of Tropical Medicine & Hygiene, 56, 99-106.

Browning, C. and Cagney, K.. (2006) Neighborhood Disadvantage andHealth. Forthcoming 2006 in Encyclopedia of Sociology. Malden, MA: Blackwell.

Cinderby S., Forrester, J. (2005) Facilitating the local governance of air pollution using GIS for participation. Applied Geography, Vol. 25, Issue: 3, pp. 143-158.

Elwood, S. (2006), Negotiating Knowledge Production: The Everyday Inclusions, Exclusions, and Contradictions of Participatory GIS Research. The Professional Geographer, Vol. 58, Issue: 2, pp. 197-208.

Fonseca et al. (2004), Emerging vectors in the *Culex pipiens* complex. Science, 303, 1535-1538.

Frumkin, H. et al. (2004) Urban Sprawl and Public Health: Designing, Planning, and Building for Health Communities. Island Press.

Frumkin, H. (2001) Beyond Toxicity: Human Health and the Natural Environment. American Journal of Preventative Medicine 20(3):234–240.

Galea, Sandro; Freudenberg, Nicholas; Vlahov, David (2005) Cities and population health. Social Science and Medicine, Vol. 60, Issue: 5, March,. pp. 1017-1033.

Hay, S.I. et al. (1996) Remotely sensed surrogates of meteorological data for the study of the distribution and abundance of arthropod vectors of disease. Annals of Tropical Medicine & Parasitology 90,1-19.

Hay, S.I., Packer, M.J. and Rogers, D.J. (1997) The impact of remote sensing on the study and control of invertebrate intermediate host and vectors for disease. International Journal of Remote Sensing 18, 2899-2930.

Hassan, M. (2005) Arsenic poisoning in Bangladesh: spatial mitigation planning with GIS and public participation. Health Policy, Vol. 74, Issue: 3, pp. 247-260.

Hembree, C. et al. (2005) The urban built environment and overdose mortality in New York City neighborhoods. Health and Place, Vol. 11, Issue: 2, pp. 147-156.

Kitron et al. (1996), Spatial analysis of the distribution of tsetse flies in the Lambwe Valley, Kenya, using Landsat TM satellite imagery and GIS. Journal of Animal Ecology, 65, 371-380.

Khasnis, Atul A.; Nettleman, Mary D. (2005) Global Warming and Infectious Disease. Archives of Medical Research, Vol. 36, Issue: 6, pp. 689-696

Klinenberg, E. (2002) Heat Wave: The Social Autopsy of a Disaster. University Of Chicago Press.

Leitner, H., (2000) Modes of GIS provision and their appropriateness for neighborhood organizations. Urban and Regional Information Systems Association Journal *pp. 43.*

Linthicum, K. J. et al. (1987), Detection of rift valley fever viral activity in Kenya by satellite remote sensing imagery. Science, 235, 1656-1659.

Linthicum, K.J. et al. (1990) Application of polar-orbiting, meteorological satellite data to detect flooding in Rift Valley fever virus vector mosquito habitats in Kenya. Medical Veterinary Entomology, 4, 433-438.

Linthicum, K.J. et al. (1991) Towards real-time prediction of Rift Valley fever epidemics in Africa. Preventive Veterinary Medicine, 11, 325-334.

Manguin, S.; Boussinesq, M. (1999) Apport de la télédéction en santé publique: l'exemple du paludisme et autres perspectives. Mâedecine et maladies infectieuses : revue de la Sociâetâe de pathologie infectieuse de langue franðcaise, 29, 318-324 (In French).

Naughton, Mary P. et al. (2002) Heat-related mortality during a 1999 heat wave in Chicago. American Journal of Preventive Medicine, Vol. 22, Issue: 4, pp. 221-227.

Pope, K.O. et al. (1994), Remote sensing of tropical wetlands for malaria control in Chiapas, Mexico. Ecological Applications, 4, 81-90.

Rogers, D.J., Hay, S.I. and Packer, M.J. (1996) Predicting the distribution of tsetse-flies in West-Africa using temporal Fourier processed meteorological satellite data. Annals of Tropical Medicine & Parasitology, 90, 225-241.

Rothman K., Greenland S. Modern Epidemiology. 2nd Ed. Lippincott-Raven, U.S.A., 1998.

Ruiz, M.O. et al. (2004), Environmental and social determinants of human risk during a West Nile virus outbreak in the greater Chicago area, 2002. International Journal of Health Geographics, 3, 1-11.

Steemers K., et al. (1997), City Texture and Microclimate, Urban Design Studies, Vol. 3, University of Greenwich, Kent.

Seto, E., et al. (2002) The use of remote sensing for predictive modeling of schistosomiasis in China. Photogrammetric Engineering and Remote Sensing, 68, 167-174.

Theophilidies, C.N., et al. (2003) Identifying West Nile virus risk areas: They dynamic continuous-area space-time system. American Journal of Epidemiology, 157, 843-854.

Thomas, C.J.; Lindsay, S.W. (2000), Local-scale variation in malaria infection amongst rural Gambian children estimated by satellite remote sensing. Transactions of the Royal Society of Tropical Medicine and Hygiene, 94, 159-163.

Thomson, M. et al. (1997) Mapping malaria risk in Africa: what can satellite data contribute? Parasitology Today 13, 313-318

Xu, B., et al. (2004) Snail density prediction for schistosomiasis control using Ikonos and ASTER Images. Photogrammetric Engineering and Remote Sensing, 70, 1285-1294.

Zhou, X.N., et al. (2002) Use of Landsat TM satellite surveillance data to measure the impact of the 1998 flod on snail intermediate host dispersal in the lower Yangtze River Basin. Acta Tropica, 82, 193-198.

7 Examining Urban Environment Correlates of Childhood Physical Activity and Walkability Perception with GIS and Remote Sensing

Gilbert C. Liu, Indiana Children's Health Services Research, Indiana University School of Medicine, Riley Hospital for Children, Indianapolis, Indiana

James Taylor Colbert, The Polis Center, Indiana University – Purdue University Indianapolis, Indianapolis

Jeffrey S. Wilson, Department of Geography, Indiana University – Purdue University Indianapolis, Indiana

Ikuho Yamada, Department of Geography, The University of Utah, Salt Lake City, Utah

Shawn C. Hoch, Department of Geography,Indiana University – Purdue University Indianapolis, Indiana

7.1 Introduction

Emerging research suggests that the built environment has potential to influence physical activity which, in turn, can have a protective effect against obesity and a positive impact on public health (Berrigan and Troiano, 2002; Atkinson et al., 2005). As a result, research on the association between the built environment and health is receiving increased attention in a variety of disciplines. Most research on the associations between the built environment and physical activity to date has focused on adults, but the potential links in children are largely unexplored. The present study examines how GIS and remote sensing can be used to enhance understanding of the relationships between physical activity and the built environ-

ment for a cohort of children from low-income urban neighborhoods in Indianapolis, Indiana.

Obesity has increased substantially in the United States over the last 30 years (Tremblay and Willms, 2000; Strauss and Pollack, 2001). The percentage of overweight children age 6 – 11 in the U.S. has grown from 4% in 1965 to 13% in 1999. The rate of overweight adolescents age 12 – 19 rose from 5% in 1970 to 14% in 1999 (Ogden et al., 2002). Obesity is costly in terms of decreased physical and psychological health (Hill and Peters, 1998; Cummins and Jackson, 2001; Liu and Hannon, 2005) -- in 1998 aggregate adult medical expenditures attributable to overweight and obesity is estimated to be $51.5 billion using Medical Expenditure Panel Survey data and $78.5 billion using 1998 National Health Accounts data. Research further suggests that increased obesity may eventually lower life expectancy in the U.S. (Olshansky et al., 2005).

Growing concerns over the obesity epidemic have prompted research into the potential effects of the built environment on physical activity and human behavior (e.g., King et al., 2000; Brownson et al., 2001; Pikora et al., 2003; Burdette and Whitaker, 2004; Hoehner et al., 2005). Promoting physical activity through environmental interventions adopts a population-based approach to behavior modification with the idea that altering the environment encountered by many people will have a greater cumulative impact on public health than individual intervention. Increasing physical activity through environmental modification that promotes routine physical activities, such as walking and cycling, may also be more effective because sedentary people are more likely to adopt moderate vs. vigorous forms of exercise and still accrue significant health benefits (Epstein et al., 1997; Frank and Engelke, 2001).

Using methods drawn mostly from transportation studies, researchers have correlated features of the built environment with both self-reported (diaries) and objective data on people's physical activity (Frank et al., 2005). Certain types of urban patterns, such as sprawl, correlate negatively with physical activity (Ewing et al. 2003). At the neighborhood level, research has identified a correlation between physical activity and street pattern, land use, and pedestrian infrastructure (Cervero and Duncan, 2003; Saelens et al., 2003[a]). An increased presence of supermarkets has been associated with increased fruit and vegetable intake in both Black and White adults (Morland et al., 2002). Studies focusing on the built environment's roles as a determinant of childhood overweight remain inconclusive.

Geospatial technologies (including GIS, GPS, and remote sensing) are increasingly employed to facilitate the collection of objective environmental measurements in support of physical activity research. These data can be both spatially and categorically comprehensive, providing information that may supplement or replace more costly field data collection. In addition, geospatial data may provide more objective observations than those obtained through the use of trained observers or self-reported by subjects. In the current study, variations in children's physical activity levels and perceptions of neighborhood walking environments were examined in relation to GIS and remote sensing measures of the built environment. Environmental variables were summarized at multiple radii around children's residences to assess how distance influences the associations between environment, physical-activity, and walkability perception.

7.2 Data and Methods

Data on 463 children's perceptions of neighborhood walkability, physical activity levels, family incomes, and body mass indices were obtained from the 2004 Summer Health Assessment Program Education (SHAPE) conducted by the Marion County Health and Hospital Corporation (Primary Investigator: Wanda S. Roddy, Marion County Department of Health). SHAPE is an annual program designed to medically evaluate and improve health care access for low-income children and their families in Indianapolis, Indiana. The program includes physical examinations conducted by physicians or nurse practitioners wherein children's height, weight, and body mass index (BMI) are determined. Children were classified by age- and gender-adjusted BMI percentiles, with a BMI greater than the 85^{th} percentile being considered 'at risk for overweight' and greater than the 95^{th} percentile being 'overweight'. The total number of children included in the 2004 program was 559.

As part of the SHAPE evaluation, children were asked to complete surveys on physical activity levels and perceptions of neighborhood walkability. Children's parents were asked to describe the child's medical history and family demographics. Demographic information collected as part of the survey included age, race, gender, annual family income, and overweight status. Five questions taken from the National Safety Council (2002) related to neighborhood walkability were included as part of the survey (Table 1). Subjects in SHAPE tend to be from poor minority families when compared to census indicators for the city as a whole. While the 2000

Census indicated that 14.7% of Indianapolis families have an annual income less than $15,000, 82.1% of SHAPE families fall below that threshold. The SHAPE survey included seven options for race. In the current study, race responses were condensed into 5 categories including one for no response. Similarly, income data were summarized in five categories including one for no response. Responses to the question, "How many days each week do you exercise, dance, or play sports?" are summarized in Figure 1.

Table 1. Subject characteristics and responses to SHAPE survey items**.**

Age	Younger than 5y		26	(6%)
	5 – 8y		192	(42%)
	9-12y	190	(41%)	
	13y and older	53	(11%)	
Sex	Female		225	(49%)
Race/ethnicity	Black	330	(71%)	
	Hispanic		42	(9%)
	White		45	(10%)
	Other			
Income	No Response	20	(4%)	
	Less than $9,000		205	(44%)
	$9,001 - $12,000		138	(30%)
	$12,001 - $15,000		35	(8%)
	Greater than $15,001	65	(16%)	
Weight Status	Normal Weight	216	(47%)	
	At Risk of Overweight	130	(28%)	
	Overweight		95	(21%)

Do you have room to walk?					
Yes	429 (93%)	No	28 (6%)	Unsure	4 (1%)
Is it easy to cross the street?					
Yes	388 (84%)	No	66 (14%)	Unsure	6 (1%)
Did drivers behave well?					
Yes	277 (60%)	No	123 (27%)	Unsure	57 (12%)
Were you able to follow safety rules?					
Yes	428 (92%)	No	14 (3%)	Unsure	17 (4%)
Was your walk pleasant?					
Yes	401 (87%)	No	21 (5%)	Unsure	28 (6%)

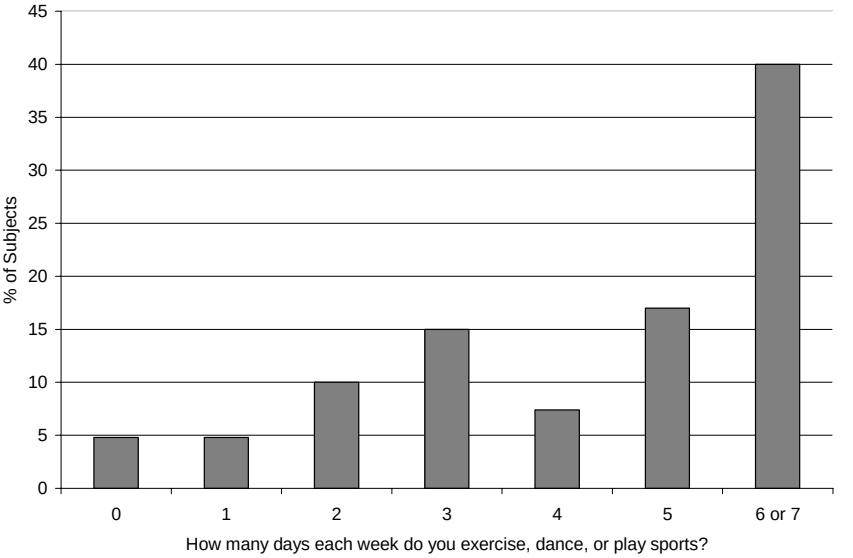

Fig. 1. SHAPE Subject Responses to the Question "How many days each week do you exercise, dance, or play sports?".

Address data collected as part of the survey were used to develop a GIS point layer representing the location of children's primary residences. Geocoding was accomplished using street center line and address range data provided by the Indianapolis Mapping and Geographic Information Service (IMAGIS) and standard GIS address matching routines. There were 559 individual children living at 390 unique addresses indicating that some locations included more than one child. Three hundred seventy of the 390 address were successfully geocoded (94.9% match rate) yielding point locations for 535 children (95.7%). Points located within a distance of 1km from the border of the Indianapolis city limits (n = 22) were excluded because of the lack of some geospatial data outside of the city boundary, leaving 513 (91.8%) child residential locations for analysis (Figure 2). Circular buffers were generated around the point locations representing children's residences at distances of 200m, 400m, 600m, 800m, and 1km in order to evaluate the influence of environmental characteristics at multiple radii. Environmental characteristics evaluated as independent variables included: crime density, street intersection density, residential population density, neighborhood median family income, speed limit and

traffic counts, sidewalk density, building offset, land use diversity, and the Normalized Difference Vegetation Index (NDVI).

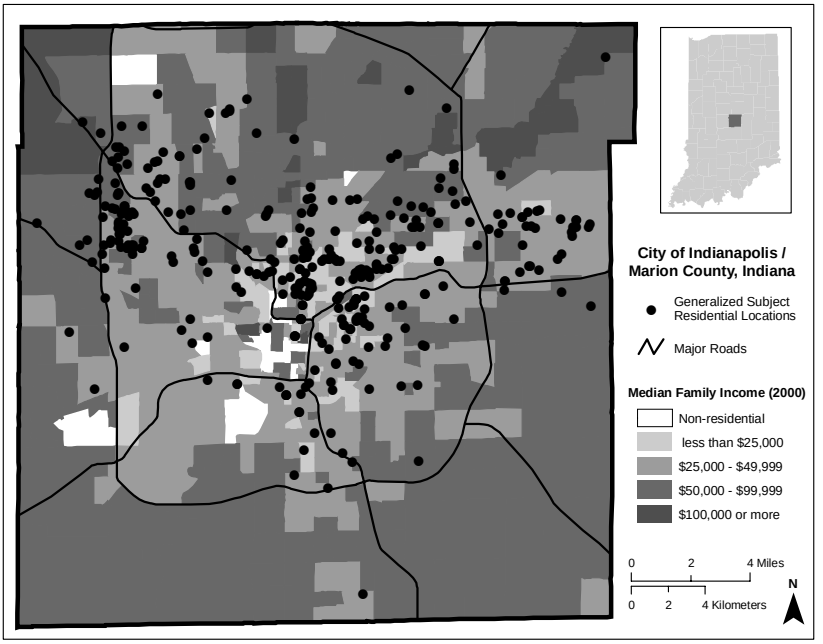

Fig. 2. Generalized SHAPE subject residential locations and median family income (2000) in Indianapolis / Marion County

Data on neighborhood median family income were obtained from the 2000 U.S. Census for all block groups in Marion County, Indiana. Median income data were converted to a 30m resolution raster and the mean value within each child's corresponding buffers was calculated. One limitation of this method is that the final calculation for a given buffer distance provides a mean of median family income and thus assumes uniform distribution of income across the region.

Building footprints and right-of-way were combined to determine average building offset. Indianapolis's public right-of-way GIS layer represents areas reserved for public use such as roads, sidewalks and bike lanes. The building footprint layer includes almost 500,000 polygons representing structures in Marion County. Buildings less than 600 square feet were excluded from the analysis to prevent small buildings, such as sheds, from

influencing the final results. Building offset was calculated by measuring the shortest distance from the edge of each building polygon to the right-of-way. The measurement was then applied to the points representing building centroids and the average of building offsets for all building centroids falling within a given buffer was calculated.

Intersection density has been used as a proxy for street connectivity, with more intersections indicating better connectivity (Frank et al., 2005). Street intersection density was calculated in the current study from a street centerline layer by determining the number of street network nodes within a given distance of children's residences. Nodes representing dead-end streets or cul-de-sacs were excluded. The number of intersections per unit area was summarized at each buffer distance around children's residences.

Traffic counts and speed limits of street segments falling within the residential buffers were averaged at each buffer distance. Traffic counts were available for arterial roads only, creating situations where some smaller buffers do not intersect with any arterial street segments generating a null value. Records that contain null values were excluded in subsequent analysis. The percentage of public streets serviced by sidewalks was estimated using Indianapolis's sidewalk and street centerline GIS layers. Street centerlines were divided into 30m sections and any section within 15m of a sidewalk was attributed as being serviced by a sidewalk. The total length of roads serviced by sidewalks was divided by the total length of all roads within a given buffer distance, yielding percent roads serviced by sidewalks around a child's residence.

Geocoded crime data for Marion County were obtained from the Indianapolis Police Department (IPD). The available data include all incident reports, regardless of whether or not a conviction was obtained. Incident reports for the 1998 calendar year were used in the current study as they were the most recent data available with complete coverage of the city. Incident report locations were aggregated by census block groups and normalized by the area of each block group yielding a value of annual crime density. This crime density variable was then rasterized to facilitate calculation of average crime density within each buffer distance around children's homes.

Current planning techniques, such as Smart Growth, are intended to reduce reliance on automobiles, increase walking for transportation, and reduce pollution from vehicle emissions (Saelens et al., 2003[b]). Mixed use neighborhoods that include amenities, such as sidewalks and crosswalks,

enhance walkability by placing people within practical walking distance of destinations (e.g. retail stores, post offices, and leisure facilities). Indianapolis zoning data were used to develop a land use dissimilarity index in the current study to capture land use heterogeneity around children's residences. These data delineate 70 unique types of zoning in the study area including 9 levels of residential density, commercial and industrial uses, and special use zones including schools, parks, cemeteries and churches. Land use data surrounding a hypothetical residential location are depicted in Figure 3 along with selected features of the built environment.

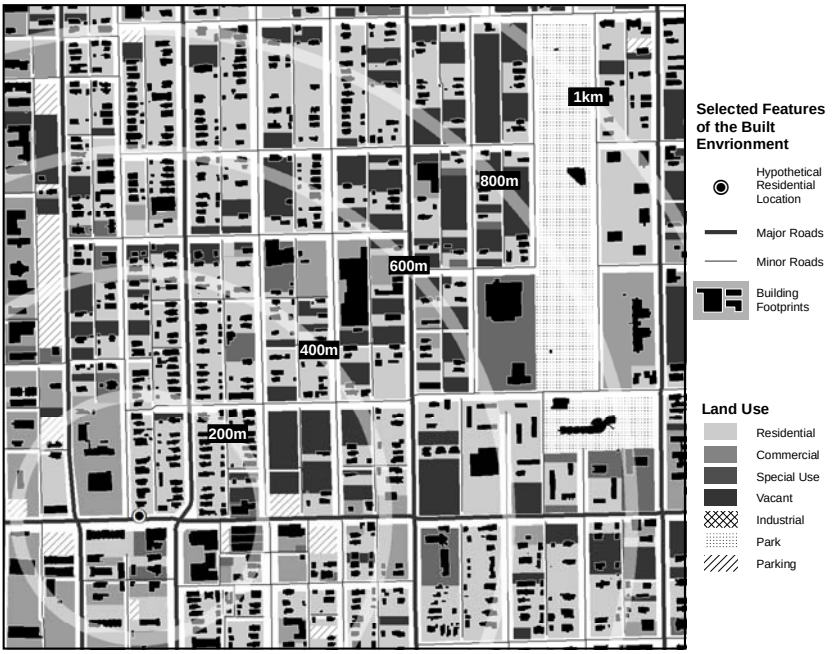

Fig. 3. Land Use and selected features of the built environment surrounding a hypothetical residential location.

The present study employed a version of the land use dissimilarity index developed by Frank et al. (2005) modified to adjust for the focus on children rather than adults. Frank's system divides land use into three categories (residential, commercial, and office) with the premise that commercial and office districts provide destinations for those living in residential areas. Recreational zones (including parks, playgrounds, and school yards) were substituted in place of office space in the current study, assuming these land use types would be more likely destinations for children than offices.

Data from the 2000 U.S. Census were used to develop residential population density estimates within children's residential buffers following the protocol of Forsyth et al. (2005). Average population density was calculated by weighting the census block group population density by the area of the block group falling within a child's residential buffer. The estimated population inside the buffer was divided by the total area of residential zoning, yielding a density measure that accounts for variation in land use.

A walkability index was calculated using the equation of Frank et al. (2005) that combines population density, intersection density and land use dissimilarity:
Walkability index = (6 x Z-score of land-use mix)
+ (z-score of net residential density)
+ (z-score of intersection density)
The results will vary based on the z-score inputs, which ranged from -15.67 to 18.47 in the current study.

The normalized differential vegetation index (NDVI) is a remotely sensed measure of the presence and condition of vegetation (Lillesand et al., 2005). NDVI values range from -1 (bare ground) to +1 (dense, healthy vegetation). NDVI data were calculated from Landsat Enhanced Thematic Mapper Plus (ETM+) imagery acquired May 8, 2001 using the equation (Near Infrared Band - Red Band) / (Near Infrared Band + Red Band) (Tucker, 1979) (Figure 4). Average NDVI values were calculated at each of the five buffer distances around children's residences.

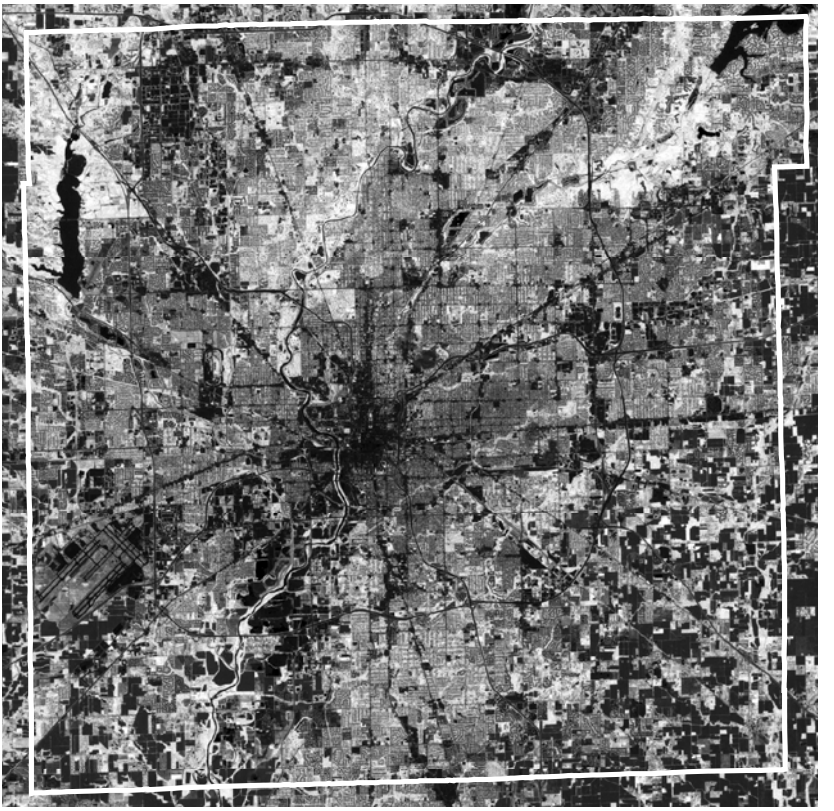

Fig. 4. Enhanced Thematic Mapper Plus (ETM+) Normalized Difference Vegetation Index (NDVI) Image of Indianapolis / Marion County, Indiana (May 8, 2001). Brighter tones indicate higher NDVI values.

Methods

Analysis was conducted in three phases designed to examine: 1) the relationship between children's perceptions of neighborhood walkability and environmental variables, 2) the relationship between children's self-reported physical activity and the built environment, and 3) the performance of regression models for predicting physical activity level of children using their perceptions of walkability, built environment variables surrounding their residences, and individual-level sociodemographics.

To examine relationship between children's perceptions of neighborhood walkability and environmental variables, T-tests were conducted to determine whether environmental variables had significantly different mean values between the groups of children responding in the affirmative or negative to the National Safety Council (NSC, 2002) neighborhood walkability survey questions. The question "Do you have room to walk?" was compared to building offset and percentage of roads serviced by sidewalks. Responses to the question "Was your walk pleasant?" were compared to NDVI and the land use dissimilarity index. Street intersection density was compared to two questions: "Were you able to follow safety rules?" and "Was it easy to cross the road?" The question related to drivers' behavior was analyzed against speed limits and traffic counts.

To examine relationship between the built environment and children's self-reported physical activity levels, T-tests compared mean values of each independent variable at multiple radii for children in high versus low physical activity categories. High physical activity was assigned to children reporting six or more days per week with physical activity. Our choice of this dichotomous outcome was informed by the U.S. Department of Health and Human Services recommendation that children and adolescents engage in 60 minutes of moderate exercise most days, preferably daily (Pate et al., 1995). We chose a cut-point of six days per week because the variation of responses was skewed and this threshold represented a natural break in the distribution (Figure 1).

Both logistic and linear regression models were developed to predict physical activity levels of children using their perceptions of walkability, built environment variables, and individual-level sociodemographics. Built environmental variables were included in the multivariate models if the associated T-test in phase 1 of the analysis was significant at the 0.05 level. Black was used as the reference category for race during regression analysis. Family income level of $9,000 - $12,000 was used as the reference category for income.

7.3 Results

Two environmental variables, NDVI and speed limit, were found to have significant relationships to children's neighborhood walking perceptions. NDVI was positively associated with the question "Was your walk pleasant?" at three buffer distances: 600m, 800m and 1km. Children who re-

ported experiencing a pleasant walk in their neighborhood had higher mean NDVI values compared to children who did not rate their neighborhood as providing a pleasant place for walking (mean NDVI at 600m buffer = 0.095, SD ± 0.087 for children with pleasant walks versus 0.054 ± 0.079 for those without, p-value 0.03). The relationship became stronger as the buffer distance increased (mean NDVI for 1km buffer = 0.093 ± 0.085 for children with pleasant walks versus 0.045 ± 0.08 for those without, p value 0.02). Speed limit had significant relationships with perceptions of driver behavior at all buffer distances. In response to the question "Did drivers behave well?" children who affirmed drivers in their neighborhoods as "well-behaved" had significantly lower mean speed limit values within their residential buffers (e.g. mean speed limit at 1km buffers 29.4 ± 2.4 mph for those children reporting well-behaved drivers versus 30.2 ± 1.9 mph, p-value 0.001).

Children living in poorer neighborhoods, with more traffic, higher speed limits, more intersections, higher crime rates, and lower NDVI tended to report higher physical activity levels. Mean intersection density (156.3 ± 71.7 intersections in high activity children versus 144.0 ± 55.4 intersections in low activity children, p-value 0.04) and average speed limit (29.9 ± 2.2mph in high activity children versus 29.4 ± 2.5 mph in low activity children, p-value 0.02) differed significantly for the physical activity subgroups only at the 1km buffer distance. Mean traffic counts differed significantly for the physical activity subgroups at 400m and 800m with the strongest relationship observed at 800m (916 ± 305 vehicles per day for high activity children at 800m versus 851 ± 326 vehicles per day for low activity children, p-value 0.03). Density of crimes showed a significant and positive relationship only at the 200m buffer distance (3164 ± 2543 annual crimes for high activity children versus 2711 ± 2287 annual crimes for low activity youth, p-value 0.04). NDVI and physical activity were negatively related at every buffer distance except 200m, with 1km showing the strongest relationship (mean NDVI at 1km buffer for high activity children 0.079 ± 0.09 versus 0.074 ± 0.01 for low activity youth, p-value 0.009). Neighborhood median family income had a negative relationship with physical activity at every buffer distance except 1km, with 200m showing the strongest relationship ($38,312 ± $14,943 for high activity children versus $42,225 ± $15,476 for low activity children, p-value 0.006).

These results led to the transfer of the following variables to the regression models in part three of the analysis: intersection density at 1km, average speed limit at 1km, traffic counts at 800m, crime density at 200m, NDVI at

1km and neighborhood median family income at 200m. The dependent variable for the linear regression model was the number of days per week children reported being physically active. The dependent variable in the logistic regression setting was a binary indication of high or low physical activity level as previously defined. Seven significant relationships emerged in the linear model and five in the logistic model. The significant relationships in the linear model were: Hispanics were less active; older children were less active; children overweight or at risk of being overweight were less active; children in the lowest family income (demographic variable) category were more active; children near roads with higher traffic counts were more active; children perceiving well-behaved drivers were more active; and children reporting pleasant walks were more active. The models estimated that children reporting presence of pleasant walks in their neighborhood would be expected to report an additional day per week of activity. Children who perceived drivers as well behaved in their neighborhoods would be expected to report an additional one half-day per week of activity than children reporting that neighborhood drivers were not well behaved. The logistic model resulted in significant relationships for age, overweight status, reported family income, traffic counts and pleasant walks, all in the same direction as the linear model. Complete regression model results are shown in Table 2.

Table 2. Regression model using demographic, neighborhood walkability, and geospatial environmental variables to predict self-reported physical activity levels (Significant at .05 in bold).

	Linear Regression $r^2 = .105$		Logistic Regression $r^2 = .197$		
	Outcome: Number of Days Active Standardized		Outcome: Physical Activity High or Low Standardized	Standard	
Variable	Beta	Sig.	Beta	Error	Sig.
Age	-0.131	**0.016**	-0.102	0.042	**0.015**
Hispanic	-0.124	**0.020**	-0.836	0.513	0.103
White	-0.073	0.233	-0.559	0.414	0.177
Other Race	0.010	0.850	0.325	0.495	0.512
Female	-0.014	0.786	-0.201	0.238	0.398
BMI > 85th percentile	-0.148	**0.004**	-0.711	0.241	**0.003**
Income under $9K	0.255	**0.000**	0.825	0.301	**0.006**
Income $12-$15K	0.045	0.423	-0.239	0.472	0.612
Income > $15K	-0.007	0.901	-0.425	0.366	0.246
Intersection Density 1km	-0.045	0.620	0.001	0.003	0.821
Speed Limit 1km	0.035	0.525	0.053	0.056	0.347
Traffic Count 800m	0.140	**0.014**	0.001	0.000	**0.008**
NDVI 1km	-0.092	0.233	-1.741	2.059	0.398
Median Family Income	-0.120	0.080	0.000	0.000	0.217
Crime 200m	-0.054	0.541	0.000	0.000	0.584
Room to Walk	0.010	0.860	0.695	0.577	0.228
Easy to Cross Road	0.023	0.687	0.487	0.413	0.238
Well Behaved Drivers	0.162	**0.007**	0.489	0.302	0.105
Follow Safety Rules	0.007	0.891	-0.453	0.808	0.575
Pleasant Walks	0.156	**0.004**	1.429	0.630	**0.023**

7.4 Discussion

This study explored the use of geospatial data to enhance analyses on the associations between physical activity levels in children and the built environment in which they live. Estimates of environmental variables quantified using GIS and remote sensing were compared to children's perceptions of their walking environments and reported levels of physical activity. Variables representing walking perceptions, demographics, and the environment were used in regression models predicting physical activity. Self reported variables representing perceptions of the built environment were significantly related to GIS variables meant to function as objective measures of the built environment. Different sizing of analytic buffers influenced associations between environmental variables and physical activity. We found that after controlling for individual demographic factors, individual weight status, and neighborhood socioeconomic

status, the following factors related to the built environment were significant predictors of physical activity in children: presence of heavier traffic, perceiving drivers to be well-behaved, and perceiving walks to be pleasant.

The findings of significant associations between child self-perceptions of driver behavior and neighborhood infrastructure comprising pleasant walking paths are, to our knowledge, novel. These findings compliment studies that heretofore have included only adult subjects. The presence of heavy traffic has been associated with higher levels of physical activity in adults and adults who perceive enjoyable scenery along walking routes have been shown to be more likely to walk (Brownson et al., 2001). In contrast to the present study, the work by Brownson et al. obtained quantification of traffic levels and pleasing scenery from subjects' survey responses, while the current study employed both survey methods as well as direct observations recorded as GIS variables.

Of particular relevance to this study are urban design initiatives such as New Urbanism (Katz 1994) and the Ahwahnee Principles (Corbet and Valesquez 1994) for resource-sufficient communities. These initiatives call for the development of integrated communities containing mixed residential, commercial, and open recreational land uses in easy walking distance of one another and transit stops; the design of streets and paths that slow traffic and result in a fully connected system of intersecting routes to all destinations; and the design of neighborhoods to support diversity in age and socioeconomic status levels of residents, encouraging the presence of people in public space. Although this study cannot elucidate how heavier traffic, driver behavior, and pleasant walks promote youth physical activity, it can be argued that each of these phenomena could be related to the availability of destinations and pedestrian safety, both of which are logically supportive of walking.

Many of the environmental factors that were examined did not emerge as significant predictors of child physical activity, in contrast to several previously published studies. Norman et al. (2006), for example, found that girls in areas of lower intersection density were more likely to exhibit lower levels of physical activity. Independent environmental variables common to both Norman's and the present study include residential density, intersection density, land use mix, and a walkability index. Gordon-Larsen et al. (2000) found serious crime was associated with a lower likelihood of moderate to vigorous physical activity in adults. The current study indicated no relationship between crime and physical activity level in children, but made no distinction between serious and non-serious crime

as in the Gordon-Larsen study. Further research is needed to assess how crime levels in neighborhoods may differentially influence persons of various ages. Consistent with the results of Hoehner et al. (2005), who found sidewalk density was not a significant predictor of walking for pleasure or transportation among adults, no significant relationship between physical activity and the presence of sidewalks was observed for children in the current study. The distinction between recreational and utilitarian walking may be especially relevant for the younger subjects in our study, as children may be more apt to engage in active play along streets in contrast to adult utilitarian purposes (e.g. walking to work or to retail areas).

Limitations of this study include the narrow demographic scope (predominance of racial minorities and families with lower incomes) of the subjects, survey questions that were not designed specifically to facilitate GIS analysis, and potentially unreliable survey answers due to the young age of the respondents. The selection of specific buffer sizes may not have resulted in the most informative variable – possibly analytic areas smaller or larger would have resulted in more explanatory variables. Another problem may have been with the process of summarizing data in the buffers – in using the mean values for the analytic buffers, we may have missed important specific information. It is possible that much of the physical activity reported occurred at a location away from the children's home, such as at school or church, to which the children took vehicular transportation.

The associations between age, gender, and differences in physical activity in our study are consistent with national surveillance of youth health behaviors (Troiano and Flegal 1998). Interpretation of the relationship between family income and physical activity must take into account that the study population was drawn mainly from children of poor families attending free or reduced-cost summer camp programs. One interpretation could be that the poorest of the poor are more physically active than those with a higher economic standing. A possible explanation for this finding is that children in the lowest socioeconomic class would have less access to sedentary forms of entertainment such as TV and video games. The association may substantially change if more representation of families with high income were included. The relationship between income and physical activity may be curvilinear as more affluent children may have better access to more expensive forms of physical activity that require costly equipment.

7.5 Conclusions

Our study adds to an accumulating body of literature suggesting that aspects of the built environment are determinants of child weight status, ostensibly influencing health behaviors such as physical activity. Geographic information systems and remote sensing have potential to complement other means of collecting exposure data for a wide range of environmental and public health analyses. The current study contributes to previous research in two ways: the focus on children as opposed to adults, and the use of GIS-based variables alongside related survey data to quantify the urban environment in terms of walkability. Examining urban environment correlates of childhood physical activity and walkability perception through GIS and remote sensing should inform future public policy aimed at the prevention and management of obesity in the United States.

References

Atkinson, J.L., Sallis, J.F., Saelens, B.E., Cain, K.L. and Black, J.B. 2005. The association of neighborhood design and recreational environments with physical activity. *American Journal of Health Promotion*, 19(4): 304-309.

Berrigan, D., and Troiano, R.P. 2002. The association between urban form and physical activity in U.S. adults. *American Journal of Preventive Medicine*, 23(2), 74-79

Brownson R, Baker E, Housemann R, Brennan L, and Bacak S. 2001. Environmental and policy determinants of physical activity in the United States. *American Journal of Public Health*. 91(12): 1995-2003.

Burdette HL, and Whitaker RC. 2004. Neighborhood playgrounds, fast food restaurants, and crime: Relationships to overweight in low-income preschool children. *Preventive Medicine.* 38(6): 57-63.

CDC. 2005. Prevalence of Overweight among U.S. Children and Adolescents. http://www.cdc.gov/nchs/data/nhanes/databriefs/overweight.pdf - last accessed September 15, 2006.

Cervero R., and Duncan M. 2003. Walking, bicycling, and urban landscapes: Evidence from the San Francisco Bay area. *American Journal of Public Health*, 93(9): 1478-1483.

.Corbett J, and Valesquez, J. September 1994. The Ahwahnee Principles: Toward More Livable Communities. *Western City.*

Epstein L, Saelens B, Myers M, and Vito D. 1997. Effects of decreasing sedentary behaviors on activity choice in obese children. *Health Psychology*. 16(2):107-113.

Ewing R, Schmid T, Killingsworth R, Zlot A, and Raudenbush A. 2003. Relationship between urban sprawl and physical activity, obesity and morbidity. *American Journal of Health Promotion, Inc*, 18(1): 47-57.

Forsyth A, Koepp J, Oakes J, Schmitz K, and Zimmerman J. 2005. *Environment and Physical Activity: GIS Protocols.* University of Minnesota, Twin Cities Walking Study, Metropolitan Design Study, 210 pp.

Frank L, Schmid T, Sallis J, Chapman J, and Saelens B. 2005. Linking objectively measured physical activity with objectively measured urban form. *American Journal of Preventive Medicine,* 28(S2): 117-125.

Frank L, and Engelke P. 2001. The built environment and human activity patterns: Exploring the impacts of urban form on public health. *Journal of Planning Literature,* 16(2): 202-218.

Gordon-Larsen P, McMurray RG, and Popkin BM. 2000. Determinants of adolescent physical activity and inactivity patterns. *Pediatrics,* 105(6): e83. http://www.pediatrics.org/cgi/content/full/105/6/e83 - last accessed September 28, 2006.

Health and Hospital Corporation of Marion County Web Page (No Date). http://www.hhcorp.org/shape.htm - last accessed September 28, 2006.

Hill J, and Peters J. 1998. Environmental contributions to the obesity epidemic. *Science,* 280(5368):1371-1374.

Hoehner C, Brennan Ramirez L, Elliot M, Handy S, and Brownson R. 2005. Perceived and objective environmental measures and physical activity among urban adults. *American Journal of Preventive Medicine,* 28(S2): 105-116.

Hume C, Salmon J, Ball K.2005. Children's perceptions of their home and neighborhood environments, and their physical activity: A qualitative and quantitative study. *Health Education Research,* 20(1): 1-13.

Jackson L. 2002. The relationship of urban design to human health and condition. *Landscape and Urban Planning,* 64(4):191-200.

Katz P. 1994. *The New Urbanism: Toward an Architecture of Community.* New York: McGraw-Hill, 288 pp.

King A, Castro C, Wilcox S, Eyler A, Sallis J, and Brownson R. 2000. Personal and environmental factors associated with physical inactivity among different racial-ethnic groups of U.S. middle-aged and older-aged women. *Health Psychology,* 19(4): 354-364.

Lillesand T, Kiefer R, and Chipman J. 2005. *Remote Sensing and Image Interpretation, 5th ed.* New York: John Wiley and Sons, 763 pp.

Liu GC, and Hannon T. 2005. Reasons for the prevalence of childhood obesity. Genetic Predisposition and Environmental Influences. *The Endocrinologist,* 15(1): 49-55.

Mokdad AH, Ford ES, Bowman BA, Dietz WH, Vinicor F, Bales VS, Marks JS. 2003. Prevelence of obesity, diabetes, and obesity-related health risk factors. Journal of the American Medical Association, 298(1): 76-79.

Morland K, Wing S, and Roux A. 2002. The contextual effect of local food environment on residents' diets: The atherosclerosis risk in communities study. *American Journal of Public Health,* 92(11): 1761-1768.

National Safety Council. 2002. *Walkable America Checklist: How Walkable Is Your Community?* http://www.nsc.org/walk/wkcheck.htm - last accessed September 15, 2006.

Norman G, Nutter S, Ryan S, Sallis J, Calfas K and Patrick K. 2006. Community design and access to recreational facilities as correlates of adolescent physical activity and body mass index. *Journal of Physical Activity and Health*, 3(S1): S118 – S128.

Olshansky SJ, Passaro DJ, Hershow RC, Layden J, Carnes BA, Brody J, Hayflick L, Butler RN, Allison DB, Ludwig DS. 2005. A potential decline in life expectancy in the United States in the 21st Century. *New England Journal of Medicine*, 35(11): 1138-1145.

Pate R, Pratt M, Blair S. 1995. Physical activity and Ppublic health -- A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. *Journal of the American Medical Association*, 273:402-407

Pikora T, Giles-Corti B, Bull F, Jamrozik K, and Donovan R. 2003. Developing a framework for assessment of the environmental determinants of walking and cycling. *Social Science and Medicine*, 56(8): 1693-1703.

Saelens[a] B, Sallis J, Black J, and Chen D. 2003. Neighborhood-based differences in physical activity: An environmental scale evaluation. *American Journal of Public Health*, 93(9): 1552-1558.

Saelens[b] BE, Sallis JF and Frank LD. 2003. Environmental correlates of walking and cycling: Findings from the transportation, urban design, and planning literatures. Annals of Behavioral Medicine, 25(2): 80-91.

Strauss RS, and Pollack HA. 2001. Epidemic increase in childhood overweight, 1986-1998. *Journal of the American Medical Association*, 286(22): 2845-2848.

Timperio A, Crawford D, Telford A, Salmon J. 2004. Perceptions about the local neighborhood and walking and cycling among children. *Preventive Medicine*, 38(1): 37-47.

Tremblay MS, and Willms JD. 2000. Secular trends in the body mass index of Canadian children. *Canadian Medical Association Journal*, 163(11): 1429-1433.

Troiano RP, Flegal KM. 1998. Overweight children and adolescents: description, epidemiology, and demographics. *Pediatrics*,101(3 Pt 2):497-504.

Tucker C. 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of the Environment*, 8(2): 127-150.

8 Mapping, Measuring, and Modeling Urban Growth

Perry J. Hardin, Department of Geography, Brigham Young University, 690 SWKT, Provo, UT 84602, USA

Mark W. Jackson, Department of Geography, Brigham Young University, 690 SWKT, Provo, UT 84602, USA

Samuel M. Otterstrom, Department of Geography, Brigham Young University, 690 SWKT, Provo, UT 84602, USA

Immediately after World War II, developers in the United States took advantage of market demand and government incentives to build new housing subdivisions for returning soldiers anxious to marry, begin families, and resume civilian life. New developments such as Levittown (New York), Park Forest (Illinois) and Lakewood (California)[1] sprang up and were quickly filled with affordable cookie-cutter homes for veterans seeking the American Dream of suburban home ownership (Hayden 2003). The baby boom followed. As a result of the boom and international immigration, the U.S. population grew from 151 million to 300 million between 1950 and 2007. To accommodate this expanding population growth, cities and towns in the U.S. rapidly spread into their rural hinterlands.

An extensive system of highway access is the primary stimulus for suburban sprawl in the United States today. The 1956 Interstate Highway Act created a system of highway construction heavily subsidized by the federal government; and multilane freeways soon radiated from every major U.S. city (Burchell et al. 2005). In suburban "bedroom communities" with easy highway access, suburban American families could build spacious homes on cheap land in new communities, and could also commute quickly to the central city for work, shopping and services (Hayden 2003, p. 165-166).

---

[1] Levittown, Park Forest and Lakewood are located respectively near New York City, Chicago, and Los Angeles.

Unmanaged urban growth has come with both an economic and social price tag. The bill includes items such as traffic congestion, social isolation, pollution, higher infrastructure costs, and economic / ethnic segregation (Burchell et al. 2005; Morris 2005; Deal and Schunk 2004; Carruthers and Ulfarsson 2003; Lindstrom and Bartling 2003). However, the price of farmland displacement and ecological damage has not yet been paid-in-full. Urban growth in the United States is predicted to consume seven million acres of farmland, seven million acres of environmentally sensitive land, and five million acres of other lands during the period 2000 to 2025 (Burchell et al. 2005, 38-39).

The problem of unmanaged urbanization is acute in newly industrialized countries. According to United Nations estimates, (UNCHS 2001), nearly half of the Earth's 6.5 billion humans are urbanites. In newly industrialized countries; high birthrates, land tenure practices, political policy, environmental degradation, epidemics, and economics are largely responsible for the continued movement of humanity from farm and pasture to city. The migrant influx can overwhelm local government's ability to provide critical sanitation, transportation, and housing infrastructure to satisfy human need. Many of the fastest growing cities in newly industrialized countries have overcrowded population cores and zones of dense squatter ghettos on their periphery. Living conditions in the countryside may be bad, but the city is a squalid cardboard and concrete kingdom of human degradation, debilitating disease, and rank injustice. Misery is king. Much will be required to force a regime change, but geographic technology for monitoring urban sprawl is one weapon to deploy.

This chapter is about mapping, measuring, and modeling cities with geospatial technologies. Evaluating the extent and structure of urban growth and land conversion over time is crucial for informed and responsive regional planning. If landcover and landuse change can be periodically mapped, future development can be forecast. Assessment of policy effectiveness is also made possible. Novel techniques for directing urban expansion can be invented, evaluated and applied to improve the lives of urban humanity.

8.1 Urban growth and planning policy

This section is an overview of planning methods used to manage city growth within the United States. As urban geographers, we do not suggest

they are globally applicable, but we do think they represent useful points for discussion.[2] Although their effectiveness is mixed, their objective is to maintain urban vitality and limit rural land loss.

In the past two or three decades, pushback against unmanaged growth in the US has been felt as cities have expanded rapidly into rural zones. This reaction has come from farmers, environmentalists, CBD business owners, social scientists, and others who noticed the unintended consequences of unbridled urban expansion. Growth management policies and laws have been subsequently developed by local, state, and federal authorities. Private companies have also taken an active voluntary interest in promoting development that follows sound planning practices.

Bengston et al. (2004) divided urban growth control policies into three categories: public acquisition, regulation, and incentives. Public acquisition involves the purchase and ownership of lands by a public entity that subsequently prohibits development of the land. Areas targeted for purchase are often environmentally sensitive or have historical or recreational importance. Regulation includes the enforcement of local zoning laws, state management acts, and planning ordinances. On a small scale, regulation can mean specification of minimum building lot sizes. On a large scale, it can include designating growth rate controls and placing moratoria on future development. Incentives and disincentives are also used for regulating urban growth. Developers might obtain benefits from creating projects that meet government planning goals. Conversely, developers might pay heavily for developments that have greater negative impacts on the community. Incentives could include housing density bonuses for providing open space or tax credits for rehabilitating historic buildings. Disincentives could include development impact fees that are charged to builders in addition to normal permit fees (Bengston et al. 2004, 275-276).

Growth management acts at the state level are found in at least eleven of the United States. Their general aim is to direct rapid population growth and development in a manner harmonious with state goals (Carruthers 2002). One common component of state management acts is the mandate for communities to develop policy instruments like urban growth boundary

[2] For international comparisons of urban growth patterns and policies, see Davis and Schaub (2005); Frenkel (2004); Tomalty (2002); Alexander and Tomalty (2002); Dawkins and Nelson (2002); Raad and Kenworthy (1998); and Summers et al. (1999).

limits, open space ordinances, and transportation plans. In part, the success of these devices relates to implementation consistency across the state, implementation concurrency, and the state's enforcement ability (Carruthers 2002; Ben-Zadok 2005). The effectiveness of growth management acts in reducing urban sprawl has been mixed (Anthony 2004; Kline 2000; Nelson 2000). Some researchers have observed that growth management plans might increase potential environmental damage in attempts to limit urban expansion (Audirac et al. 1990; Burby et al. 2001).

The concept of "smart growth" encompasses a popular set of ideas and policies to promote compact community development. Although the goals of smart growth supporters sometimes diverge, the movement has been increasing its calls for a new planning paradigm (Downs 2001; Lindstrom and Bartling 2003, 65-67). Typical elements of smart growth include more mixed land-use areas, more compact building footprints, multiple housing types, walkable urban designs, and multiple modes of public and private transportation. Smart growth frequently focuses on redevelopment of existing neighborhoods (Smart Growth Network 2003; Burchell et al. 2005). Unfortunately, smart growth principles are sometimes difficult to implement because they conflict with the tradition of low-density urban development in the United States (Downs 2001). Additionally, some smart growth policies have been shown to encourage sprawl (Irwin and Bockstael 2004).

Other investigations have focused on different aspects of growth policy. These include:

- Studies of policy to increase open space, protect wildlife habitat, and limit environmental impact (Kline 2006; Radeloff et al. 2005; Dwyer and Childs 2004; Howell-Moroney 2004; Robbins and Birkenholtz 2003; Johnson 2001; Bernstein 1994; Lewis 1990).
- Studies of how rural areas are confronting urban sprawl (Mattson 2002; Theobald 2005; Weiler and Theobald 2003).
- Investigations of agricultural land preservation efforts (Kashian and Skidmore 2002; Nelson 1992; Nelson 1990).
- Research to document the effects of growth on transportation. This includes modifying transportation plans to focus or channel growth (Buliung and Kanaroglou 2006; Kuby et al. 2004; Cervero 2001; Willson 1995).

The growth management studies and policies cited above have a strong geographic component – landscape space is at the heart of urban growth problems. Most of the studies also illustrate the need for accurate geospatial data mated with geocomputational techniques to monitor expansion

over time. With a proper set of geographic data and tools and data, informed policy can be made. In addition, the success of growth management schemes can be accurately assessed and quickly modified to meet management objectives.

8.2 Mapping urban growth

8.2.1 Change detection

The goal of any urban change detection method is to identify the changes with problem relevance while ignoring the inconsequential. Some kind of remote sensing is almost always used. Several of the most popular change detection methods using overhead imagery are described below.

Image overlay is a simple method to visually identify the location and extent of change (Jensen 2005). It is most commonly done by inserting single bands of imagery from different dates into the red, green, or blue (RGB) image planes of a graphics card for display on a monitor (Banner and Lynham 1981). When examining change between two dates of imagery, a single band (from one date) can be inserted into both the blue and green image planes whilst the same band (for the second date) is placed in the red image plane. Landtype changes that modulate reflectance in the selected bands will appear bright red while areas with little change will appear gray. For change assessment between three dates, it is possible to insert corresponding bands into all three RGB image planes. Resulting shades of cyan, yellow and magenta would then show areas of change while gray would indicate unchanged zones.

This simple approach can be powerful in urban change detection, particularly when monitoring urban fringe areas where vegetation conversion to impervious surface produces a large reflectance change. Although image overlay does not quantify change, it allows the analyst to see its extent and location. She can then focus subsequent quantitative procedures on the areas and change classes revealed. By noting the bands in which change is obvious, image overlay is also a rapid way to select bands for inclusion in automated change detection techniques. Finally, this method does not require specialized image processing software and is easily understood by those unfamiliar with image processing.

Post-classification comparison is currently the most popular method of urban change detection (see Table 8.1). In post-classification comparison, each date of rectified imagery is independently classified to fit a common landtype schema. Following classification, the resulting landcover maps are overlaid and compared on a pixel-by-pixel basis. The result is a map of landtype change. This per-pixel comparison is also summarized in a "from-to" change matrix (Jensen 2005). The change matrix displays every landtype change class possible under the original classification schema and shows the acreage of each change class.

The utility of the change matrix helps explain the popularity of post-classification comparison. Its conceptual simplicity is also appealing. Furthermore, when each image is independently classified, atmospheric corrections are not necessary (Kawata et al. 1990; Song et al. 2001). This method is also appropriate when dates of imagery are not phenologically compatible, have substantially different sun or look angles, or come from different sensors (Yang and Lo 2002).

Despite its prevalence, post-classification comparison is error prone. A major source of error is the method's high dependence upon the accuracy of the independently classified landtype maps (Aspinall and Hill 1997; Serra et al. 2003) For example, if two landtype maps each have a producer's accuracy of 90%, the accuracy of the post-classification change map accuracy will be about 81% (Stow et al. 1980).

Another method used extensively for change detection is **spectral-temporal classification**. This method uses a composite image created by adding (i.e., "stacking") the bands from multiple dates of imagery together to form a single image composite. If two six-band images are stacked, the result is a 12 band composite. The composite image is then classified into change classes using traditional image processing techniques.

Simple image differencing (i.e., **image subtraction)** is also a widely used urban change detection algorithm. The technique requires the selection of corresponding bands from two dates of imagery. The difference image $D(x)$ is created by subtracting the brightness values (x) of one image from the other on a per-pixel basis. While all bands can be differenced in this way, bands can also be deliberately chosen to maximize detection of important change classes.

The numerical values in the output difference image are the signed (i.e., $-\infty$ to $+\infty$) changes in reflectance between the two image dates. Complexities

aside, unchanged areas will have values near zero while areas with significant change will be progressively positive or negative depending on the change direction and magnitude. The difference image can then be used to create a binary mask of change vs. no change. This mask *M(x)* is created by defining a threshold value (T) such that:

$$M(x) = \begin{cases} 1, & if \left|D(x)\right| > \mathrm{T} \\ 0, & otherwise \end{cases}$$

The threshold choice is crucial. It is selected to separate relevant and real changes from those irrelevant or created by noise. If the threshold value is too small, the difference mask will overestimate the area of change. Conversely, an overly large threshold would exclude many bonafide areas of change from the mask.

Table 8.1 A summary of urban change detection methods

Method	Reported usage
Change vector analysis	Chen and Gong et al. 2003
Decision trees	Im and Jensen 2005; Chan et al. 2001
Econometric panel	Kaufmann and Seto 2001
IHS transform	Chen and Chen et al. 2003
Image differencing	Todd 1977; Quarmby and Cushnie 1989; Fung 1990; Ridd and Liu 1998; Sunar 1998; Bruzzone and Prieto 2000; Masek et al. 2000; Maktav and Erbek 2005; Liu et al. 2004
Image ratioing	Maktav and Erbek 2005; Liu et al. 2004
Image regression	Ridd and Liu 1998; Liu et al. 2004
Kauth-Thomas image differencing*	Fung 1990; Ridd and Liu 1998; Kaufmann and Seto 2001; Seto and Fragkias 2005
Learning vector quantization	Chan et al. 2001
Artificial Neural networks	Dai and Khorram 1999; Chan et al. 2001; Liu and Lathrop 2002; Pijanowski et al. 2005; Bruzzone 1999
PCA	Fung and LeDrew 1987; Fung 1990; Fung 1992; Yeh and Li 1997; Li and Yeh 1998; Sunar 1998; Liu et al. 2004
Post-classification comparison	Fung 1992; Jensen et al. 1995; Li and Yeh 1998; Sunar 1998; Ward et al. 2000; Chan et al. 2001; Madhavan et al. 2001; Yang 2002; Yang and Lo 2002; Chen and Gong et al. 2003; Chen and Chen et al. 2003; Weber and Puissant 2003; Alberti et al. 2004; Liu et al. 2005; Mundia and Aniya 2005; Xiao et al. 2006; Yu and Ng 2006
Spectral-temporal classification	Schneider et al. 2005

Vegetation index comparison	Maktav and Erbek 2005

* Also known as tasseled-cap image differencing

The analyst usually determines the threshold value empirically (Jensen 2005; Singh 1989). To set thresholds that make the change map meaningful, the analyst needs to understand both the target and the project goals. (Schowengerdt 1997). Researchers have also proposed statistical methods of selecting the threshold. These include Fung and LeDrew (1988), Ingram et al. (1981), Morisette and Khoram (2000), Smits and Annoni, (2000), Rosin (2002), Rosin and Ioannidis (2003), Coppin et al. (2004), Singh (1989), and Yuan et al. (1999).

Despite these statistical aids, problems interpreting change magnitude often make impossible the definition of a global threshold value. For example, a reflectance difference of 30% in the urban periphery may indicate conversion from agriculture to urban residential land, while the same reflectance change in the CBD may result from street repaving.

Input values with vastly different magnitudes can also create the same difference value in the subtraction process. Although reflectance *magnitude* on the two images would indicate different change class involvement, change classes generating the same reflectance *difference* would not be differentiated in the change product. Coppin and Bauer (1994) suggest dividing the difference by the sum of the input values to mitigate this problem. It is also possible to produce separate change masks for different types of change through the use of density slicing (Singh 1989).

In addition to the problems cited above, the technique sensitivity to noise and bi-temporal illumination differences make it a questionable choice for urban change detection. It is a particularly poor choice when using high resolution imagery wherein shadows and slight illumination changes cause large changes in radiance recorded by the sensor.

Image ratioing is also commonly used for change detection. It proceeds like image differencing with simple division being used in place of subtraction. The resulting per-pixel quotients constitute a ratio image with image pixels unchanged between image dates taking a value of 1.0 and areas of change producing higher or lower values depending on the change classes involved. Because the output quotient distribution is non-Gaussian and bi-modal, it is impossible to set a numeric threshold of meaningful change (Coppin and Bauer 1994).

Image ratioing was first used in vegetation studies with Landsat Multispectral Scanner data (Rouse et al. 1974). It was successfully used in an urban environment by Todd (1977) who reported 91.4% correct change classification of Atlanta, Georgia. Howarth and Boasson (1983) found that a bitemporal ratio of TM5[3] provided information on urban change while a similar ratio of TM7 failed to provide information on urban change in Hamilton, Ontario.

Image differencing and ratioing are commonly performed on images previously transformed to enhance important change classes. These transforms include the following:

- The Kauth-Thomas or tasseled-cap transform (Ridd and Liu 1998; Rogan et al. 2002)
- The Normalized Differenced Vegetation Index (NDVI) (Cakir et al. 2006; Du et al. 2002; Masek et al. 2000; Song et al. 2001)
- The Principal Component Analysis transform (Du et al. 2002; Millward et al. 2006)
- Miscellaneous ratios and transformations. For example, Schott et al. (1988) found that an infrared to red ratio was effective in separating urban pixels from water and vegetation.

Obviously a number of transformations can be concocted. The transformation selected is often dependent upon the scene properties and project design. The right method is the one that works to solve the problem. Determining what works usually requires trial and error.

[3] See Table 8.2 for satellite and band specifications and nomenclature.

Table 8.2 Satellites and sensors discussed in the chapter.

	TM / ETM+	**SPOT HRV & HRVIR**	**IKONOS**
Satellites	Landsat 4, 5, 7	SPOT 1, 2, 3, 4	IKONOS-2
Pixel size	30 m	20 m	4 m
Blue band	TM1 0.45-0.52 μm		XS1 0.45-0.52 μm
Green band	TM2 0.52-0.60 μm	SPOT1 0.50-0.59 μm	XS2 0.51-0.60 μm
Red band	TM3 0.63-0.69 μm	SPOT2 0.61-0.68 μm	XS3 0.63-0.70 μm
IR band	TM4 0.76-0.90 μm	SPOT3 0.79-0.89 μm	
SWIR band	TM5 1.55-1.75 μm	VGT 1.58 – 1.75 μm (SPOT 4 HRVIR)	
MIR band	TM7 2.08-2.35 μm		
Pan band	EPAN (Landsat 7 ETM)	SPAN	IPAN

The operating assumption of change detection through **image regression** is that image reflectance values recorded for the same location on two dates are linearly related. To detect change using this logic, a linear least-squares regression is performed using the pixel values of image at time t_1 as the independent variable and the corresponding pixel values at t_2 as the dependent variable. The change map is a map of regression residuals. It is created by subtracting the regression-predicted t_2 reflectance values from the actual t_2 pixel reflectance values. Pixels with large residuals (based on a threshold) are used to generate a change/no-change mask (Singh 1989).

Image regression has some theoretical advantages. The regression accounts for differences in reflectance mean and variance between dates. It also reduces the effects of different sun angles and atmospheric transparency (Coppin et al. 2004). Despite these advantages, Ridd and Liu (1998) found that image regression techniques performed only slightly better than simple image differencing in a Salt Lake City study. Because of the linearity assumption, this technique is not recommended if a large proportion of the target area has changed between the two image dates.

Principal components analysis (PCA) is a multivariate statistical method for data summarization and reduction. The fundamental assumption of

PCA is the existence of some underlying structure between a set of physical variables that can be described by a smaller set of synthetic variables. PCA is used to reduce image dimensionality by defining new, uncorrelated bands composed of the principal components (PCs) of the input bands. PCs are computed by examining the correlation between input image bands, grouping highly correlated bands, and then calculating new bands that summarize the information contained in the original band set. See Duda et al. (2001) for computational details.

The creation of new components can be envisioned as the translation and rotation of the original image axes in feature space to match patterns in the input variables. The first principal component (PC1) is calculated to capture the maximum variation resident in the original data. Each subsequent component is calculated such that it accounts for the maximum possible variance remaining, and it is orthogonal to the antecedent PC axes. In the process of PCA, new PC bands are created by projecting the original data values in terms of the new axes (Jensen 2005). The interpretation of PC bands is coaxed from a table of factor loadings that summarize the correlation between input bands and PC bands.

Principal components can be calculated two ways. Standardized PCA uses the input bands' correlation matrix whereas unstandardized PCA uses the covariance matrix (Singh and Harrison 1985; Eastman and Fulk 1993). Various studies have examined the utility of both variants in change detection with general agreement that standardized PCA is superior; it produces increased signal to noise ratio (SNR) and effective image enhancement (Eklundh and Singh 1993; Fung and LeDrew 1987; Singh and Harrison 1985).

Application of PCA in urban change detection is based on the assumption that pixel brightness values will remain constant through time if landtype remains unchanged. If this condition holds, a high per-pixel correlation will be found between different image dates. The most common method of utilizing PCA in change detection is to combine two images with n bands into a single image with $2n$ bands and subject it to PCA. The computation will produce $2n$ components. For targets where change is limited in spatial area, PC1 will portray the unchanged portion of the landscape, while successive components will represent areas of change. If a scene is dominated by change then PC1 will highlight the changed areas. Various landtype changes may also be sufficiently dissimilar that each will be captured in different components (Byrne et al. 1980, Fung and LeDrew 1987).

Sources of change detection error

Practitioners of urban change detection should remember that changes in apparent reflectance can be caused by factors other than landtype conversion. According to Du et al. (2002) and Paolini et al. (2006), these error factors include:

- Atmospheric attenuation
- Misregistration between multiple image dates
- Topographic attenuation
- Different phenological stage and/or seasonal variability
- Sensor spatial, spectral and radiometric resolution differences
- Changes in sensor response for same sensor due to drift or age
- Changes in viewing and/or sun angle.

Of these error sources, atmospheric differences and image misregistration are the most detrimental to change detection (Coppin and Bauer, 1994). Apparent differences due to changes in sun angle and phenostage can often be eliminated or reduced to ignorable levels by using imagery collected by the same sensor on anniversary dates (Jensen, 2005). Unfortunately, the paucity of cloud free imagery for many world regions prevents the satellite collection of anniversary images. In a sample of 32 urban change detection projects we surveyed, only 14 used near-anniversary dates.

8.2.2 Comparative studies

It is important to remember that no single change detection approach can be globally recommended. Since urban change detection is a machine learning problem, two theorems of Duda et al. (2001, 454-461) come into play as alternative methods are compared. The first is the *No Free Lunch Theorem*.[4] In consequence of the theorem, any apparent accuracy advantage of a classifier is strictly a function of the fit between algorithm and problem dataset, not an inherent advantage of the classifier itself. Consider two competing classifiers used to detect change. For every problem where the first classifier produces higher accuracy, another dataset could be devised where the converse was true. Over an infinite number of problems, the average accuracy difference between the two algorithms would be zero. Following Duda et al., we recommend that change detection practitioners be leery of any novel machine learning method purporting to be

[4] Duda et al. (2001, p. 456) attribute the theorem name to David Haussler.

universally superior. In addition to classifier selection, careful consideration of training sample size, noise minimization, etc. is required for urban change detection success.

The *Ugly Duckling Theorem* is the corresponding statement for feature representation. "In the absence of assumptions there is no privileged or 'best' feature representation, and … even the notion of similarity between patterns depends implicitly on assumptions that may or may not be correct" (Duda et al. 2001, 458). This suggests we limit expectations built upon past experience. Even when guidelines seem reasonable based on previous success; the Ugly Duckling Theorem invites us to explore alternative features, feature representations, feature sets, and similarity measures. We should not expect one combination to be superior to another under all conditions; the results will be dependent on the data. The same applies to spectral band selection. Different band combinations will be optimal for detecting change in different contexts.

Given the trial and error required for successful urban change detection, a detailed review of comparative studies is instructive. These projects illustrate the experimentation necessary before adopting a single methodological solution for a particular project.

The goal of Ridd and Liu (1998) was to map urban change in Salt Lake City, UT between 1896 and 1990 by comparative analysis of two TM images. Results from four change detection algorithms were compared, three of which share general popularity: 1) image differencing, 2) image regression, and 3) tasseled-cap image differencing. Using 585 field validation sites, Ridd and Liu found that no combination of algorithm or TM band was "absolutely superior to the others" (p. 100). For simple binary change (i.e., changed vs. unchanged) mapping, image differencing and image regression using TM2 and TM3 produced the highest accuracy values. Detection via image differencing of TM4 produced the lowest accuracy.

A more complex picture of algorithm competency emerged when specific change categories were examined. For example, image differencing of TM2 was effective in detecting conversion from construction site to new residential land, new residential sites to vegetated residential land, and conversion from farmland to construction sites. In contrast, the green tasseled-cap difference image was superior for detecting the transformation from green to dry farm. Generally speaking, temporal alterations that modulated relative percentages of bare ground and vegetation were most easily detected by all three algorithms (Ridd and Liu, 1998).

Fung (1990) provides another comparative example. The objective was to determine whether Landsat TM would be a useful data source for monitoring landscape change in the Kitchener-Waterloo area, Ontario. After visual study, ten change classes were described. Eight of the ten were related to change between agricultural landtypes (e.g., pasture to grain conversion), whereas the remaining two focused on urbanization: 1) change from construction to residential land and, 2) construction site replacement of agricultural land.

The change detection methods compared by Fung included image differencing, PCA, and tasseled-cap image differencing. Image differencing of all TM bands between the two summer image dates (1985 and 1986), yielded six difference images from which change could be assessed. A standardized PCA was also performed on the two stacked images. Although the process created 12 PCs, only PC3, PC5, and PC6 were related to change. PC3 was interpreted as a brightness change image, PC5 displayed near-infrared reflectance change, and PC6 was related to wetness differences. Because the Kitchener-Waterloo region was predominantly agricultural, tasseled-cap image differencing was also tried. Like the retained PCs, the three tasseled-cap image layers were labeled brightness, greenness, and wetness change respectively.

The accuracy of the change as depicted on the 12 change images was assessed by field investigation and comparison with air photography. Accuracy was measured on both a per-change class basis and a reduced binary change scale. As expected, no image depicted all change classes with equivalent accuracy. For example, difference images produced from TM1, TM2, and TM3 were very competent to detect rural to urban conversion and construction to residential transformations, but were ineffective in detecting change between agricultural land types. In contrast, the TM4 difference image detected change between some agricultural classes but not others.

Fung observed that temporal changes in the area primarily involved variation in target brightness and greenness. In evaluating brightness change, Fung concluded that the tasseled-cap brightness difference image was superior to PC3. The tasseled-cap greenness difference image was also superior to greenness principal component PC5. The tasseled-cap greenness difference image provided the overall best information for monitoring change in the agricultural categories.

In an effort to map the change in Instanbul, Turkey, Sunar (1998) compared Landsat images dated August 1984 and September 1992 using four common methods. These included the visual image overlay of TM3, image differencing, PCA of stacked images, and spectral-temporal classification. Visual analysis of the results demonstrated different strengths of the different approaches but no specific conclusions were offered. Sunar suggests that production ease, product information content, and product interpretability are the primary criteria by which change detection procedures should be judged. We recommend two additional criteria. The first criterion is the ability to accurately map change on a binary scale. This implies a method that equally minimizes errors of both change underestimation and overestimation. The second criteria is tougher – to detect change uniformly and accurately across all the change classes of importance. This uniform accuracy requirement follows from Anderson et al. (1976, 5).

8.2.3 Innovative approaches

As shown in Table 8.1, innovative approaches to urban change detection have been introduced in recent years. All are more complicated than popular traditional methods. One of the research challenges faced by urban change detection practitioners is making these innovative approaches available to the broader user community; only then can the domain of their effectiveness be delimited.

Change vector analysis represents one innovative strategy for landtype change mapping. Following Chen and Gong et al. (2003), the concept of a change vector is best described by example. Spectral information in a single pixel is represented in its reflectance vector. In a study utilizing TM, the vector can be expressed as $\mathbf{R} = (r_1, r_2, r_3, r_4, r_5, r_7)^T$ where r_1 is the reflectance for the pixel in TM1, r_2 the reflectance in TM2, etc. Given two dates of imagery, reflectance vectors for the same pixel location on the two dates can be expressed as $\mathbf{R_1}$ and $\mathbf{R_2}$. The change vector is then defined as $\mathbf{\Delta R = R1 - R_2}$. Calculation of the vector's magnitude and direction follow simple vector formulae, resulting in a two band difference image where one band records the magnitude of the change vectors and the second records the direction (Johnson and Kasischke, 1998). As with simple image differencing, a threshold value defining relevant change in magnitude and direction must be chosen. The research question of Chen and Gong et al. (2003) was how to best specify this limit. Chen and Gong et al. present the threshold determination method and apply it to a CVA case study in Haidian District, Beijing, PRC. Using TM imagery from 1991 and 1997,

the authors calculated the change vectors using TM1-TM5 and TM7. The threshold of change was specified and the changed areas mapped. Using the 1991 TM imagery, a hybrid classification generated a 1991 landtype map of nine classes. The classes included urban fabric, barren land, water, paddy fields, wheatland, vegetables, forest, shrub / grassland, and other land types. Based on this classification and the change vector analysis previously performed, the specific change classes were enumerated. Two types of change dominated the scene. The first was the conversion of agricultural land to fishponds whereas the second was archetypical urban intrusion into fringing agricultural zones. Both change types were a result of economic development in the Haidian District.

The research of Chen and Gong et al. (2003) is notable for the rigorous accuracy assessment. The validation sample included 425 sites of documented change and 1975 sample areas that remained unchanged between the two dates. Only 13% of the areas of known change were missed by the CVA. Only 2% of the unmodified sites were mistakenly classified as changed. Overall, the classification produced a kappa of 0.87. A post-classification comparison using the same images produced a kappa of only 0.69 and severe overestimation of change acreage. Classwise accuracy for the CVA was dependent on the change class considered. Conversion of water to other landtypes was correctly mapped 93% of the time whereas conversion of paddy fields was correctly mapped only 81% of the time. Conversion of other seven landtype classes ranged between those two values.

Spectral mixture analysis (SMA) is a method of quantifying urban landcover components within a pixel. It is based on the principle that the signature of a single pixel is a combination of the signatures of the pixel integrants weighted by the area of each component. For example, consider an urban pixel covered by 30% grass and 70% concrete. The signature of the pixel would be a composite – the sum of the signatures of the grass and concrete components, each multiplied by 0.30 and 0.70 respectively to account for the relative amounts of both materials. In SMA parlance, pure grass and pure concrete in this example are spectral endmembers.

Alberti et al. (2004) is an excellent example of SMA use in urban change detection. Their first objective was to differentiate between categories of urban land change within the Puget Sound area using Landsat TM and ETM+ imagery acquired in 1991 and 198. The study was focused on lowland basins in the area experiencing landtype alteration due to urbanization. This second goal was to conduct a detailed analysis of landcover

change between the two dates. The methodology was complex, detailed, and used both supervised classification and SMA. The supervised classification was designed to differentiate between broad landtype categories such as paved urban, mixed urban, vegetation, and bare soil. Using shade, pavement, and green vegetation spectral endmembers, the urban class was then subdivided according to impervious material (i.e., pavement). Similarly, grass and forest classes within the vegetation class were sorted out using bare soil, shade and green vegetation endmembers. Significant spectral confusion plagued the discrimination of bare agricultural soils, pavement, and clearcuts; a sequence of heuristic filters and rules were applied to limit the impact of the confusion. The landtype classes were then identified according to their spectral mixture components. For example, the paved urban class was constituted from the urban and mixed urban classes containing more than 75% impervious material.

The landtype classifications for 1991 and 1999 reported in Alberti et al. produced kappa coefficients of 0.92 and 0.88 respectively. A per-pixel overlay of the two maps was done to determine landtype change between the two dates. The largest source of error in the resulting map was seasonality differences between the two years. Although the second image (1999) was acquired on the anniversary date of the first (1991), record precipitation totals created significantly greener 1999 vegetation. Problems with this seasonality coupled with registration errors and spatial heterogeneity required a resampling to 90m resolution for the final change map product. Comparison with orthophotography revealed that the accuracy of the change detection map was 85%. Areas of no change were mapped very accurately. Otherwise, accuracy differed by change class:

- Forest to mixed urban; accuracy = 90%
- Forest to bare soil; accuracy = 88%
- Grass / bare soil to mixed urban; accuracy = 65%
- Mixed urban to paved urban; accuracy = 80%
- Bare soil to paved urban; accuracy = 83%.

Like SMA and change vector analysis, **artificial neural networks** (ANNs) are another innovation with apparent promise for urban change detection. As the name implies, ANNs attempt to model the low-level structure of the brain. As such, they are the product of research into artificial intelligence. Although the brain's computational method is very different from a computer (Haykin 1994), ANNs tend to mimic the biologic brain's fault tolerance and learning capacity. The neuron is the basic unit of an ANN. A neuron stores knowledge by the application of a learning process algo-

rithm. Interneuron connection strengths (i.e., synaptic weights) store the knowledge that is learned. ANNs are frequently used as alternatives to statistical procedures such as regression and discriminant analysis because they estimate functions adaptively rather than through some constrained mathematical algorithm like ordinary least squares. As summarized by Jensen and Hardin (2005), "ANNs have been used in remote sensing applications to classify images (Bischof et al. 1992; Hardin 2000), and incorporate multisource data (Benediktsson et al. 1990). ANN classifiers have been successfully used with remote sensing data because they take advantage of the ability to incorporate non-normally distributed numerical and categorical GIS data and image spatial information (Jensen et al. 2000)" (p. 23).

In supervised classification, feed-forward back-propagation ANNs can replace the typical maximum-likelihood classifier used to assign image pixels to a training class. Other ANNs genres can serve as unsupervised classifiers or hybrid combinations of both strategies. For example, the Learning Vector Quantizer is a Kohonen self-organizing neural network map that can be adapted to either supervised or unsupervised learning strategies (Kohonen 1995).

As mentioned by Liu and Lathrop (2002), ANNs have promise for detecting change between satellite images, but the high cost of network training hinders their widespread adoption. Because of this problem, the authors developed a method for training such ANNs that is quicker than traditional approaches while also facilitating efficient feature extraction. Although the article focuses on the training algorithm, the success of the neural network is also noteworthy. The data source for their training experiment was two anniversary TM images of Barnegat Bay, New Jersey acquired 11 years apart. Thirteen of the possible 49 landtype change combinations among seven broad landtype classes required discrimination.

Two feed-forward back-propagation ANNs were tested. Although the two networks differed in the number of hidden neurons, the primary difference was the pattern vectors presented to each. The first network was trained with a stacked 12-band image created from the visible and infrared bands of the two TM images. The second network was trained with a six band PC set from the 12-band image. When compared to the accuracy of the mapping produced by post-classification overlay, the results of the artificial neural networks were superior. In addition, the network trained with the PCA vectors produced better results than the stacked 12-band image.

In similar work, Dai and Khorram (1999) used a four-layer feed-forward back-propagation neural network to detect and classify change in Wilmington, NC between 1988 and 1994. The basic change detection methodology was spectral spatial image classification with a four layer network (i.e., two hidden layers) serving in lieu of a maximum likelihood classifier. The reported accuracy of the change detection (96%, 16 change classes) was higher than the accuracy produced by simple post-classification comparison (87%) of the same two images. However, the high accuracy figures may be due in part to the use of training data as validation data and the small image size (15 km x 15 km). Nonetheless, Dai and Khorram (1999) is an excellent example of how a simple neural network can effectively substitute for a parametric classifier. Finally, see Bruzzone et al. (1999) for a more advanced approach to change detection using feed-forward networks.

Decision trees are another promising method for urban change detection. Whereas ANNs were spawned by research into artificial intelligence, decision trees are used to represent rules in expert systems – systems nominally designed to model human expert knowledge and decision making strategies. Like ANNs, decision trees are useful for classification. In their general incarnation, decision trees have nodes that represent distinguishing attributes. These tree nodes are decision points. Nodes in different levels of the tree are connected by arcs. Arcs exiting nodes represent different decisions made by examining the nodes' attribute state. To classify a pattern, a path is followed through the tree from root to leaf examining nodal attributes and making decisions along the way. A decision at a given node selects an arc and moves the decision to a leaf (representing a final class label) or to another node (requiring another conditional examination).

The objective of Chan et al. (2001) was to test the relative effectiveness of decision trees with three other machine learning algorithms (i.e., multilayer ANN, maximum likelihood, learning vector quantizer) to map urban change on Tsing-yi Island, Hong Kong. Data for the comparison included two triband SPOT images acquired on January 14, 1987 and Feb 5, 1995. Twelve change classes and four no-change classes were constructed from the following landtype categories: 1) built-up land, 2) construction sites, 3) vegetation, and 4) water. Supervised training was used on the stacked multitemporal image for all four learning algorithms.

Accuracy assessment used a validation sample of 1200 sites. For the binary change mapping, accuracy was 88%, 77%, 83%, and 91% for the decision tree, maximum likelihood, ANN, and learning vector approaches re-

spectively. Validation of a traditional post-classification comparison with the same 1200 points produced an accuracy of 69%. When specific change classes were examined, accuracy was much lower. The same four classifiers produced a classification accuracy of 85%, 71%, 73%, and 88%. Accuracy from the post-classification comparison was 81%. Classifier performance differed by change class. The best classifiers for each change class follow:

- Construction site to vegetation: All were equivalent; accuracy = 100%
- Vegetation to construction site: ANN; accuracy = 88%
- Vegetation to built-up land: Maximum likelihood classifier; accuracy = 82%
- Construction site to built-up land: Decision tree classifier; accuracy = 78%
- Built-up land to water: Learning vector quantizer; accuracy = 95%
- Water to built-up land: Maximum likelihood classifier; accuracy = 90%
- Water to construction site: ANN; accuracy = 100%
- Built-up land to vegetation: Learning vector quantizer; accuracy = 58%
- Built-up land to construction site: Decision tree classifier; accuracy = 59%

The **Intensity-Hue-Saturation (IHS) transform** has long been used for image fusion, and yet represents another innovative method of detecting urban change between two satellite images.[5] As described by Chen and Chen et al. (2003), IHS change detection is based on an observation from image fusion i.e., when a forward and inverse IHS transform are performed for fusion purposes, a noticeable color distortion will result if one of the images exhibits any spectral or textural differences related to temporal change.

In order to test the utility of an IHS change detection methodology, Chen and Chen et al. (2003) conducted case studies in Guangzhou and Nanyang cities, PRC. The study required that two disparate image types be compared (i.e., 1996 TM and 1998 SPAN). The processing was straightforward. After initial rectification, the SPAN image was contrast stretched exponentially to enhance urban-vegetation differences and high-pass filtered for edge enhancement. The TM bands were histogram equalized. Successive combinations of three TM bands were then transformed into IHS components. For each IHS TM product, the intensity component was replaced by the SPAN image and transformed inversely from IHS to RGB space.

[5] For more information on the IHS transform, see Harris and Murray (1990). For an introduction to image fusion in remote sensing, see Pohl and Van Genderen (1998).

Cluster analysis was performed on each RGB image, change classes were identified, and the best TM band subset retained for further analysis. For Guangzhou, the three best RGB bands used in the IHS fusion were TM7, TM4 and TM3. Using the classification from this fused image, error assessment was performed on two change classes; arable land which converted to urban land, and all other landtypes. The overall classification accuracy was 93% (kappa = 0.88) for 1000 validation sample sites. Transforms using band sets {TM4, TM3, TM2} and {TM5, TM4, TM3} for the R, G, and B color space components also produced excellent results. Post-classification comparison using SPOT HRV multiband imagery acquired on similar dates produced an accuracy of only 86% (kappa = 0.72) on the same 1000 Guangzhou sites. Results for Nanyang city were comparable.

As an additional innovation, the objective of Kaufmann and Seto (2001) was not just to detect urban change between dates, but to estimate the date the changes occurred using **econometric panel techniques**. The objective was driven by the need to match yearly socioeconomic data with its concurrent annual landcover change. Although the obvious approach would be to acquire satellite imagery yearly, the researchers feared that repetitive pairwise comparison of images would beget unacceptable error rates across the experiment. Thus, to prevent the compounding of errors associated with post-classification comparison, it was important to use a longitudinal comparison scheme that avoided repetitive pairwise image procedures.

The Pearl River Delta, PRC was chosen as a study area. Nine images dated 1988 to 1996 were utilized. The landtype classification scheme had seven categories. In order to assign change dates to image pixels displaying change between 1988 and 1996 (and avoid the pairwise comparisons), Kaufmann and Seto used econometric panel techniques. The process required three steps: “In the first step, regression equations [were] estimated for each of the six DN bands for each of the seven stable land-cover classes. In the second step, the estimated regression equations for each class [were] used to calculate DN values for change land-cover classes for each of the eight possible dates of chan. In the third step, the date of change [was] identified by comparing a pixel’s DN values against the eight possible dates of change using tests for predictive accuracy” (p. 97). On an absolute scale of accuracy, the results were mixed, however the results of the econometric panel technique were superior to a bi-image assessment of change predicated on a multidate tasseled-cap approach.

8.3 Quantification and modeling of urban growth

8.3.1 Urban growth measurement with landscape metrics

Landscape ecology metrics are used to numerically describe landscape structure. Accurate quantification is a prerequisite to understanding how structures affect system interactions in a heterogeneous landscape. Measurement permits landscapes to be compared. It also facilitates the recognition and monitoring of landscape change (Turner 1989; O'Neill et al. 1999). Finally, quantification of landscape structure enables the scientific transition from an inductive to deductive logic model wherein hypotheses can be formed and rigorously tested (Curran 1987; Dietzel et al. 2005). Hobbs (1999) noted that the shift in environmental ecology from a qualitative to quantitative basis was fueled by metrics developed to quantify natural landscape structure and pattern. Many of the metrics and methods used in landscape ecology are borrowed from other sources such as information and complexity theory, fractal geometry, spatial statistics, and image analysis (Turner l989). Most common metrics can be subdivided into two classes:

1. Measurement of individual patch characteristics (e.g., size, shape, perimeter, perimeter-area ratio, fractal dimension).
2. Measurement of whole landscape characteristics (e.g., richness, evenness, dispersion, contagion). Metrics of landscape characteristics are typically more computationally and analytically complex than individual patch metrics (Farina 1998).

Natural landscape metrics have found an important application in quantifying urban growth, sprawl, and fragmentation. Like natural ecology, urban change detection focus has shifted from detection to quantification of change, measurement of pattern, and casual analysis. Herold et al. (2002) represents an early landmark in this shift by establishing that low density residential, high density residential, and commercial zones could be discriminated by landscape metrics such as fractal dimension, landscape percentage, patch density, edge density, patch size standard deviation, contagion index, and area weighted average patch fractal dimension. This discrimination ability led to the creation of landscape metric signatures for Santa Barbara, California landtypes. The metrics were also capable of quantifying land conversion at two California test sites.

Schneider et al. (2005) also illustrate the recent shift in focus from urban change detection to change quantification. To begin this study of urban land conversion in Chengdu, Western PRC, the researchers initially created landtype change maps using satellite imagery. Using the digital change maps, fragmentation statistics such as mean patch size and landscape shape index were calculated. The statistics were not calculated for the entire image, but rather along seven transportation corridors radiating from the Chengdu CBD. Five periods of time, corresponding to five of the image dates, were examined (i.e., 1988, 1991, 1995, and 2000). The statistics highlighted the effect of economical, social, and government policy forcing mechanisms on urban structure.

As illustrated by Schneider et al. (2005) and articulated by Seto and Fragkias (2005), the quantification of urban growth with landscape metrics represents a significant enhancement to the calculation of yearly landtype acreage changes. Although area figures enable change rate calculations, landscape metrics make possible the evaluation of changing urban spatial pattern; an important additional piece of information for planners seeking to control urban growth.

Seto and Fragkias (2005) used landscape metrics to quantify change in four cities of southern PRC over an 11 year period (1988 to 1999). Using satellite imagery, maps of change for several years were constructed. Urban growth rates for the four cities were then annualized. Using the annualized change images, six spatial landscape pattern metrics were calculated for three concentric buffer zones centered on each of the four cities. The metrics selected were intended to describe urban form complexity and size and included total urban area, edge density, urban patch count, mean patch fractal dimension, average patch size, and patch size coefficient of variation. Calculation of urban change rates during the ten year period was also done. As detailed by Seto and Fragkias, key aspects of urban development in the two cities were illuminated by the metrics. Envelopment and multiple nuclei growth were revealed as the primary urban expansion processes. Changing administrative practices to control (or not) land use development were likewise reflected in the metrics. These results illustrate that investigating temporal urban change via landscape metrics is a valuable procedure both to quantify change and link its spatial pattern to cultural practices and processes. See Yu and Ng (2006) for a similar study.

Herold et al. (2005) coined the term “spatial metric growth signatures” to describe this use of landscape metrics in a Santa Barbara urban growth

project. Metrics were not only used to describe historical structure and predicted future urban form, but were also competent to measure goodness of fit between an urban growth model and historical data. The authors also illustrated the use of spatial metrics as a means of visualizing urban modeling results.

8.3.2 Urban growth modeling with cellular automata

The measurement of urban change is an obvious extension of urban change detection. Contemporary research also includes the use of cellular automata (CA) to model urban growth and predict future urban expansion. Remotely sensed maps of urban change are data sources that (usually) calibrate and (rarely) validate the model.

As described by Torrens (2000), cellular automata are an outgrowth of efforts by John von Neumann and Stanislaw Ulam to mathematically model the self-reproduction that characterizes living things. John Conway's CA Game of Life is familiar to most introductory computer science students as a laboratory programming assignment to understand two-dimensional arrays. Although the game is simple, fascinating repetitive patterns and regularities appear. The Game of Life illustrates one of the characteristics of many CA models – large scale self-organization and patterning as a result of actions by a multitude of individual actors. Because they are adaptive, reflect emergence, and can model complexity, CA form an excellent basis for modeling urban growth.

Perhaps the most popular North American CA urban growth model was developed by Keith Clarke at UCSB (Clarke et al. 1997). Named SLEUTH, Clarke's modeling system is freely available to other researchers. This is important because model sharing is essential to model improvement. While SLEUTH can be used to model historical growth (Clarke and Gaydos, 1998) SLEUTH flexibility enables prediction of future urban extent and pattern predicated on "what if" management scenarios (Clarke et al. 1997). Candau and Goldstein (2002) exercised this potential by investigating five growth options for Santa Barbara, California designed to model change over three decades. These included scenarios of 1) historical growth extended to the future, 2) historical patterns supplemented by an expanded road network, 3) moderate environmental protection of sensitive lands, 4) strong environmental protection, and 5) constrained urban expansion. Spatial metrics were effectively used by Candau

and Goldstein to quantify and describe the patterns produced under the various management plots.

When viewed together, Yang (2002) and Yang and Lo (2002, 2003) form an important show-and-tell of a complete project from historical urban change mapping through modeling of future urban growth. In Yang and Lo (2003), the researchers detail their use of SLEUTH to model the urban growth of Atlanta, Georgia. Three future visions of urban growth were entertained. The first was a status quo growth scenario; the unconstrained progression of current patterns into the future. The second scenario modified the first to allow environmental protection of sensitive areas. The third scenario defined active management to constrain Atlanta's urban form and slow its sprawling growth. Results of the SLEUTH modeling were mixed but encouraging. Yang and Lo thought that the model needed improvement, particularly to account for externally imposed constraints on growth such as zoning and regional development policies. An inability to effectively model human behavior and race were other cited limitations.

Other examples of SLEUTH CA urban modeling under different presumed management conditions can be found.

- Syphard et al. (2005) modeled urban expansion in the Santa Monica Mountains National Recreation Area under fictitious slope development constraints of 25%, 30% and 60%.
- Jantz et al. (2003) modeled the impact of management policies on Chesapeake Bay water quality. The model was based on Washington D.C. regional growth regulated by three possible scenarios: 1) cloistered resource growth, 2) spatially focused protected resource growth, and 3) status quo growth.
- Xian and Crane (2005) adapted SLEUTH to predict suburban expansion in the Tampa Bay, FL watershed for the year 2025. Using impervious surface as a surrogate for urban land, Xian and Crane detected and mapped urban growth through satellite image analysis, calibrated the CA model with the image analysis results, and then predicted future growth with the calibrated model.

8.4 Future research directions

From the foregoing review, future research directions are suggested. Certainly efforts to enhance urban modeling via cellular automata should con-

tinue unabated. We also believe that the study of landscape metrics should continue with discussions regarding their correct urban application and interpretation. The five research needs we suggest below require investigation at a more fundamental level. We admit that some of them require reinvestigation of old problems with new eyes and rekindled enthusiasm to make significant contributions.

The bulk of this review was devoted to urban change mapping. This was intentional because the quality of urban change mensuration and modeling depends on accurate data. Therefore a significant and thorny research challenge is to improve the accuracy of landtype maps produced through satellite image analysis. Although other agendas are sexier, this is a critical and persistent need for advancing urban change studies. The problem of inaccurate landtype maps is particularly acute for practitioners mapping change via post-classification comparison wherein errors in each landtype map are compounded in the overlay process (Stow et al. 1980). In their landmark work, Anderson et al. (1976) state that "the minimum level of interpretation accuracy in the identification of land use and land cover categories from remote sensor data should be at least 85 percent. The accuracy of interpretation for the several categories should [also] be about equal" (p. 5). This standard is insufficient for maps used in post-classification comparison. Application of these guidelines for a *bi-temporal landtype change map* requires the *original landtype maps* to have a producer's accuracy of at least 93%, with errors equally distributed between all classes.

The second research need is related to the first. There is a need for rigorous accuracy assessment of landtype change products. In a large minority of projects we reviewed, no such assessment was reported. In others, accuracy assessment was done using the same data that fed the classification training phase. In addition, we suggest that sensitivity analysis be performed and reported. Accuracy claims can then be accompanied by an assessment of data uncertainty's effect. This is an old idea (Fisher, 1999) that should be considered standard procedure in urban change detection, quantification, and modeling. Such sensitivity analysis may provide sobering insights into the limits of our mapping.

Accuracy assessment, error estimation, and sensitivity analyses can be computationally intensive, particularly in CA modeling. The third research need is for the development of numerical methods, computing data structures, and computer algorithms to reduce computational burdens imposed by urban modeling and analysis.

The fourth research need is to continue development of novel change detection methodologies that mitigate the accuracy problem inherent in post-classification comparison. Since a single "best method" does not exist, more exploration is warranted. We believe that continued efforts to explain urban change phenomena in context of physical models (e.g., VIS) and urban fabric reflectance (e.g., SMA) will be important components of these novel methods. The use of econometric panel methods (Kaufmann and Seto, 2001) should also be evaluated more completely by applying them to a variety of urban settings. The role and use of high-resolution hyperspectral imagery for urban change detection should also be evaluated. As mentioned previously, making novel methods accessible to more practitioners is also necessary for the sake of replication and general assessment.

We conclude with the fifth research need — the need to focus on urban humanity's condition. As the metropolis grows into the megalopolis and becomes part of a worldwide gigalopolis, monitoring urban growth using geographic technologies can preserve and improve the quality of life for the family Hominidae. This means a reinvigorated focus on the needs of developing countries. Consider the newly industrializing nation where a burgeoning urban population outpaces the government's ability to provide essential health and service infrastructure essential to forestalling social meltdown. Maps provided by satellite remote sensing can provide fundamental ecological, population, housing, and life quality data (albeit imperfect) necessary for rational and equitable urban management. Further modeling of the urban growth can empower decision makers to create informed policy that directs growth in a sustainable direction. These would be policies that anticipate critical need before crises arise. Although the tasks of urban growth detection, measurement, and modeling are satisfying academic exercises, the exercises must lead to judicious management of our urban world.

References

Alberti, M., Weeks, R. and Coe, S. (2004) Urban land-cover change analysis in central Puget Sound, *Photogrammetric Engineering and Remote Sensing* 70:1043-52.

Alexander, D. and Tomalty, R. (2002) Smart growth and sustainable development: Challenges, solutions and policy directions, *Local Environment* 7:397-409.

Anderson, J. R., Hardy, E. E., Roach, J. T. and Witmer, R. E. (1976) *A Land Use and Land Cover Classification System for Use with Remote Sensor Data*, USGS Geological Survey Professional Paper 964. Washington, DC: USGS

Anthony, J. (2004) Do state growth management regulations reduce sprawl? , *Urban Affairs Review* 39:367-97.

Aspinall, R.J. and Hill, M.J. (1997) Land cover change: a method for assessing the reliability of land cover changes measured from remotely sensed data, *Proceedings of the 1997 International Geoscience and Remote Sensing Symposium*, Singapore, August 4-8, 1997.

Audirac, I., Shermyen, A. H. and Smith, M. T. (1990) Ideal urban form and visions of the good life: Florida's growth management dilemma, *Journal of the American Planning Association* 56:470-82.

Banner, A. V. and Lynham, T. (1981) Multitemporal analysis of Landsat data for forest cut over mapping – a trial of two procedures., *Proceedings of the 7th Canadian Symposium on Remote Sensing , Winnipeg, Manitoba, Canada*:233-40.

Benediktsson, J. A., Swain, P. H. and Ersoy, O. K. (1990) Neural network approaches versus statistical methods in classification of multisource remote sensing data *IEEE Transactions on Geoscience and Remote Sensing* 28:540-51.

Bengston, D. N., Fletcher, J. O. and Nelson, K. C. (2004) Public policies for managing urban growth and protecting open space: policy instruments and lessons learned in the United States, *Landscape and Urban Planning* 69:271-86.

Ben-Zadok, E. (2005) Consistency, concurrency and compact development: three faces of growth management implementation in Florida, *Urban Studies* 42:2167-90.

Bernstein, J. D. (1994) *Land Use Considerations in Urban Environmental Management*. Washington, D.C.: Urban Development Division, World Bank.

Bischof, H., Shneider, W. and Pinz, A. J. (1992) Multispectral classification of Landsat-images using neural networks, *IEEE Transactions on Geoscience and Remote Sensing* 30:482-90.

Bruzzone, L. and Prieto, D. F. (2000) Automatic analysis of the difference image for unsupervised change detection, *IEEE Transactions on Geoscience and Remote Sensing* 38:1171-82.

Bruzzone, L., Prieto, D. F. and Serpico, S. B. (1999) A neural-statistical approach to multitemporal and multisource remote-sensing image classification, *IEEE Transactions on Geoscience and Remote Sensing* 37:1350-9.

Buliung, R. N. and Kanaroglou, P. S. (2006) Urban form and household activity travel behavior, *Growth and Change* 37:172-8.

Burby, R. J., Nelson, A. C., Parker, D. and Handmer, J. (2001) Urban containment policy and exposure to natural hazards: is there a connection? *Journal of Environmental Planning and Management* 44:475-90.

Burchell, R. W., Downs, A., McCann, B. and Mukherji, S. (2005) *Sprawl Costs: Economic Impacts of Unchecked Development*. Washington, DC: Island Press.

Byrne, G. F., Crapper, P. F. and Mayo, K. K. (1980) Monitoring land-cover by principal components analysis of multitemporal Landsat data, *Remote Sensing of Environment* 10:175-84.

Cakir, H. I., Khorram, S. and Nelson, S. A. C. (2006) Correspondence analysis for detecting land cover change, *Remote Sensing of Environment* 102:306-17.

Candau, J. C. and Goldstein, N. (2002) Multiple scenario urban forecasting for the California south coast region, *Proceedings from the 40th Annual Conference of the Urban and Regional Information Systems Association*, Chicago, IL, October 26-30.

Carruthers, J. I. (2002) The impacts of state growth management programmes: A comparative analysis, *Urban Studies* 39:1959-82.

Carruthers, J. I. and Ulfarsson, G. F. (2003) Urban sprawl and the cost of public services, *Environment and Planning B: Planning and Design* 30:503-22.

Cervero, R. (2001) Efficient urbanization: economic performance and the shape of the metropolis, *Urban Studies* 28:1651-71.

Chan, J. C. W., Chan, K. P. and Yeh, A. G. O. (2001) Detecting the nature of change in an urban environment: a comparison of machine learning algorithms, *Photogrammetric Engineering and Remote Sensing* 67:213-25.

Chen, J., Gong, P., He, C., Pu, R. and Shi, P. (2003) Land-use / land-cover change detection using improved change vector analysis, *Photogrammetric Engineering and Remote Sensing* 69:369-79.

Chen, Z., Chen, J., Shi, P. and Tamura, M. (2003) An IHS-based change detection approach for assessment of urban expansion impact on arable land loss in China, *International Journal of Remote Sensing* 24:1353-60.

Clarke, K. C. and Gaydos, L. J. (1998) Loose-coupling a cellular automaton model and GIS: long-term urban growth prediction for San Francisco and Washington / Baltimore, *International Journal of Geographical Information Science* 12:699-714.

Clarke, K. C., Hoppen, S. and Gaydos, J. (1997) A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area, *Environment and Planning B: Planning and Design* 24:247-61.

Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B. and Lambin, E. (2004) Digital change detection methods in ecosystem monitoring: a review, *International Journal of Remote Sensing* 25:1565-96.

Coppin, P. R. and Bauer, M. E. (1994) Processing of multitemporal Landsat TM imagery to optimise extraction of forest cover change features, *IEEE Transactions on Geoscience and Remote Sensing* 32:918-27.

Coppin, P. R. and Bauer, M. E. (1994) Processing of multitemporal Landsat TM imagery to optimize extraction of forest cover change features, *IEEE Transactions on Geoscience and Remote Sensing* 32:918-27.

Curran, P. J. (1987) Remote sensing methodologies and geography, *International Journal of Remote Sensing* 8:1255-75.

Dai, X. and Khorram, S. (1999) Remotely sensed change detection based on artificial neural networks, *Photogrammetric Engineering and Remote Sensing* 65:1187-94.

Davis, C. and Schaub, T. (2005) A transboundary study of urban sprawl in the Pacific Coast region of North America: the benefits of multiple measurement methods, *International Journal of Applied Earth Observation and Geoinformation* 7:268-83.

Dawkins, C. J. and Nelson, A. C. (2002) Urban containment policies and housing prices: an international comparison with implications for future research, *Land Use Policy* 19:1-12.

Deal, B. and Schunk, D. (2004) Spatial dynamic modeling and urban land use transformation: A simulation approach to assessing the costs of urban sprawl, *Ecological Economics* 51:79-95.

Dietzel, C., Herold, M., Hemphill, J. J. and Clarke, K. C. (2005) Spatio-temporal dynamics in California's Central Valley: Empirical links to urban theory, *International Journal of Geographic Information Science* 19:175-95.

Downs, A. (2001) What does 'smart growth' really mean? *Planning* 67:20-5.

Du, Y., Guindon, B. and Cihlar, J. (2002) Haze detection and removal in high resolution satellite images with wavelet analysis, *IEEE Transactions on Geoscience and Remote Sensing* 40:210-7.

Duda, R. O., Hart, P. E. and Stork, D. G. (2001) *Pattern Classification*. New York: Wiley.

Dwyer, J. F. and Childs, G. M. (2004) Movement of people across the landscape: A blurring of distinctions between areas, interests and issue affecting natural resource management, *Landscape and Urban Planning* 69:153-64.

Eastman, J. R. and Fulk, M. (1993) Long sequence time series evaluation using standardized principal components, *Photogrammetric Engineering and Remote Sensing* 59:991-6.

Eklundh, L. and Singh, A. (1993) A comparative analysis of standardised and unstandardised principle components analysis in remote sensing, *International Journal of Remote Sensing* 14:1358-70.

Farina, A. (1998) *Principles and Methods in Landscape Ecology*. London: Chapman & Hall.

Fisher, P. (1999) Models of uncertainty in spatial data, in Longley, P. A., Goodchild, M. F., Maguire, D. J. and Rhind, D. W. (eds.) *Geographical Information Systems: Principles, Techniques, Management, and Applications*. New York, NY: John Wiley, pp. 191-205.

Frenkel, A. (2004) The potential effect of national growth-management policy on urban sprawl and the depletion of open spaces and farmland, *Land Use Policy* 21:357-69.

Fung, T. (1990) An assessment of TM imagery for land-cover change detection, *IEEE Transactions on Geoscience and Remote Sensing* 28:681-4.

Fung, T. (1992) Land use and land cover change detection with Landsat MSS and SPOT HRV data in Hong Kong, *Geocarto International* 3:33-40.

Fung, T. and LeDrew, E. (1987) Application of principal components analysis to change detection, *Photogrammetric Engineering and Remote Sensing* 53:1649-58.

Fung, T. and LeDrew, E. (1988) The determination of optimal threshold levels for change detection using various accuracy indices, *Photogrammetric Engineering and Remote Sensing* 54:1449-54.

Hardin, P. J. (2000) Neural networks versus nonparametric neighbor-based classifiers for semisupervised classification of Landsat Thematic Mapper imagery, *Optical Engineering* 39:1898-908.

Harris, J. R. and Murray, R. (1990) IHS transform for the integration of radar imagery with other remotely sensed data, *Photogrammetric Engineering and Remote Sensing* 56:1631-41.

Hayden, D. (2003) *Building Suburbia: Greenfields and Urban Growth 1820-2000*. New York: Pantheon Books.

Haykin, S. (1994) *Neural Networks: a Comprehensive Foundation*. Upper Saddle River, NJ: Prentice-Hall.

Herold, M., Couclelis, H. and Clarke, K. C. (2005) The role of spatial metrics in the analysis and modeling of urban change, *Computers, Environment, and Urban Systems* 29:339-69.

Herold, M., Scepan, L. and Clarke, K. C. (2002) The use of remote sensing and landscape metrics to describe structures and change in urban land uses, *Environment and Planning A* 34:1443-58.

Hobbs, R. J. (1999) Clark Kent or Superman: Where is the phone booth for landscape ecology? in Klopatek, J. M. and Gardner, R. H. (eds.) *Landscape Ecological Analysis: Issues and Applications*. New York, NY: Springer.

Howarth, P. J. and Boasson, E. (1983) Landsat digital enhancements for change detection in urban environments, *Remote Sensing of Environment* 13:149-60.

Howell-Moroney, M. (2004) Community characteristics, open space preservation and regionalism: Is there a connection? *Journal of Urban Affairs* 26:109-18.

Im, J. and Jensen, J.R. (2005) A change detection model based on neighborhood correlation image analysis and decision tree classification, *Remote Sensing of Environment*, 99:326-340.

Ingram, K., Knapp, E. and Robinson, J. (1981) Change detection technique development for improved urbanized area delineation, *Technical memorandum CSC/TM-81/6087 - Computer Sciences Corporation, Silver Springs, Maryland*.

Irwin, E. G. and Bockstael, N. E. (2004) Land use externalities, open space preservation, and urban sprawl, *Regional Science and Urban Economics* 34:705-25.

Jantz, C. A., Goetz, S. J. and Shelley, M. K. (2003) Using the SLEUTH urban growth model to simulate the impacts of future policy scenarios on urban land use in the Baltimore-Washington metropolitan area, *Environment and Planning B: Planning and Design* 30:251-71.

Jensen, J. R. (2005) *Introductory Digital Image Processing: A Remote Sensing Perspective*. Upper Saddle River, NJ: Prentice-Hall.

Jensen, J. R., Qui, F. and Ji, M. (2000) Predictive modeling of coniferous forest age using statistical and artificial neural network approaches applied to remote sensor data, *International Journal of Remote Sensing* 20:2805-22.

Jensen, J. R., Rutchey, K., Koch, M. S. and Narumalani, S. (1995) Inland wetland change detection in the Everglades water conservation area 2A using a time series of normalized remotely sensed data, *Photogrammetric Engineering and Remote Sensing* 61:199-209.

Jensen, R. R. and Hardin, P. J. (2005) Estimating urban leaf area using field measurements and satellite remote sensing data, *GIScience and Remote Sensing* 42:229-52.

Johnson, P. E. (2001) Environmental impacts of urban sprawl: a survey of the literature and proposed research agenda, *Environment and Planning A* 33:717-35.

Johnson, R. D. and Kasischke, E. S. (1998) Change vector analysis: a technique for the multispectral monitoring of land cover and condition, *International Journal of Remote Sensing* 19:411-26.

Kashian, R. and Skidmore, M. (2002) Preserving agricultural land via property assessment policy and the willingness to pay for land preservation, *Economic Development Quarterly* 16:75-87.

Kaufmann, R. K. and Seto, K. C. (2001) Change detection, accuracy, and bias in a sequential analysis of Landsat imagery in the Pearl River Delta, China: econometric techniques, *Agriculture, ecosystems, and environment* 85:95-105.

Kawata, Y., Ohtani, A., Kusaka, T. and Ueno, S. (1990) Classification accuracy for the MOS-1 MESSR data before and after the atmospheric correction, *IEEE Transactions on Geoscience and Remote Sensing* 28:755-60.

Kline, J. D. (2000) Comparing states with and without growth management analysis based on indicators with policy implications comment, *Land Use Policy* 17:349-55.

Kline, J. D. (2006) Public demand for preserving local open space, *Society and Natural Resources* 19:645-59.

Kohonen, T. (1995) *Self-Organizing Maps*. Berlin: Springer.

Kuby, M., Barranda, A. and Upchurch, C. (2004) Factors influencing light-rail station boardings in the United States, *Transportation Research Part A: Policy and Practice* 38:223-47.

Lewis, S. (1990) The town that said no to sprawl, *Planning (APA)* 56:14-9.

Lindstrom, M. J. and Bartling, H. (2003) *Suburban Sprawl: Culture, Theory, and Politics*. Lanham, MD: Rowman and Littlefield.

Li, X. and Yeh, A. G. O. (1998) Principal component analysis of stacked multitemporal images for the monitoring of rapid urban expansion in the Pearl River delta, *International Journal of Remote Sensing* 19:1501-18.

Liu, X. and Lathrop, R. G., Jr. (2002) Urban change detection based on an artificial neural network, *International Journal of Remote Sensing* 23:2513-8.

Liu, Y., Nishiyama, S. and Yano, T. (2004) Analysis of four change detection algorithms in bi-temporal space with a case study, *International Journal of Remote Sensing* 25:2121-39.

Madhavan, B. B., Kubo, S., Kurisaki, N. and Sivakumar, T. V. L. N. (2001) Appraising the anatomy and spatial growth of the Bangkok Metropolitan area using a vegetation-impervious-soil model through remote sensing, *International Journal of Remote Sensing* 22:789-806.

Maktav, D. and Erbek, F. S. (2005) Analysis of urban growth using multitemporal satellite data in Istanbul, Turkey, *International Journal of Remote Sensing* 26:797-810.

Masek, J. G., Lindsay, F. E. and Goward, S. N. (2000) Dynamics of urban growth in the Washington DC metropolitan area, 1973-1996, from Landsat observations, *International Journal of Remote Sensing* 21:3473-86.

Mattson, G. A. (2002) *Small Towns, Sprawl and the Politics of Policy Choices: The Florida Experience*. Lanham, MD: University Press of America.

Millward, A. A., Piwowar, J. M. and Howarth, P. J. (2006) Time-series analysis of medium-resolution, multisensor satellite data for identifying landscape change, *Photogrammetric Engineering and Remote Sensing* 72:653-63.

Morisette, J. F. and Khorram, S. (2000) Accuracy assessment curves for satellite based change detection, *Photogrammetric Engineering and Remote Sensing* 66:875-80.

Morris, D. E. (2005) *It's a Sprawl World After All*. Gabriola Island, Canada: New Society Publishers.

Mundia, C. N. and Aniya, M. (2005) Analysis of land use/cover changes and urban expansion in Nairobi city using remote sensing and GIS, *International Journal of Remote Sensing* 26:2831-49.

Nelson, A. C. (1990) Economic critique of US prime farmland preservation policies. Towards state policies that influence productive, consumptive and speculative value components of the farmland market to prevent urban sprawl and foster agricultural production in the United States, *Journal of Rural Studies* 6:119-42.

Nelson, A. C. (1992) Preserving prime farmland in the face of urbanization: lessons from Oregon, *Journal of the American Planning Association* 58:467-88.

Nelson, A. C. (2000) Comparing states with and without growth management analysis based on indicators with policy implications, *Land Use Policy* 16:121-7.

O'Neill, R. V., Ritters, K. H., Wichham, J. D. and Jones, K. B. (1999) Landscape pattern metrics and regional assessment, *Ecosystem Health* 5:225-33.

Paolini, L., Grings, F., Sobrino, J., Jiménez Muñoz, J. and Karszenbaum, H. (2006) Radiometric correction effects in Landsat multi-date/multi-sensor change detection studies, *International Journal of Remote Sensing* 27:685-704.

Pijanowski, B. C., Pithadia, S., Shellito, B. A. and Alexandridis, K. (2005) Calibrating a neural network-based urban change model for two metropolitan areas of the Upper Midwest of the United States, *International Journal of Geographical Information Science* 19:197-215.

Pohl, C. and Van Genderen, J. L. (1998) Multisensor image fusion in remote sensing: concepts, methods, and applications, *International Journal of Remote Sensing* 19:823-54.

Quarmby, N. A. and Cushnie, J. L. (1989) Monitoring urban land cover changes at the urban fringe from SPOT HRV imagery in south-east England, *International Journal of Remote Sensing* 10:953-63.

Raad, T. and Kenworthy, J. (1998) The US and us: Canadian cities are going the way of their US counterparts into car-dependent sprawl, *Alternatives* 24:14-22.

Radeloff, V. C., Holcomb, S. S., McKeefry, J. F., Hammer, R. B. and Stewart, S. I. (2005) The wildland-urban interface in the United States, *Ecological Applications* 15:799-805.

Ridd, M. K. and Liu, J. (1998) A comparison of four algorithms for change detection in an urban environment, *Remote Sensing of Environment* 63:95-100.

Robbins, P. and Birkenholtz, T. (2003) Turfgrass revolution: Measuring the expansion of the American Lawn, *Land Use Policy* 20:181-94.

Rogan, J., Franklin, J. and Roberts, D. A. (2002) A comparison of methods for monitoring multitemporal vegetation change using Thematic Mapper imagery, *Remote Sensing of Environment* 80:143-56.

Rosin, P. L. (2002) Thresholding for change detection, *Computer Vision and Image Understanding* 86:79-95.

Rosin, P. L. and Ioannidis, E. (2003) Evaluation of global image thresholding for change detection, *Pattern Recognition Letters* 24:2345-56.

Rouse J.W., R.H. Haas, J.A. Schell, D.W. Deering, and Harlan, J.C. (1974) *Monitoring the Vernal Advancement and Retrogradation (Greenwave Effect) of Natural Vegetation*. College Station (TX): Texas A&M University, Remote Sensing Center. Report RSC 1978-4.

Schneider, A., Seto, K. C. and Webster, D. R. (2005) Urban growth in Chengdu, Western China: applications of remote sensing to assess planning and policy outcomes, *Environment and Planning B: Planning and Design* 32:323-45.

Schott, J. R., Salvaggio, C. and Volchock, W. J. (1988) Radiometric scene normalization using pseudoinvariant features, *Remote Sensing of Environment* 26:1-16.

Schowengerdt, R. A. (1997) *Remote Sensing; Models and Methods for Image Processing*. New York: Academic Press.

Serra, P., Pons, X. and Sauri, D. (2003) Post-classification change detection with data from different sensors: some accuracy considerations, *International Journal of Remote Sensing* 24:3311-40.

Seto, K. C. and Fragkias, M. (2005) Quantifying spatiotemporal patterns of urban land-use change in four cities of China with timer series landscape metrics, *Landscape Ecology* 20:871-88.

Singh, A. (1989) Digital change detection techniques using remotely-sensed data, *International Journal of Remote Sensing* 10:989-1003.

Singh, A. and Harrison, A. (1985) Standardized principal components, *International Journal of Remote Sensing* 6:883-96.

Smart Growth Network (2003) *Getting to Smart Growth: 100 Policies for Implementation*: International City/County Management Organization (ICMA).

Smits, P. C. and Annoni, A. (2000) Toward specification-driven change detection, *IEEE Transactions on Geoscience and Remote Sensing* 38:1484-8.

Song, C., Woodcock, C. E., Seto, K. C., Pax-Lenney, M. and Macomber, S. A. (2001) Classification and change detection using Landsat TM data: when and

how to correct for atmospheric effects? *Remote Sensing of Environment* 75:230-44.

Stow, D. A., Tinney, L. R. and Estes, J. E. (1980) Deriving land use/land cover change statistics from Landsat: a study of prime agricultural land, *Proceedings of the 14th International Symposium on Remote Sensing of the Environment*. Ann Arbor, MI: Environmental Research Institute of Michigan, pp. 1227-37.

Summers, A. A., Cheshire, P. C. and Senn, L. (1999) *Urban Change in the United States and Western Europe: Comparative Analysis and Policy*. Washington DC: Urban Institute Press.

Sunar, F. (1998) Analysis of changes in a multidate set: a case study in Ikitelli Area, Istanbul, Turkey, *International Journal of Remote Sensing* 19:225-35.

Syphard, A. D., Clarke, K. C. and Franklin, J. (2005) Using a cellular automaton model to forecast the effects of urban growth on habitat pattern in southern California, *Ecological Complexity* 2:185-203.

Theobald, D. M. (2005) Landscape patterns of exurban growth in the USA from 1980 to 2020, *Ecology and Society* 10:34.

Todd, W. J. (1977) Urban and regional land use change detected by using Landsat data, *Journal of Research by the U.S. Geological Survey* 5:529-34.

Tomalty, R. (2002) Growth management in the Vancouver Region, *Local Environment* 7:431-45.

Torrens, P. M. (2000) How cellular models of urban systems work (1. Theory). *Centre for Advanced Spatial Analysis Working Paper Series* 28.

Turner, M. G. (1989) Landscape Ecology: The Effects of Pattern on Process, *Annual Review of Ecological Systems* 20:171-97.

UNCHS (2001) *The State of the World's Cities*. Nairobi, Kenya: United Nations Centre for Human Settlements.

Ward, D., Phinn, S. R. and Murray, A. T. (2000) Monitoring growth in rapidly urbanizing areas using remotely sensed data, *Professional Geographer* 52:371-86.

Weber, C. and Puissant, A. (2003) Urbanization pressure and modeling of urban growth: example of the Tunis Metropolitan Area, *Remote Sensing of Environment* 86:341-52.

Weiler, S. and Theobald, D. (2003) Pioneers of rural sprawl in the Rocky Mountain West, *Review of Regional Studies* 33:264-83.

Willson, R. W. (1995) Suburban parking requirements: a tacit policy for automobile use and sprawl, *Journal of the American Planning Association* 61:29-42.

Xian, G. and Crane, M. (2005) Assessments of urban growth in the Tampa Bay watershed using remote sensing data, *Remote Sensing of Environment* 97:203-15.

Xiao, J., Shen, Y., Ge, J., Tateishi, R., Tang, C., Liang, Y. and Huang, Z. (2006) Evaluating urban expansion and land use change in Shijiazhuang, China, by using GIS and remote sensing, *Landscape and Urban Planning* 75:69-80.

Yang, X. (2002) Satellite monitoring of urban spatial growth in the Atlanta metropolitan area, *Photogrammetric Engineering and Remote Sensing* 68:725-34.

Yang, X. and Lo, C. P. (2002) Using a time series of satellite imagery to detect land use and land cover changes in the Atlanta, Georgia metropolitan area, *International Journal of Remote Sensing* 23.

Yang, X. and Lo, C. P. (2003) Modelling urban growth and landscape changes in the Atlanta metropolitan area, *International Journal of Geographical Information Science* 17:463-88.

Yeh, A. G. O. and Li, X. (1997) An integrated remote sensing and GIS approach in the monitoring and evaluation of rapid urban growth for sustainable development in the Pearl River Delta, China, *International Planning Studies* 2:193-210.

Yu, X. and Ng, C. (2006) An integrated evaluation of landscape change using remote sensing and landscape metrics: a case study of Panyu, Guangzhou, *International Journal of Remote Sensing* 27:1075-92.

Yuan, D., Elvidge, C. D. and Lunetta, R. S. (eds.) (1999) *Survey of multispectral methods for land cover change analysis.* London: Taylor and Francis.

9 Deer-Vehicle Collisions Along the Suburban-Urban Fringe

Rusty A. Gonser, Department of Life Sciences, Indiana State University, Terre Haute, Indiana.

J. Scott Horn, Utah Geological Survey, Salt Lake City, Utah

9.1 Introduction

Deer-Vehicle Collisions (DVCs) are a significant problem in many areas of the United States (Conover et al. 1995). In 2002 alone, there were over 1.5 million DVCs resulting in over 1 billion dollars in damages, 150 human fatalities, and approximately 1.5 million white-tailed deer deaths (Curtis and Hedlund 2005). In sum, there are roughly 4,100 accidents/day resulting in over 2.7 million dollars of damage/day. DVCs are an increasing hazard to motorists as human development spreads into rural areas where residents commute daily to urban locations. Furthermore, through urban sprawl cities further encroach into environments where wildlife have no choice but to interact with humans. In many cases these interactions can have negative impacts for both humans and wildlife.

The main factors predicting DVCs (e.g. number of deer, number of vehicles, and miles traveled) are increasing (Insurance Institute for Highway Safety 1993). In order to reduce the number of DVCs, many countermeasures have been developed, most of which have focused on the prevention of DVCs through mitigation and driver awareness. However, few countermeasures have been empirically validated and, as a result, the effectiveness of these efforts remains unknown (Deercrash 2006). An emerging trend in DVC mitigation is to use spatial (e.g. location of DVC) and temporal (e.g. time and date of DVC) data to determine where and when to use countermeasures (DeerCrash 2006; Hubbard et al. 2000; Roseberry and Woolf 1998). Historically, DVC research has focused on factors such as

traffic volume and traffic speed which are thought to have major impacts on DVCs (Bashore et al. 1985; Allen and McCullough 1976; Pojar et al. 1975), however, they may not be as good of predictors of DVCs as once thought. Data now suggest that landscape structure, rather than traffic volume, might be a better indicator of DVC incidence (Horn 2005; Hubbard et al. 2000).

This chapter focuses on human-environment interactions, especially those between motorists and white-tailed deer (i.e. human-wildlife interactions), and the use of Geographic Information Systems technology to predict the locations of DVCs to better facilitate the use of mitigation strategies. Additionally, since white-tailed deer are social organisms, it would be logical to use GIS to analyze the autocorrelation between incidences of DVC, as well as their relationship to the types of habitat that they occur in.

9.1.1 Background

In this country when we discuss wildlife, the usual mindset is one based on the conservation of species and/or habitats. Conservation is a relatively easy thing to manage if there are vast tracts of uninhabited land that we can set aside for that purpose. However, in today's world we are faced with the reality that those uninhabited areas are vanishing. In fact, the three principles of conservation biology, 1) that evolution is the thread that unites all disciplines of biology, 2) that the ecological world is dynamic, and 3) that the presence of humans must be a part of all conservation efforts (Groom et al. 2006), remind us that humans are an integral part of today's ecological landscape. In modern times, principle 3 has become extremely important as human intervention can change spatial and temporal dynamics by altering landscapes, species interactions, and the energy flow in environments (principles 1 and 2; Meffe et al. 2006). Many anthropogenic changes occur so quickly that organisms are incapable of adapting to these changes before facing local or global extinction. Given that the presence of humans can vastly change the health of an ecosystem and the very course of evolution itself, it is essential that we evaluate the strength of such anthropogenic effects.

The natural world is a far different place now than it was 1,000 years ago – this is primarily due to anthropogenic effects as humans have altered almost every natural ecosystem on the planet (Meffe et al. 2006). As human populations grew, there was an increased demand for natural resources that generated a conflict between the exploitation and the preservation of habi-

tats (Rothely 2001). European colonists exploited North American forests for their lumber (used for ship masts, charcoal, gum, and turpentine) and furs (from beaver, white-tailed deer, and mink.; Meffe et al. 2006; Harlow and Guynn 1994). Early settlers viewed the North American forest as vast and endless, a resource that could never be exhausted, and as a result, exploitation became the norm. Many species of wildlife were adversely affected by the cultural environment of the times. Some species became extinct while others, such as the white-tailed deer (*Odocoileus virginianus*), reached critically low numbers. Given the number of DVCs we experience today, this is hard to imagine.

White-tailed deer (*Odocoileus virginianus*) were nearly extirpated from much of their North American ranges by 1900 (Woolf and Roseberry 1998). Two episodes of over-harvesting contributed to this severe decline in numbers. The first episode occurred during the early years of colonial times, and was primarily due to over-harvesting for the fur trade. It is estimated that during this time, Native Americans and colonial settlers killed between 4.6 and 6.4 million deer annually to support the demand for furs (Harlow and Guynn 1994). The second episode of over-harvesting occurred during the last half of the 19th century and was fueled by subsistence hunting and the demands of ever-increasing markets. During this period, which had the highest recorded hunting pressure in North American history (Harlow and Guynn 1994), white-tailed deer numbers decreased from an estimated 18 million to roughly 500,000 individuals nationwide (Curtis and Hedlund 2005; Hubbard et. al. 2000).

The recovery of many deer populations across the United States is a result of a combination of management policies and several key characteristics of their species, including adaptability, mobility, reproductive vigor, and a lack of natural predators. Improved conservation efforts and planned management (e.g. hunting seasons and limits, habitat protection, law enforcement, etc.) helped deer populations recover (Hubbard et al. 2000). In addition, two other anthropogenic factors affected deer population recovery. The first was the migration of rural human populations to the cities, thereby leading to the abandonment of approximately 65% of agricultural lands (Harlow and Guynn 1994). This resulted in a reduction in hunting (humans were the only predation threat to deer after the eradication of natural predators) and a change in landscape structure, both of which made more land available for deer.

The story of white-tailed deer population recovery is the opposite of what one might think regarding human's disturbance of habitat. Because white-

tailed deer require open grassland with nearby woodlands for protection, they thrive on human disturbance of forested habitat. When forest is the dominant habitat type it supports fewer deer because the canopy prevents the growth of understory browse, which is preferred by deer (Rothley 2001). Moderate amounts of human disturbance leave fragmented sections of forest and actually enhance deer abundance by increasing the amount of forage and reducing the amount of land available to hunters (Roseberry and Woolf 1998). As forests, which had been cleared for timber and agriculture, began to regenerate they became ideal deer habitat (Hubbard et al. 2000) because these habitats became dominated by favored browse plants including shrubs, vines, woody seedlings, legumes, and grasses (Masters et. al. 1996; Jacobson, 1994). Together, all of these factors contributed to one of the most successful population recoveries ever documented. By the end of the twentieth century the deer numbers had risen from 500,000 in 1900 to an estimated 30 million deer nationwide (Curtis and Hedlund 2005; Woolf and Roseberry 1998).

It is difficult to determine actual deer herd sizes because estimates typically come from annual harvest numbers (Strickland et. al. 1996), which are a function of several factors not associated with increasing population size (e.g. increased number of hunters). However, despite these uncertainties, it is clear that the last three decades of the 20th century have produced a large rate of increase in US deer populations (Curtis and Hedlund 2005). The white-tailed deer success story has put wildlife management agencies at odds with other competing stakeholders (Sullivan and Messmer 2003). As deer populations increase, so has the number of deer-human interactions resulting in an increase in crop damage, disease (i.e. Lyme Disease), and DVCs (Kilpatrick et. al. 2002). In addition, over the last 100 years, Departments of Transportation (DOTs) have developed and improved transportation corridors throughout North America (Sullivan and Messmer 2003). The construction of roads may be the single biggest threat to wildlife health and diversity as they negatively affect wildlife by:

- Increasing mortality and injury
- Compacting soils
- Affecting bodies of water due to oil and grime run-off
- Disrupting the physical and chemical environments of soil and water
- Acting as a corridor for the spread of exotic species and disease
- Fragmenting habitat and increasing edge effects
- Disrupting social structure

- Disruption of dispersal and migtration process that maintain genetic flow with species populations
- Increasing human access (Adams et al. 2006; Donaldson and Bennett 2004; Trombulak and Frissell 2000; Forman and Alexander 1998)

White-tailed deer thrive on edge habitat (Harlow and Guynn 1994). Therefore anthropogenic activities that fragment the landscape ultimately create ideal edge habitat for deer. Roads often create edge habitat; therefore roads bring humans and deer together in the form of Deer-Vehicle Collisions (Conover et al. 1995). During the 1980's and 1990's several states observed a 51-69% increase in DVCs (Curtis and Hedlund 2005). There are several reviews of Mitigation and DVC prevention. Some variables involved in predicting DVCs include: habitat, traffic volume and speed, visibility, deer population size, number of bridges, number of buildings (Nielsen et al. 2003; Hubbard et al. 2000; Bashore et al. 1985; Case 1978; Pojar et al. 1975). Most DVC studies focus on reduction and not an in depth look at landscape structure and how it might influence DVCs.

The state of Indiana is an excellent venue for investigating this relationship. The number of DVCs in 1981 for Indiana was approximately 2,000. However, by 2004 that number had grown to almost 15,000 (McNew 2005a). Currently Indiana is in the top 10 of number of DVCs in the nation (HuntingNet 2006). Indiana has many characteristics (e.g. a mixture of agricultural land, forested patches, riparian corridors, and growing subdivisions) that have allowed the deer population to grow and remain high. Indiana DVC numbers have steadily increased since 1997 with the number of DVCs increasing at a greater rate than traffic volume (McNew 2005a). Indiana has 92 counties and in 2004 eleven counties reported more than 300 DVCs compared to 10 in 2003. Forty-five counties showed an increase in DVCs over 2003 (24 counties > 15% increase) while forty-four showed a decrease (12 counties > 15% decrease; McNew 2005a). Furthermore, the number of collisions per billion miles was greater in 2003 than in 1995 when the deer population was at a peak (see Figure 1; McNew 2005a, b). One potential explanation for the increase in DVCs is change in the landscape structure of each county.

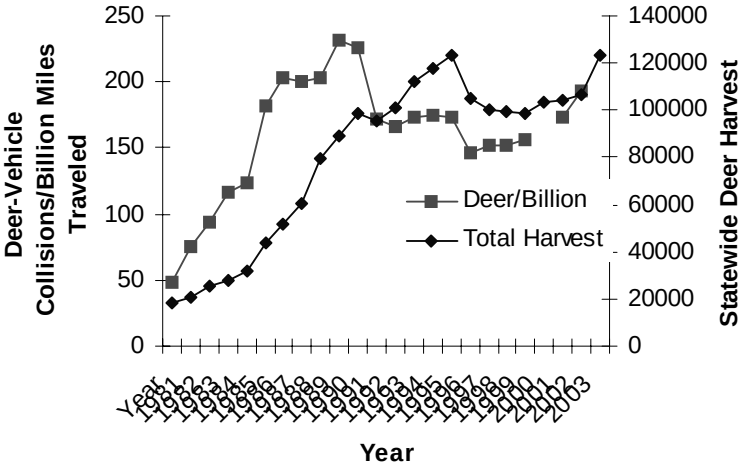

Fig. 1. Deer harvest and Deer-Vehicle collision demographics for the past 22 years in the state of Indiana (from McNew 2005a).

9.1.2 Mitigation: Reducing the number of DVCs

Several mitigation strategies have been tested with varied results (for full review see DeerCrash 2006). Mitigation strategies are typically placed into two categories: reducing the number of deer on highways, and increasing driver awareness or perception. The first category includes strategies such as fencing and wildlife passages (Bruinderink and Hazebroek 1996). While the second category ranges from the passive (e.g. permanent deer crossing signs) to the relatively more active (e.g. seasonal temporary flashing warning signs) and often are not analyzed for their effectiveness in preventing DVCs (Putnam 1997). Despite the best efforts of mitigation strategies, predicting where DVCs may occur is still an enigma. However, geospatial technologies could aid in predicting the location of DVCs thus indicating where mitigation strategies should be implemented**.**

9.1.3 Spatial autocorrelation and Likelihood maps

One strategy that might help the implementation of mitigation efforts in preventing DVCs would be to determine areas where DVCs might occur more frequently. Deer live and thrive in very specific habitat types and

landscape structures and humans often create this kind of habitat. Therefore landscape structure is important and influences the activity patterns of white-tailed deer. White-tailed deer require two types of habitat: forest for cover, and open areas for food. Therefore deer frequently associated with edge habitat. Human disturbances such as roads, agriculture, and development fragment habitats creating edge effects where white-tailed deer tend to aggregate.

One method to mitigate DVCs would be to examine the relationship between individual DVCs and the habitat type they occur. DVCs are unique because they only occur in spatially finite areas. As a rule of thumb vehicles typically only travel on roads and therefore DVCs only occur on roads. This provides the opportunity to examine the spatial relationship between DVC and habitat along a finite space. Therefore, one prediction about DVCs might be that they are spatially autocorrelated.

Spatial autocorrelation is an extension of autocorrelation into a two dimensional field (Ebdon 1987). Spatial autocorrelation can be loosely defined as random variables taking values that are more or less similar than would be expected for randomly distributed observations (Legendre1993), or the value of the spatial data is non-randomly related or interdependent over space (McGrew and Monroe 2000). There are three potential outcomes of spatial autocorrelation analysis: the points are similar and are positively correlated, the points are dissimilar and are negatively correlated, or the points are randomly distributed and show no correlation (Mcgrew and Monroe 2000; Ebdon 1987). Other analyses that are similar include spatial interpolation and spatial interaction (Miller 2004).

Analyzing the spatial relationship between DVCs would provide an index of “hotspots” that would identify where DVCs might occur more frequently. The first step in this anlaysis was to obtain records and locations of DVCs and geo-code the location of each DVC. By mapping location of DVCs, a deer-vehicle collision likelihood map generated in a GIS was created. The map has multiple layers that include DVC data, habitat maps, and landscape metrics to determine the probability of DVC in several areas throughout a defined area. The map was used to create an interpolated map DVC risk in any given area and is useful to land managers, conservationists, road planners, etc. as a practical tool that can aid in development, road placement, and even insurance rates. Additionally, buffer zones were created around DVCs to determine relationships between DVCs and habitat type.

The following steps were taken to investigate spatial autocorrelation of DVCs:

- Obtain DVC information and geo-code DVC locations for analysis of spatial autocorrelation.
- Examine spatial autocorrelation of DVCs
- Use land-cover maps and DVC data to create interpolated surface maps that identify "hotspots" of DVC and to correlate with habitat type.

9.2 Methods

9.2.1 Location

West central Indiana was chosen as a reference area to conduct this analysis because it is in the center of the white-tailed deer geographic range. Prior to western expansion in the United States Indiana was heavily forested with a small percentage of prairies and wetlands. Today the landscape is very different with 60% of the land being used for crops, 15% as managed forests, and the remaining land cover is urbanized (Jackson 1997). Within Indiana, Vigo County (39.467°N 37.414°W) was chosen as a representative location to examine spatial autocorrelation of DVCs (Figure 2). The largest city in Vigo County is Terre Haute, with a 2004 population estimate of 57,224. The population in the area has slowly been shrinking in recent years but there has been a shift in population from the central city area of Terre Haute to the South and to the East. Overall Vigo County is 1044km and had a human population of 104,77 in 2001 (US. Census, 2004). Vigo County consistently has over 200 DVCs annually.

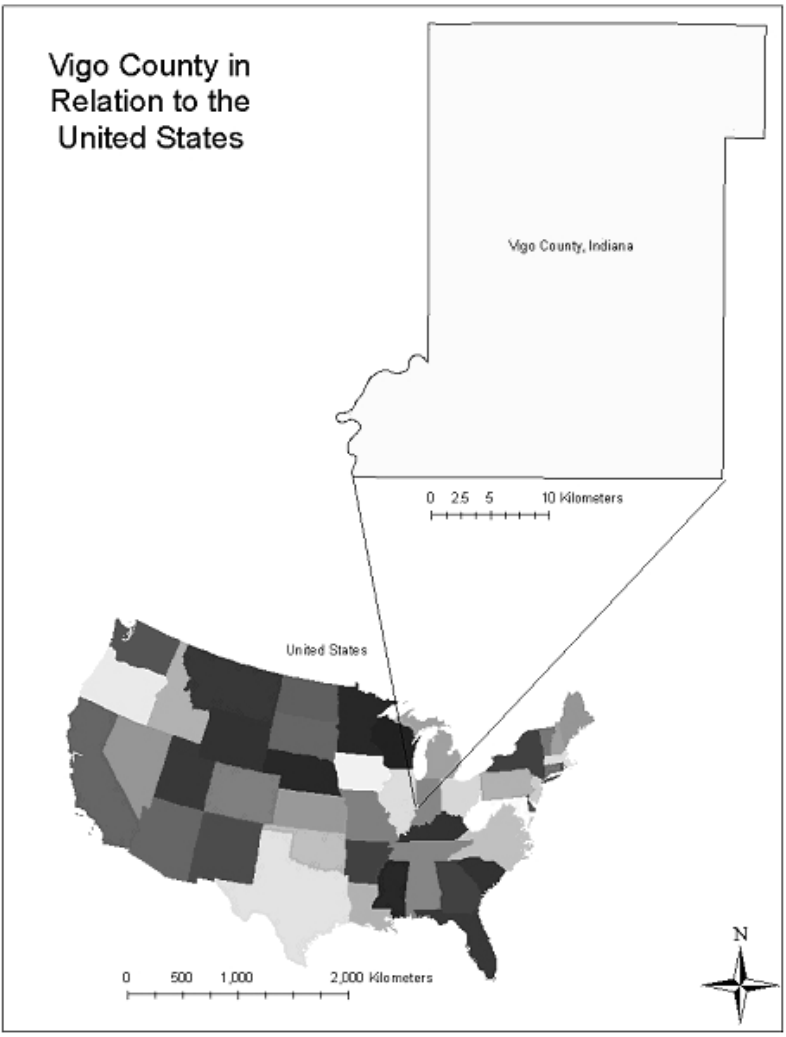

Fig. 2. The location of Vigo County, Indiana in the United States (from Horn, 2005).

9.2.2 Data

Deer-vehicle collision locations for the year 2000 were obtained from the Indiana Department of Transportation (IDOT). Location data were obtained from reports from law enforcement officers. All Indiana law enforcement officers are required to submit an "Officer's report" following a DVC investigation that involved either property damage over $750 or personal injury. Officer reports contain the location of the accident, road condition, light and weather conditions, and time of day.

The political boundary maps of Vigo County, Indiana and the tiger line road files were obtained from the United States Census GIS data website. The habitat and LULC data set were gathered from a 30m Landsat Thematic Mapper image obtained from the USGS at Indiana University's GIS data website.

9.2.3 Data Analysis

By combining deer habitat and road location data, a DVC likelihood map was produced using the inverse distance weighted (IDW) method in ArcGIS's geostatistical analysis package to interpolate the data. There are many equations and variations of equations used to determine the magnitude of spatial autocorrelation between data points. The most common equations are: Geary's *c*, Join-Count, G statistic, and Moran's *I*. We chose Moran's *I* as the most robust and commonly used spatial autocorrelation statistic. Moran's *I* is applied to zones or points with continuous values associated with them. It then compares the value at that point with the values at all other locations. The returned *I* value is between -1 and +1 with values of 1 indicating positive spatial autocorrelation and values of –1 indicating negative spatial autocorrelation. Values near zero have little or no spatial autocorrelation.

Inverse Distance Weighting (IDW) interpolation explicitly implements the principle of spatial autocorrelation to predict a value for an unmeasured location. IDW uses the measured values surrounding the predicted location to create values. Those measured values closest to the predicted location will have more influence on the predicted value than those farther away. Thus, IDW assumes that each measured point has a local influence that diminishes with distance. It weights the points closer to the prediction location greater than those farther away, hence the name inverse distance weighting.

The software used to accomplish this analysis was ArcGIS for Geo-coding. ArcGIS also contains a geo-statistical analysis extension that was used to complete Moran's *I* statistic and IDW interpolation. Crystal Reports was used to gather database information when selecting habitat criteria.

Several preparatory steps were needed to determine if DVCs are spatially autocorrelated. The first step was to geo-code all DVCs. Geo-coding uses the physical address of an accident and then assigns it a location on the map based on the physical address' location. Of the 230 DVCs in Vigo County during 2000, only 179 of them had sufficient address data to be geo-coded (Figure 3).

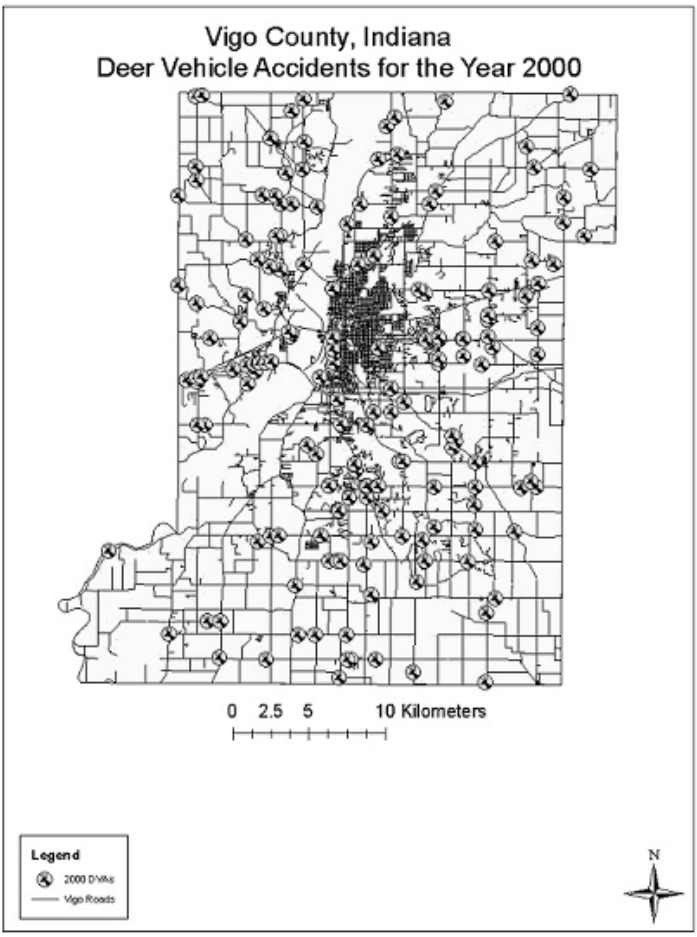

Fig. 3. Geo-coded locations of DVCs in 2000 for Vigo County, IN (from Horn 2005).

For this study we used a random sample of 30 DVCs selected out of the 179 DVCs that were Geo-coded. To determine the type of land cover near the accidents, a buffer was created around each DVC in order to overlay land cover data layers. Six buffer distances were used for the radius of the

accident buffers: 250m, 500m, 750m, 1000m, 1250m, and 1500m. These distances were chosen to represent the immediate area of a DVC and an approximate home range of white-tailed deer at the 1500m radius buffer. Once these distances were chosen, a buffer was placed around the sample of accidents and overlaid on a land use land cover map of the county (Figure 4). Each of the buffers was used to extract the land cover type through the LULC layer and analyzed using Crystal Reports. Crystal Reports analysis used the area of each specific land cover and divided it by the total area that it was selected from.

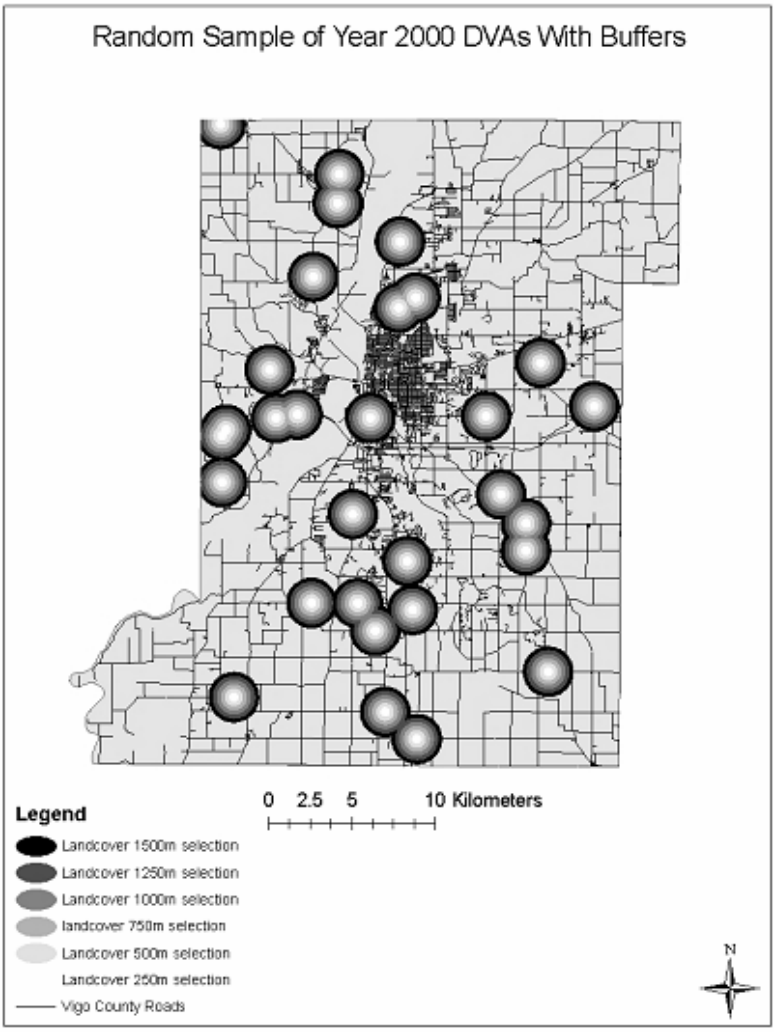

Fig. 4. Buffer zones for the randomly selected 30 DVCs used to analyze land cover and habitat type (from Horn 2005).

9.3 Results

There was no clear indication of spatial autocorrelation between DVCs for the year 2000 in Vigo County, Indiana. Four different pixel sizes were used and none showed a significant relationship between data points (Table 1). Habitat analyses indicated that nearly all the accidents in the study took place in or near the edge of forest and agricultural land. Approximately 60% of the land cover regardless of buffer size was agricultural land with 20% being forested (Table 2). Finally, an interpolated risk map was created that did show zones of higher frequency of accidents relative to other areas (Figure 5).

Table 1. Moran's *I* Results for each of the 4 pixel sizes used to determine the degree of spatial autocorrelation between DVCs**.**

Pixel Size	Moran's *I*
15m	0.0001
100m	0.0023
250m	0.053
500m	0.0685

Table 2. Land cover relationships to DVCs at different buffer zone widths.

Cover Type/Total area	**250m**	**500m**	**750m**	**1000m**	**1250m**	**1500m**
Agriculture	66.50%	71.30%	66.10%	63.50%	64.10%	64.00%
Forest	23.50%	21.20%	26.60%	28.40%	27.80%	27.70%
Residential/Urban Grass	10.00%	7.40%	5.10%	5.40%	5.30%	5.50%
Commercial	0.00%	0.00%	2.20%	0.00%	0.00%	0.00%
Woody Wetlands	0.00%	0.00%	0.00%	2.70%	2.80%	2.80%

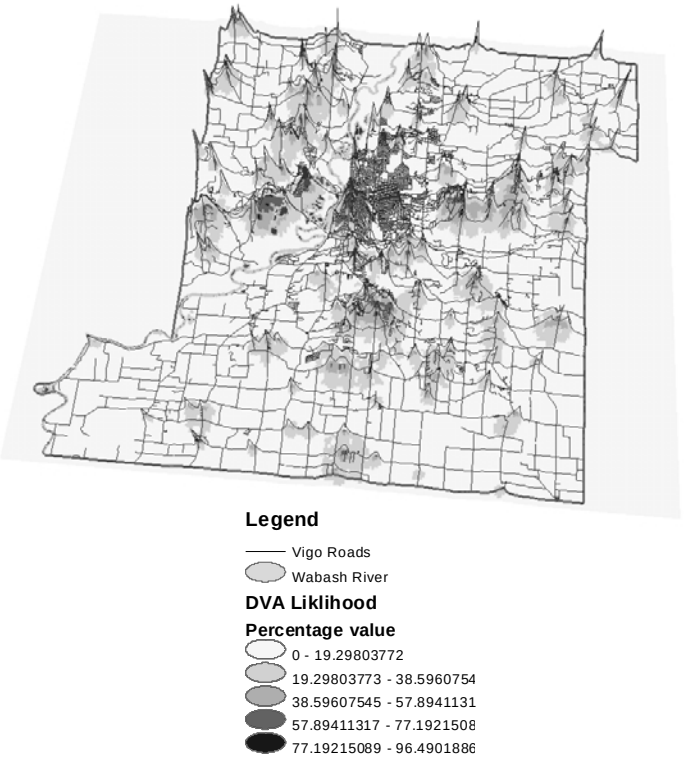

Fig. 5. Map of Vigo County, IN showing the interpolation of DVC incidences (from Horn 2005).

9.4 Discussion

The location of DVCs throughout Vigo County, IN was not significantly autocorrelated. Instead, DVCs seemed to be more evenly distributed across the area. A lack of autocorrelation could be attributed to the type of habitat primarily associated with DVC. In our study, most DVC occurred near agricultural fields, which serve as prime foraging habitats for white-tailed deer. Large agricultural fields may allow deer to forage in many places and as a result deer may cross roads in multiple points along the field rather than one central corridor. The highly distributed nature of a

their food source may explain the lack of spatial autocorrelation in DVCs for Vigo County.

Although in this study we did not demonstrate spatial autocorrelation between DVCs, our analyses did set the stage for future investigations of this type. Even though we only used 30 points, other analyses using this data set provided similar results. Therefore the major problem with this type of analysis is the limitations of the data including:

- Not all DVCs are reported due to a cost minimum for mandatory reports (cost minimums vary from state to state).
- Not all information is recorded (e.g. missing dates, and times). Much of the missing information is crucial to understanding DVC patterns. DVCs are influenced temporally. White-tailed deer vary their movement based not only on the time of year but the time of day as well.
- There is a time lag for reports. Databases are not readily accessible to query for DVCs, or accident reports are not filed electronically by the officer and are transcribed at a later date.
- Actual locations of DVCs are estimated or just provide the name of the road. Accurate and precise data on the location of DVCs are crucial for spatial autocorrelation analysis.
- May need several years worth of data from the same temporal segments (i.e. just the month of November) to use for analysis. This will effectively increase the data size but would also allow for a more robust analysis. White-tailed deer have seasonal movement patterns and DVCs increase during these seasonal activity periods. By focusing on DVCs that occur during certain temporal segments (e.g. November) we could generate more accurate interpolated risk maps.

These are just a few of the weaknesses of current DVC data. However, the biggest impediment to analyzing DVC data is lack of accurate and precise locations of DVCs. If spatially precise data could be obtained then our analyses would produce new insights that could help reduce the number of DVCs. One solution that would allow for the collection of spatially accurate data would be to better utilize global positioning system (GPS) technology. GPS technology has become relatively inexpensive and could be an easily added feature to police vehicles and/or Department of Transportation vehicles (carcass clean-up) that would be responding to DVCs. Additionally, the creation of a DVC reporting system where motorists could

submit DVC information without risk of penalty would greatly enhance data analysis It is believed that only about half of all the DVCs are actually reported (Conover et al. 1995; also see Deercrash 2006). A combination of many of the suggestions we have made may help provide the information needed to create an accurate risk map in order to reduce the number and severity of DVCs.

9.5 Conclusions and Broader Impacts

The spatial analysis made possible by GIS/RS data from this study was used to create an interpolated risk map, which can indicate areas that are more likely to produce DVCs. This information could then be used with further GIS analysis to indicate what landscape features (e.g. characteristics of the roadway or ecological habitat) are important in determining DVCs. This information will be valuable to urban planners and transportation officials when designing subdivisions, roads, and other forms of development and implementing mitigation strategies such as warning signs, speed limits, and public awareness.

References

Adams, C.E., Lindsey, K.J., and Ash, S.J. 2006. Urban Wildlife Management. Taylor & Francis Group, LLC. New York.

Allen, R. E. and McCullough, D.R. 1976. Deer-car accidents in southern Michigan. Journal of Wildlife Management 40: 317-325.

Bashore, T.L., Tzilkowski, W.M., and Bellis, E.D. 1985. Analysis of deer-vehicle collision sites in Pennsylvania. Journal of Wildlife Management 49: 769-774.

Bruindernink, G.W.T.A. and Hazebroek, E. 1996. Ungulate traffic collisions in Europe. Conservation Biology 10: 1059-1067.

Case, R.M. 1978. Interstate highway road killed animals: a data source for biologists. Wildlife Society Bulletin 6: 8-13.

Conover, M.R., Pitt, W.C., Kessler, K.K., DuBow, T.J., and Sanborn, W.A. 1995. Review of human injuries, illness, and economic losses caused by wildlife in the United States. Wildlife Society Bulletin 23: 407-414.

Curtis, P.D. and Hedlund, J.H. 2005. Reducing Deer-Vehicle Crashes, Wildlife Damage Management Fact Sheet Series, Cornell Cooperative Extension, Cornell University, Ithaca, NY.

DeerCrash. 2006. Countermeasures toolbox. Deer-Vehicle Crash Information Clearinghouse, University of Wisconsin, Madison, Wisc. www.DeerCrash.com, accessed January 02, 2006.

Donaldson, A. and Bennett, A. 2004. Ecological Effects of roads: Implications for the internal fragmentation of Australian parks and reserves. Parks Victoria Technical Series No. 12. Parks Victoria, Melbourne.

Ebdon, D. 1987. Statistics in Geography 2nd Edition. Basil Blackwell, Inc.

Forman, R.T.T. and Alexander, L.E. 1998. Roads and their major ecological effects. Annual Reviews of Ecology and Systematics 29: 207-231.

Groom, M.J., Meffe, G.K., and Carroll, C.R. 2006. Principles of Conservation Biology 3rd Edition. Sinauer Associates, Inc. Sunderland, MA.

Harlow, R. F. and Guynn, Jr., D.A. 1994. Population change and loss of habitat, pp. 218-224 *In* D. Gerlach, S. Atwater, and J. Schnell (ed.), The Wildlife Series: Deer, Stackpole Books, Mechanicsburg, PA.

Horn, S. 2005. Spatial autocorrelation of deer/vehicle accidents in Vigo County. MS. Thesis. Indiana State University.

Hubbard, M.W., Danielson, B.J., and Schmitz, R.A. 2000. Factors influencing the location of deer-vehicle accidents in Iowa. Journal of Wildlife Management 64: 707-713.

HuntingNet. 2006. Top 10 states for deer collisions. www.huntingnet.com/news/pf_news.aspx?news_id=524, accessed January 06, 2006.

Insurance Institute for Highway Safety. 1993. Deer, moose collisions with motor vehicles peak in spring and fall. Status Report April Volume 28. Number 4.

Jackson, M.T. 1997. The Natural Heritage of Indiana. Indiana University Press. Bloomington, IN.

Jacobson, H.A. 1994. Feeding behavior and diet. pp. 192-202. *In* D. Gerlach, S. Atwater, and J. Schnell (ed.), The Wildlife Series: Deer, Stackpole Books, Mechanicsburg, PA.

Kilpatrick, H.J., LaBonte, A.M., and Seymour, J.T. 2002. A shotgun-archery deer hunt in a residential community: evaluation of hunt strategies and effectiveness. Wildlife Society Bulletin 30: 478-486.

Legendre, P. 1993. Spatial autocorrelation: trouble or a new paradigm? Ecology 74: 1659-1673.

McGrew, Jr., J.C. and Monrore, C.B. 2000. An Introduction to Statistical Problem Solving in Geography. McGraw-Hill.

McNew, L.B. 2005a. White-tailed deer population Management: Deer-vehicle collisions. Indiana Division of Fish and Wildlife Federal Aid Program Report W-26-R-37. Job 13-B-7. 10pp.

McNew, L.B. 2005b. White-tailed deer population Management: Mandatory Deer Check Stations. Indiana Division of Fish and Wildlife Federal Aid Program Report W-26-R-36. Job 13-B-7. 10pp.

Meffe, G.K., Carroll, C.R., and Groom, M.J. 2006. What is conservation biology? pp. 1-25. *In* M.J. Grrom, G.K. Meffe, and C.R. Carroll (ed.), Principles of Conservation Biology 3rd Edition. Sinauer Associates, Inc. Sunderland, MA.

Miller, H.J. 2004. Tobler's first law and spatial analysis. Annals of the Association of American Geographers 94: 284-289.

Nielsen, C.K., Anderson, R.G., and Grund, M.D. 2003. Landscape influences on deer-vehicle accident areas in an urban environment. Journal of Wildlife Management 67:46-

Pojar, T.M. Prosence, R.A. Reed, D.F., and Woodward, RH. 1975. Effectiveness of a lighted, animated deer crossing sign. Journal of Wildlife Management 39: 87-91.

Putman, R.J. 1997. Deer and road traffic accidents: options of rmanagement. Journal of Environmental Managemnent 51: 43-57

Roseberry, J.L and Woolf, A. 1998. Habitat-population density relationships for white-tailed deer in Illinois. Wildlife Society Bulletin 26: 252-258.

Rothely, K. 2001. Manipulative, multi-standard test of a white-tailed deer habitat suitability model. Journal of Wildlife Management 65: 953-963.

Strickland, M.D., Harju, H.J., McCafffery, K.R., Miller, H.W., Smith, L.M., and Stoll, R.J. 1996. Harvest Management. pp. 445-473 In T.A. Bookhout (ed.), Research and Management Techniques for Wildlife and Habitats. The Wildlife Society, Bethesda, Md.

Sullivan, T.L. and Messmer, T.A. 2003. Perceptions of deer-vehicle collision management by state wildlife agency and department of transportation administrators. Wildlife Society Bulletin 31: 163-173.

Trombulak, S.C. and Frissell, C.A. 2000. Review of ecological effects of roads on terrestrial and aquatic communities. Conservation Biology 14: 18-30.

US.Census. 2004. Vigo County, IN Census. http://quickfacts.cesus.gov/qfd/states/18/18167.html. Accessed Fall 2004.

Woolf, A. and Roseberry, J.L. 1998. Deer management: our profession's symbol of success or failure? Wildlife Society Bulletin 26: 515-521.

10 Scale and Spatial Autocorrelation From A Remote Sensing Perspective

J. Scott Spiker Department of Geography, University of Wisconsin - Parkside, Kenosha, Wisconsin

Timothy A. Warner Department of Geology and Geography, West Virginia University, Morgantown, West Virginia

10.1 Introduction

One of the challenges for urban and regional planners and other users of remotely sensed imagery is how to select the appropriate data for a particular monitoring or mapping problem. In the past, the dearth of available imagery meant that the problem itself usually had to be adapted to fit the data, which was typically limited to either high spatial resolution film-based aerial imagery, or coarse-spatial resolution digital satellite imagery. Today, a vast range of aerial and satellite imagery is available (Kramer, 2002), opening a new range of potential scales of problems that can be investigated. However, these new options also place additional burdens on the remote sensing user, who, in selecting data, has to consider differences in spectral, temporal, radiometric, and spatial characteristics of the imagery. Spatial properties are particularly important, and the pixel size of current sensors varies over more than three orders of magnitude (from 0.6 m to 1 km and larger) (Kramer, 2002).

When imagery is manually interpreted, there is a general expectation that finer spatial resolution makes interpretation easier and more accurate. However, it is obvious that too much detail, in addition to increasing the cost of data acquisition, will result in an overwhelming amount of data. Furthermore, when it comes to digital image analysis, it has been found that greater spatial resolution may lead to lower classification accuracy. Perhaps most importantly, it is important to match the scale of the data

used to the scale of the spatial patterns of the physical phenomenon of interest, in order to avoid the problem of ecological fallacy (Marceau, 1999). Thus the ability to measure and quantify spatial pattern is crucial for any mapping activity.

One of the most important measures of spatial structure of a variable is spatial autocorrelation. A non-random spatial pattern may show either positive or negative spatial autocorrelation. In the case of positive spatial autocorrelation, the value of a variable at a given location tends to be similar to the values of that variable in nearby locations. In other words, if the value of some variable is low in a given location, the presence of positive spatial autocorrelation indicates that nearby values are also likely to be relatively low. Conversely, negative spatial autocorrelation is characterized by a tendency for dissimilar values to cluster in proximate locations. For example, areas exhibiting low values for a particular variable may be surrounded by high values when negative spatial autocorrelation exists. The absence of spatial autocorrelation indicates that the spatial arrangement of the variate values is random.

In a remote sensing context, the variate values of interest in a spatial autocorrelation analysis are commonly the digital number (DN) values of the pixels in an image. In general, adjacent pixels are more likely to display similar values than pixels that are more spatially separated (Woodcock et al., 1988). Spatial autocorrelation in an image can be affected by object size, spacing, and shape (Jupp et al., 1988). In addition, the sensor's spatial resolution will have an effect on the overall spatial autocorrelation present in a remotely sensed image (Jupp et al., 1988). The term "regularization" is used to describe the imposition of a discrete sampling template, such as a pixel, on the underlying phenomenon (Jupp et al., 1989).

To the remote sensing practitioner, spatial autocorrelation is of interest for two important reasons. Firstly, the presence spatial autocorrelation indicates that the assumption that the data are independently distributed, a requirement for many statistical procedures, is violated. As a result, many traditional statistical procedures may be inappropriate for analyzing these data. Instead, methods of analysis that account for the inherent spatial nature of the data should be applied.

Secondly, the spatial information itself may prove useful in image analysis. Examples of the use of autocorrelation statistics include the reduction of noise in imagery (Switzer and Ingebritsen, 1986), planning ground data collection (Curran, 1988), analysis of tree canopy (Cohen et al., 1990) and

forest structure (St-Onge and Cavayas, 1996), estimation of optimal image spatial resolution (Hyppänen, 1996), feature selection (Warner & Shank, 1997), image compression (Warner, 1999), coral reef health analysis (LeDrew et al., 2004), sensor calibration (Bannari et al., 2005), mapping vineyards and orchards (Warner and Steinmaus, 2005), and classification of urban land-use (Wu, et al., 2006) .

This chapter provides an overview of the different methods of quantifying spatial autocorrelation in remotely sensed images, using as a case study an image of a small urban community. We start with a discussion of global measures of spatial autocorrelation, which summarize the spatial pattern across an entire image. We will show how the scale of the phenomena imaged can be observed in autocorrelation data, and how that scale varies as the pixel size changes. Global measures of spatial autocorrelation may mask the complexity and non-uniformity of the underlying spatial structure. Therefore, following the discussion of global measures of autocorrelation, we will describe local measures of spatial autocorrelation, and show how the local statistics complement the analysis of the global statistics.

10.2 Measuring Spatial Autocorrelation

10.2.1 Global measures of spatial Autocorrelation

Three commonly used global methods for measuring spatial autocorrelation have been defined: semivariance (Matheron, 1971), Geary's *c* (Geary, 1954), and Moran's *I* (Moran, 1948). For each of these measures, the distance over which spatial association is investigated is called the *lag*. When semivariance is plotted against lag, a *variogram* is produced. For Geary's *c* and Moran's *I* the equivalent graph is termed a *correlogram*. Because each of the spatial autocorrelation metrics utilizes different operational approaches to quantify the degree of spatial autocorrelation present, each of these three approaches is discussed in more detail below.

Semivariance is a measure of the average variance of the differences between all pairs of measurements that are separated by the lagged distance (Curran 1998).

$$\gamma(h) = \frac{1}{2N}\sum_{1}^{N}[Z(x_i) - Z(x_i + h)]^2 \qquad \text{eqn. 1}$$

Where: h is the spatial lag
$Z(x_i)$ is the DN value at location x_i
and N is the number of DN pairs at lag h.

The semivariogram is the relationship between semivariance and lag. Two key characteristics of a typical simple semivariogram are the *range* and *sill*. The sill is the plateau, the range of lags where semivariance does not vary with lag. The range is the lag at which the sill is reached. Multiple sills may indicate a repetitive pattern in the data, such as a gridded street network or rows of crops in an agricultural field. A key advantage with the semivariogram is that it has been studied extensively, and a strong theoretical understanding exists about its behavior (e.g. Jupp et al., 1988).

The Geary's *c* statistic is based on the squared difference between spatially lagged pairs of pixels, normalized by the overall scene variance. In a remote sensing context, Geary's *c* statistic can be defined as:

$$c = \frac{(n-1)}{2\sum_i\sum_j W_{ij}} \bullet \frac{\sum_i\sum_j W_{ij}(x_i - x_j)^2}{\sum(x_i - \bar{x})^2} \qquad \text{eqn. 2}$$

Where: n is the number of observations
W_{ij} is the weighted spatial lag between pixels i and j
and x_i and x_j are the DN values of pixels i and j.

Values for Geary's *c* range from 0 to 2, with 0 indicating maximum positive spatial autocorrelation and 2 maximum negative spatial autocorrelation. The expected value of 1 signifies a lack of spatial autocorrelation.

A third measure of spatial autocorrelation is Moran's *I*. In this approach, spatial autocorrelation is calculated as a function of the covariation between pixels *i* and *j*. This statistic is calculated as follows:

$$I = \frac{n}{\sum_i\sum_j w_{ij}} \bullet \frac{\sum_i\sum_j w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum(x_i - \bar{x})^2} \qquad \text{eqn. 3}$$

Where: W_{ij} is the weighted spatial lag between pixels *i* and *j*
and x_i and x_j are the DN values of pixels i and j.

Moran's *I* values can range from a maximum of +1, which indicates strong positive spatial autocorrelation, to a minimum of -1, which indicates strong negative spatial autocorrelation. The expected value of 0 denotes a random spatial arrangement (the precise expectation is -1/(n-1) (Goodchild, 1986), which tends to 0, as n gets large).

Although Moran's *I* and Geary's *c* are attempts to measure a conceptually similar property, using respectively covariation and squared differences, the two statistics are noted from empirical evidence to exhibit a strong, but not perfectly linear, inverse relationship (Sawada, 2004).

As with semivariance, both Moran's *I* and Geary's *c* can be plotted as a function of lag. These correlograms give valuable information about spatial structure, in a fashion analogous to semivariance. One advantage of the correlograms over the semivariograms, is that the former can be used to differentiate positive and negative autocorrelation.

10.2.2. Local measures of spatial autocorrelation

One criticism of semivariance, as well as Moran's *I* and Geary's *c* statistics, is that these global measures of spatial autocorrelation potentially ignore important local variation in the data. In response to this shortcoming, Getis and Ord (1992) introduced a local autocorrelation measure, the G_i statistic. Anselin (1995) subsequently proposed *local indicators of spatial autocorrelation* (LISA) as a general means for decomposing global autocorrelation measurements so that the individual contribution of each observation can be assessed, and local "hot spots," or areas of spatial nonstationarity, identified (Anselin, 1995).

The Getis and Ord (1992) *G* statistic compares pixel values at a given location with those pixels at a lag, *d*, from the original pixel at location *i*. Getis has devised two methods of calculating the local *G* statistic, G_i and G_i^*. The difference between these two approaches is that the G_i measure does not include the i^{th} observation as part of the calculation, while the G_i^* measure does include that value. The calculation of the statistic in the context of remotely sensed data is as follows (Ord and Getis, 1995, Wulder and Boots, 1998):

$$G_i(d) = \frac{\sum_j w_{ij}(d)x_j - W_i x}{s[W_i(n - W_i)/(n-1)^{1/2}}, j \neq i \qquad \text{eqn. 4}$$

Where: $W_i = \Sigma_j w_{ij}(d)$
$w_{ij}(d)$ is the weighted spatial lag distance "d" between pixels *i* and *j*
and x_i and x_j are the DN values of pixels i and j.

In the case of both G_i and G_i^*, high positive values indicate a clustering of high values, while highly negative values reveal clusters of low values. Under the null hypothesis, the expected value of 0 indicates that no clustering is occurring at the specified spatial lag *d*. Wulder and Boots (1998) have noted that the Getis-Ord G_i statistic gives information about both the degree of clustering and the average values of the cluster. As a consequence, the G_i cannot differentiate between a lack of autocorrelation, and a cluster of average values.

Since Moran's *I* and Geary's *c* have been already been defined above, we will not give the specific formula for the LISA versions of these statistics. However, it is important to note that the Anselin (1995) LISA statistics have as a defining property that the sum of all the local values is proportional to the global measure. Thus, the Anselin LISA measures provide an excellent overview of the spatial distribution of the constituents of the global measure. The Anselin local Geary's *c* tends to highlight boundaries in images, whereas the Anselin local Moran's *I*, has been described as a way to identify "hot spots" (Anselin 1995) that represent clusters of high or low values (which are not differentiated). Unlike with the Getis-Ord G_i, the local Moran's *I* and Geary's *c* can be used to identify negative autocorrelation, although this property is rather rare in images.

10.3 Data and Processing

We will illustrate the use of autocorrelation measures with a case study focusing on Morgantown, West Virginia, USA (Figure 1). This area is characterized by predominantly low-density residential land-use, with an area of relatively high-density urban land-use in downtown Morgantown adjacent to a portion of the Monongahela River. A few forested areas and agricultural fields are found in the outskirts of the region. The data used for

this case study (Figure 1) is derived from QuickBird panchromatic satellite imagery with a nominal 0.7 meter spatial resolution, and spectral sensitivity from 445 to 900 nm (i.e. blue-green to near-infrared).

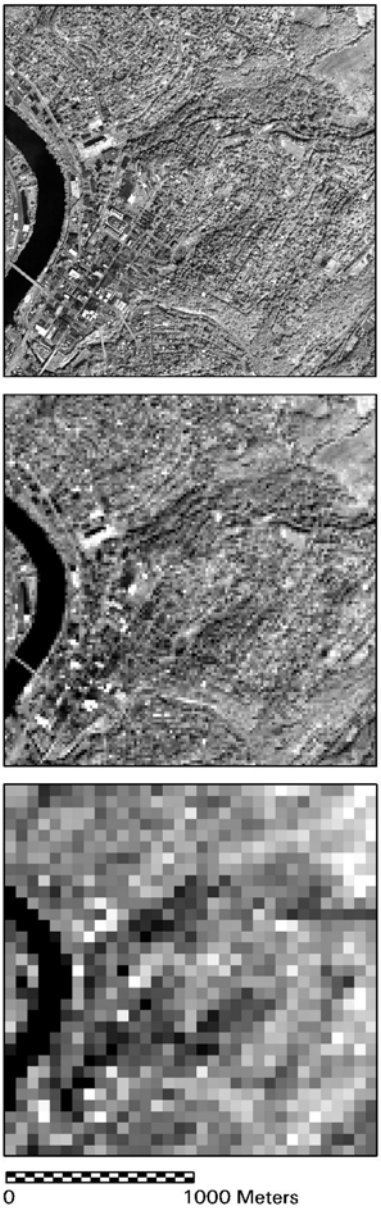

Fig 1. Study Area - Morgantown WV at three different pixel sizes. Top: 0.7m. Middle: 15 m. Bottom: 60 m.

The image processing was undertaken using the remote sensing software packages Leica's Erdas Imagine and ITT Visual Information Solutions' ENVI. Imagine was used to subset the imagery, and generate images at four coarser scales through pixel aggregation (Table 1). This results in images at a total of five scales, each representative of a typical group of sensors (Table 1).

Table 1. Spatial resolutions used in this study, and representative sensors that have a similar spatial resolution.

Pixel Size	Representative sensors with similar resolution
0.7 m	QuickBird Panchromatic, IKONOS Panchromatic
5 m	IRS RESOURCESAT LISS IV, SPOT-5
15 m	Landsat ETM+ Panchromatic, ASTER VNIR bands
30 m	Landsat Multispectral bands, ASTER SWIR bands
60 m	IRS RESOURCESAT AWiFs, LANDSAT MSS

ENVI was used to calculate the global and local statistics. For all the statistics, the Queen's neighborhood criterion was used. The Queen's criterion defines the neighborhood as comprising all eight adjacent pixels, in other words, adjacent pixels in the directions of the rows, columns, and diagonals. The advantage with this definition of neighborhood is that it gives a measure of the pattern averaged over all directions.

10.4 Discussion

Spatial resolution clearly has a major influence on the scale of objects that can be resolved in an image. Figure 1 shows that, for example, even with the intermediate-sized 15 meter pixels, a major 4 lane bridge across the Monongahela River is only just discriminated, and with 60 meter pixels, the bridge is not visible. What is perhaps less obvious is that as the pixel size becomes coarser, the overall variance in the image is decreased. This is shown by the variograms (Figure 2), where the larger pixels are associated with a reduced sill height, an observation that corresponds to theoretical and experimental observations (Woodcock et al. 1988). Note that the variogram starts flattening out at approximately 30 meters, but only reaches the sill at a lag of approximately 200 meters. (The 0.7 m data could only be calculated out to a lag of 200 pixels, due to an apparent limitation in the ENVI software.) If we compare the semivariance of the images at the lag of one pixel (Table 2), it is noticeable that the 0.7 m data does not follow the trend of the remaining scales, where coarser pixel sizes are associated with a gradual reduction in semivariance. This suggests that there is a qualitative difference in the scale of the objects resolved between the 0.7 m data and the coarser resolution images.

Some interesting differences are evident in comparing the Moran's *I* and Geary's *c* correlograms (Figure 2). The Geary's *c* graph shows the values reaching a plateau very close to 1 (i.e. a random spatial structure) at approximately 200 m. The 0.7 m data are more random (higher values) than the coarser pixel data. On the other hand, the Moran's *I* graph does not show the data quite reach a plateau, even at the maximum lag of 300 m. This is especially noticeable for the coarsest spatial resolution 60 m pixels. Conversely, there is a relatively consistent trend for both Geary's *c* and Moran's *I* that autocorrelation at the scale of one pixel is generally reduced as the pixel size increases from 0.7 to 15 m, and with almost no change from 15 to 60 m (Table 2). This again emphasizes that the very finest scales of data are qualitatively different from coarser scales.

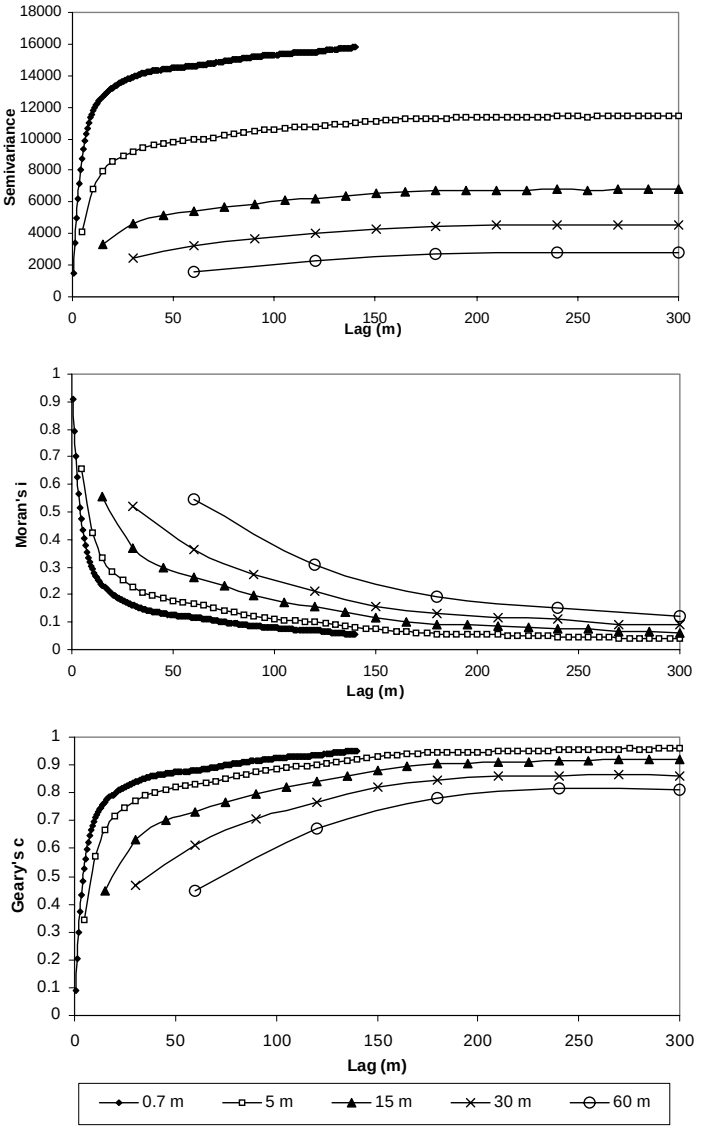

Fig. 2. Spatial structure of the Morgantown image. Top: Semivariogram. Middle: Moran's i correlogram. Bottom: Geary's c correlogram..

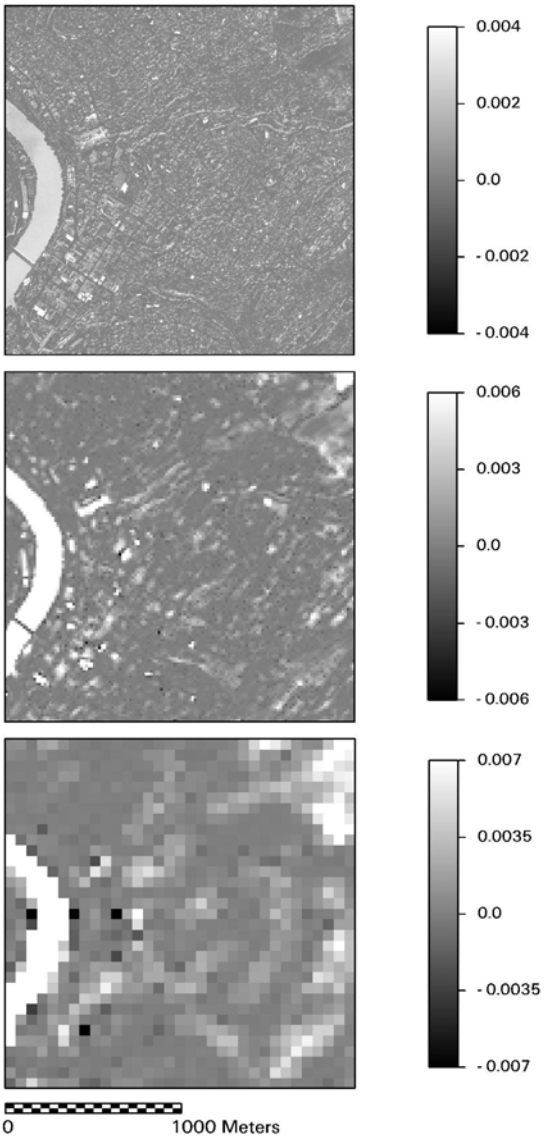

Fig. 3. Anselin's local Moran's I images, at a lag of 1 pixel. Top: 0.7 m pixels. Middle: 15m. Bottom: 60m.

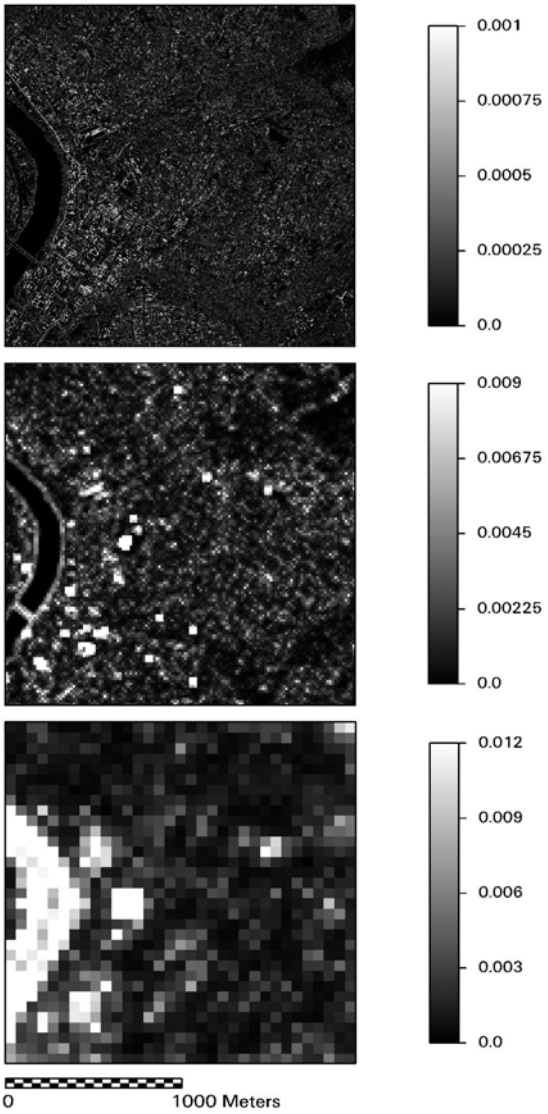

Fig. 4. Anselin's local Geary's *c* images at a lag of 1 pixel. Top: 0.7 m pixel sizes. Middle: 15m. Bottom: 60m.. .

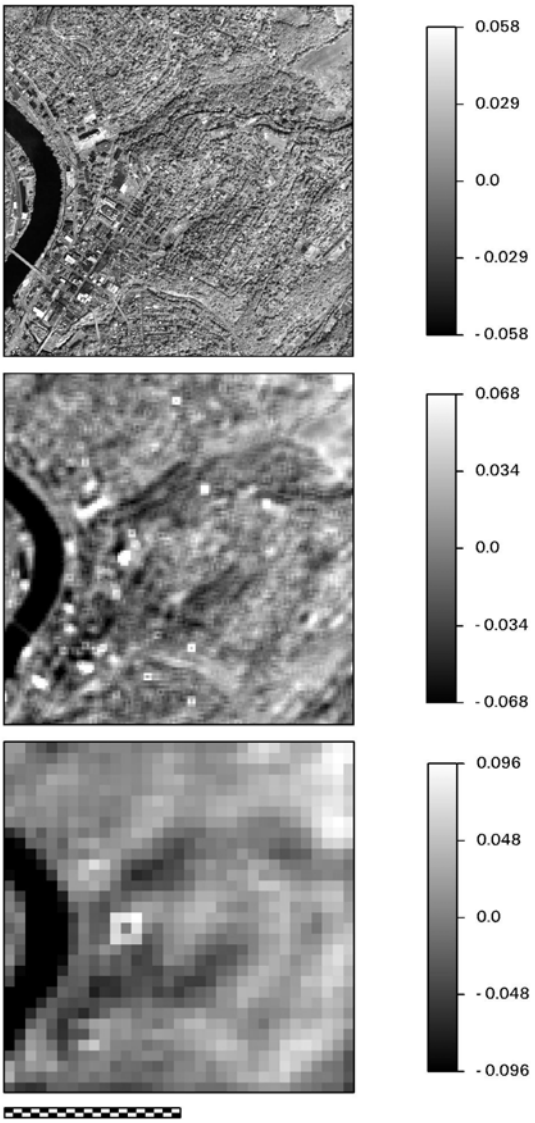

Fig.5. Getis-Ord G_i images at a lag of 1 pixel. Top: 0.7 m pixel sizes. Middle: 15m. Bottom: 60m.

Table 2. Global spatial autocorrelation at a lag of one pixel.

Measure	0.7 m	5 m	15 m	30 m	60 m
Semivariance	1494	4081	3293	2452	1533
Geary's *c*	0.09	0.34	0.45	0.47	0.45
Moran's *I*	0.91	0.66	0.55	0.52	0.54

The images of the local autocorrelation measures help explain the patterns observed with the global measures. With 0.7 m pixels, the local Moran's I is dominated by intermediate values (i.e. close to 0), with scattered small clusters of high values associated with individual buildings and other urban features (Figure 3). As the pixel is coarsened to 15 m and 60, the river, and to a lesser extent the contrast between urban and rural land use, becomes the major contributor to the overall variation (note from the legend that the 60 m data are associated with greater overall variability than the 0.7 m data). Thus, the global Moran's I is quantifying the broad landscape pattern with 60 m pixels, whereas with 0.7 m pixels, elements of the urban infrastructure are important.

The local Geary's *c* images (Figure 4) show a somewhat similar pattern of the changing resolution of objects imaged, with an even more distinctive changing role of the river. At 0.7 m, the river has the lowest values (i.e. highest autocorrelation), whereas at 60 m it has the highest (i.e. lowest autocorrelation). However, a few distinctive urban features still have an influence at the coarsest scale, and the contrast between urban and rural areas seen in the local Moran's *I* is not observed.

The Getis-Ord G_i images show a superficial resemblance to the original satellite images. This is partly because the G_i statistic is dependent in part on the average brightness of the local group of pixels over which it is calculated. Thus, the river remains a low value region at all scales. The distinctive bright donut shape in the 60 m local G_i image is a product of the fact that for the G_i statistic, the central pixel of the neighborhood is not included in the calculation. To avoid this artifact, the G_i* may be a better choice for remote sensing analysis (Wulder and Boots, 1998).

10.5 Concluding Remarks

The issue of matching the scale of image acquisition to the scale of the phenomenon to be mapped is important. Autocorrelation measures provide a way to quantify the scale and spatial structure of images at different scales. Global measures are useful, but local measures are necessary to

understand the dominant contributors to the global metrics. Local autocorrelation statistics help to overcome some of this masking effect by revealing areas of spatial non-stationarity.

The empirical examples presented here provide insight as to how scale can be evaluated. The 0.7 m imagery of this case study provided a fundamentally different spatial scale of information than coarser scale imagery that had pixel sizes 5 m and greater. This may have implications not just for the scale of phenomenon that can be observed, but also for the nature of the automated analysis that can be applied. Future research exploring these relationships in more detail will help to shed light on the interaction between scale and autocorrelation.

Acknowledgements

The authors would like to thank West Virginia View (www.wvview.org), AmericaView and the USGS for the QuickBird imagery used in this study.

References

Anselin, L., 1995. Local Indicators of Spatial Autocorrelation – LISA. Geographical Analysis 27(2):93-15.

Bannari, A., K. Omari, P. Tellet, and G. Fedosejevs, 2005. Radiometric Uniformity and Stability of Test Sites Used for the Calibration of Earth Observation Sensors. IEEE Transactions on Geosciences and Remote Sensing 43(12):2918-2926.

Cohen, W. B., T. A. Spies, and G. A. Bradshaw, 1990. Semivariograms of Digital Imagery for Analysis of Conifer Canopy Structure. Remote Sensing of Environment 34:167-178.

Curran, P. J., 1988. The Semivariogram in Remote Sensing: An Introduction. Remote Sensing of Environment 24: 493-507.

Geary, R., 1954. The contiguity ratio and statistical mapping. The Incorporated Statistician 5**:**115-145.

Getis, A., and J. K. Ord, 1992. The analysis of spatial association by use of distance statistics. Geographical Analysis 24(3):189-206.

Goodchild, M. F., 1986. Spatial Autocorrelation. Geo, Norwich, United Kingdom, 56 pp.

Hyppänen, H., 1996. Spatial Autocorrelation and Optimal Spatial Resolution of Optical Remote Sensing Data in Boreal Forest Environment. International Journal of Remote Sensing 17(17):3441-3452.

Jupp, D. L. B., A. H. Strahler, and C. E. Woodcock, 1988. Autocorrelation and Regularization in Digital Images I. Basic Theory. IEEE Transactions on Geoscience and Remote Sensing 26(4):463-473.

Jupp, D. L. B., A. H. Strahler, and C. E. Woodcock, 1989. Autocorrelation and Regularization in Digital Images II. Simple Image Models. IEEE Transactions on Geoscience and Remote Sensing 27(3):247-258.

Kramer, H. J., 2002. Observation of the Earth and its Environment. Springer, Berlin, Germany, 1510pp.

LeDrew, E. F., H. Holden, M. A. Wulder, C. Derksen, C. Newman, 2004. A spatial statistical operator applied to multidate satellite imagery for identification of coral reef stress. Remote Sensing of Environment 91:271-279.

Marceau, D., 1999. The scale issue in social and natural sciences. Canadian Journal of Remote Sensing 25(4): 347-356.

Matheron, G., 1971. The theory of regionalized variables and its applications. Les Cahiers du Centre de Morphologie Mathematiques de Fontainebleau No. 5, Fontainebleau, France.

Moran, P. A. P., 1948. The interpretation of statistical maps. Journal of the Royal Statistical Society, series B:246-251.

Ord, J., and A. Getis, 1995. Local spatial autocorrelation statistics: Distributional issues and an application. Geographical Analysis 27**:**286-306.

Sawada, M. 2004. Global Spatial Autocorrelation Indices - Moran's I, Geary's C and the General Cross-Product Statistic. Research paper from the Laboratory for Paleoclimatology and Climatology at the University of Ottawa. (as posted online at: http://www.lpc.uottawa.ca/publications/moransi/moran.htm)

St-Onge, B. A. and F. Cavayas, 1997. Automated Forest Structure Mapping from high Resolution Imagery Based on Directional Semivariogram Estimates. Remote Sensing of Environment 61:82-95.

Switzer, P., and S. E. Ingebritsen, 1986. Ordering of Time-Difference Data from Multispectral Imagery. Remote Sensing of Environment 20(1): 85-94.

Warner, T. A., 1999. Analysis of spatial patterns in remotely sensed data using multivariate spatial correlation. Geocarto International 14(1):59-65.

Warner, T. A. and M. C. Shank, 1997. Spatial Autocorrelation Analysis of Hyperspectral Imagery for Feature Selection. Remote Sensing of Environment 60:58-70.

Warner, T. and K. Steinmaus, 2005. Classification of orchards and vineyards with high spatial resolution panchromatic imagery. Photogrammetric Engineering and Remote Sensing 71(2):179-187.

Woodcock, C. E., A. H. Strahler, and D. L.B. Jupp, 1988. The Use of Variograms in Remote Sensing: I. Scene Models and Simulated Images. Remote Sensing of Environment 25:323-348.

Wu, S-S., B. Xu, and L. Wang, 2006. Urban Land-use Classification Using Variogram-based Analysis with an Aerial Photograph. Photogrammetric Engineering & Remote Sensing 72(7): 813-822.

Wulder, M. and B. Boots, 1998. Local spatial autocorrelation characteristics of remotely sensed imagery assessed with the Getis statistic. International Journal of Remote Sensing 19(11):2223-2231.

11 The Spatial Imperatives of Environmental Justice

Trevor Fuller, Department of Geography, Geology, & Anthropology, Indiana State University, Terre Haute, Indiana, USA

Jay D. Gatrell, Department of Geography, Geology, & Anthropology, Indiana State University, Terre Haute, Indiana, USA

E. LaFary, School of Geography & Environmental Sciences, University of Auckland

The nexus between environmental justice and geo-technologies is an evolving one. That is to say, geographic information systems, remote sensing, and other technologies have the capacity to locate and situate the politics and place-based dangers of environmental risk within a broader conceptual and policy framework. Conceptually, GIScience has the capacity to chart new geographies of environmental risk across the urban and rural landscape. Empirically, GIScience has the capacity to map heretofore disparate datasets in an attempt to unlock the socio-economic determinants of "who's at risk and where?" In this paper, we build on the earlier work of Buzzelli to explore the socio-spatial dynamics of environmental risk in Terre Haute, Vigo County, Indiana. Using GIS, remote sensing, census, and environmental data, the paper presents a framework for unlocking the spatial dynamics of socioeconomic status and environmental risk across urban and rural neighborhoods in Vigo County.

11.1 Study Area

The study area is located in the state of Indiana located within the United States. Situated on the banks of the Wabash River, Terre Haute, Indiana is the county seat of Vigo County (Fig. 1). Terre Haute had a 2000 population of 69,614 with an observed county wide median income of $33,184 and a median housing value of $72,500 (U.S. Census, 2002). There is

considerable variety in the land use encountered in Terre Haute and Vigo County with dense and mixed urban, parks, suburban, and rural/agricultural regions present. In this respect, Terre Haute and Vigo County are typical of moderate Midwestern metropolitan areas.

Fig. 1. Vigo County, Indiana

11.2 Placing & Scaling Environmental Justice

Environmental justice has been defined many ways over the last 30 years. The dominant narrative suggests that specific populations, particularly marginalized groups, are being subjected to a disproportionate amount of risk from environmental disamenities. Disamenities, as used here, refer to commonly used indicators of environmental quality, such as the location of certain types of facilities and spills or releases to the environment. Most often in such research the differentiation among the population occurs by means of socioeconomic/demographic characteristics. The United States Environmental Protection Agency (U.S. EPA) defines environmental justice as "the fair treatment and meaningful involvement of all people regardless of race, color, national origin, or income with respect to the development, implementation, and enforcement of environmental laws, regulations, and policies." (U.S. EPA Office of Environmental Justice 2006).

The wide variability in the strength of correlation between socioeconomic status and environmental disamenities further fuels the controversy as to whether certain populations do indeed bear a disproportionate amount of environmental risk (Cutter, et al. 2001). Such variability is exemplified in research that has indicated people of color were disproportionately exposed, especially working-class Latinos (Pulido 2000), while other research demonstrated a strong relationship between environmental risk and dwelling value, as well as lone-parent families (Buzzelli, et al. 2003). Such mixed results have been revealed over the years, with significant shifts in the relative role of demographic and socioeconomic conditions in determining disproportionate environmental risk (Cutter, et al. 2001). Some research has not revealed any direct relationship between minority populations and disproportionate environmental risk (Anderton et al. 1994). In addition to the changing place of demographics within research, another contributor toward the varying results was the wide array of study areas used in environmental justice research. In the following paragraphs we will discuss the issue of an appropriate scale of analysis.

The area of analysis has varied widely within environmental justice research, with much of the research focusing on a city-wide analysis (Mohai and Bryant 1992, Buzzelli, et al. 2003, Pulido 2000). Some research in the environmental justice field has been designed to model environmental risk at the county, state, and even national level (Margai 2001, Pastor et al. 2001). As can be expected, the findings within environmental justice re-

search have therefore not only been highly variable, but contradictory as well. Early research regarding environmental justice often focused on a larger area such as the zip code, in order to examine distribution of risk (United Church of Christ 1987). Later investigations revealed the phenomenon of ecological fallacy, in which the heterogeneity of a particular area of study is often missed due to the area being too large (Anderton et al. 1994). In Anderton, et al. (1994), researchers utilized the census tract as the area of analysis in an attempt to capture the heterogeneity present within the study area. In an attempt to reduce the risk of ecological fallacy for this project we used a smaller area of analysis, the U.S. Census block group, which was the smallest area at which we could still obtain critical Census data. The challenge with selection of area of analysis for this project as with all environmental justice research is the development of a model which efficiently and effectively characterizes any disproportionate amount of environmental risk endured by any particular segment(s) of the population.

11.3 GIScience: GIS, RS & GWR

GIS has established itself as a tool well-suited for spatial analysis of environmental quality investigations, such as assessing questions of environmental justice. As environmental justice examines the geographical distribution of both status and risk, the benefits of using GIS are apparent. With state and federal government agencies realizing the importance of geographical data, a wealth of information has become available, including the locations of various facilities or sites which have been subjects of government enforcement. Such location data has proven useful when assessing questions of residential proximity to environmental risk or the siting of various facilities known or perceived to create environmental risk. Remote sensing technologies have also proven their effectiveness at revealing relationships perhaps otherwise not seen, such as that of quality of life and vegetation (Gatrell and Jensen 2002). In general, data gathered using remote sensing software can be combined with GIS data for effective modeling of environmental issues (Longley 2002). The combination of these technologies is what has been used here in order to compare data gathered by both GIS and remote sensing technologies.

In addition to the combination of GIS and remote sensing, another aspect of this research, which is discussed in greater detail later in this chapter, is the challenge of effectively representing the statistical interactions of risk

and status across the study area. To address this issue we used, in part, the statistical technique known as geographically weighted regression (GWR). Whereas standard regression provides global statistics implying uniformity across space, GWR effectively calculates local statistics at regression points across a study area, which aids in visualization of phenomena (Fotheringham, et al. 2002). Given the variability in environmental justice research results, with all sizes of study areas considered, GWR is an important tool in examining the variation of environmental risk across and throughout the U.S. Census Block Groups. Indeed, GWR may unlock heretofore unseen relationships and/or problematize existing assumptions. The following pages will provide further insight into the uses of GIS, remote sensing, and GWR to investigate issues which lay at the intersection of humans and their environment.

11.4 Data and Methods

The primary objective of the investigation discussed here was to assess the relative efficacy of both environmental quality data and a normalized difference vegetation index (NDVI) as metrics of socioeconomic conditions. To follow is a discussion of the methods used, including the environmental quality data, socioeconomic variables, and statistical techniques, as well as a discussion of the creation of the NDVI for the study area, Vigo County, Indiana.

11.4.1 Environmental Data Sets

The United States Environmental Protection Agency (U.S. EPA) has required reporting of certain information under the guidance of environmental regulations for several decades. This information has provided extensive data sets for research relating to environmental quality. The first data set used for this investigation was the EPA's Toxics Release Inventory Program (TRI), which includes information regarding reported releases from regulated facilities throughout the United States. In particular, releases to air, soil, and surface water were used by first asking whether there has been a release, answered with a yes or no, and then adding the amounts released (air, soil, and water) to make one reported number or quantity. In this way, there was no differentiation between routes of release. Rather, the total amount of released contaminants from each facility or site is used. By not parceling out the release information by medium, we avoided an investigative slippery slope regarding the route of release,

which leads to a consideration of the medium, meteorological, and hydrological factors.

Treatment, storage, and disposal facilities (TSDFs) databases were the second environmental quality data source used. TSDFs are regulated under the EPA's Resource Conservation and Recovery Act (RCRA), which in part, was designed to monitor the flow of hazardous waste from generation through to the time of disposal, a process commonly referred to as "cradle to grave".

The third environmental quality data set included the locations of Superfund sites within Vigo County, Indiana. This data consists of sites that are currently on the U.S. EPA's National Priorities List (NPL). Sites are placed on the NPL after regulatory officials investigate each site by following the Superfund cleanup process, beginning with notification to EPA of possible releases of hazardous substances. After each site is investigated it is either designated as needed no further remedial action or it is proposed for placement on the NPL.

TRI, TSDF, and Superfund data are location-based in their application to environmental justice research. The proximity of such facilities to particular communities or segments thereof is interpreted by many researchers as an indication of environmental risk, usually disproportionately distributed among the study area population. The TRI data was acquired from the U.S. EPA via its online data download library. The information is provided in the form of ESRI shapefiles and associated files, which was imported into ESRI's ArcMap software for display and analysis. TSDF data and Superfund site data were acquired from the Indiana Department of Environmental Management via the Indiana Geological Survey's online GIS data download library.

The fourth data set used provides levels of the metal lead (Pb) found in the blood of children within Vigo County, provided in the form of number of children within each zip code whose blood-lead levels were above a previously set criteria level. This data set could not only be a potential indicator of environmental quality, but it also represents actual human exposure to environmental contamination, as opposed to the other environmental quality data sets used here, which reflect a potential for environmental risk. Blood-lead level data were acquired from the Vigo County Health Department in the form of a hard-copy spreadsheet, with the data then being entered into a computer-based database and imported into ESRI's ArcMap software. The data set provides each zip code, as opposed to census block

group, where an elevated level was revealed. In order to use this data at the block group level, the mean blood-lead level of results for those zip codes affected was assigned for each block group within that zip code.

11.4.2 NDVI

A NDVI map was created with the use of the remote sensing software ERDAS Imagine 8.7 (ERDAS) and a satellite-produced image of Vigo County. The satellite image was produced by Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) using the Terra satellite, yielding a spatial resolution of 15-meters in the near-infrared and visible portions of the electromagnetic spectrum. Within ERDAS a NDVI was calculated by incorporating the near-infrared and red channels into the following formula:

$$\frac{\text{Near-infrared} - \text{red}}{\text{Near-infrared} + \text{red}}$$

NDVI is based upon the principle that the red (visible) portion of the electromagnetic spectrum is highly absorbed by chlorophyll present within plants or vegetation, while the near-infrared energy is reflected at high levels by a plant's mesophyll leaf structure (Tucker 1979). The calculated vegetation index then indicates the relative strength or reflectance of vegetation throughout the satellite image of the study area. A higher NDVI value indicates a more robust presence of vegetation. NDVI is unique in that it normalizes the various reflectance values by converting them to a value between –1 and 1 for each pixel in the image, with –1 representing no vegetation and 1 indicating robust vegetation. This investigation examines a NDVI of Vigo County, Indiana to determine its efficacy as a metric for socioeconomic status, and then compares the resulting capacity to that of the environmental quality data. Specifically, NDVI variables used in this analysis were the following:

Standard deviation of NDVI values within a block group;
Minimum NDVI value observed within a block group;
Maximum NDVI value observed within a block group;
Range of NDVI values observed within a block group;
Interaction of NDVI with population density; and,
Mean NDVI value

These values were assigned to each of the census block group polygons within the spatial database.

11.4.3 Socioeconomic/Demographic Characteristics

Specific socioeconomic and demographic variables selected for use in this investigation have been applied in much of the earlier research regarding environmental justice. Such variables are often used as indicators of socioeconomic status. The following socioeconomic and demographic variables were acquired from the U.S. Census Bureau's online data library for the year 2000 and integrated into this analysis as indicators of socioeconomic status: (1) Median Household Income; and (2) Median Household Value.

11.5 Methods

This investigation uses three approaches: correlation, weighted least squares regression, and geographically weighted regression. Correlation—Pearson's R—was used to explore the relationships between variables and the significance of these variables. Using the Pearson's R results as a guide, weighted least squares regression models were tested using both enter and step-wise approaches. The weighted least squares regression was performed using population density as the weighting variable.

Geographically weighted regression was used as standard regression statistical techniques often treat phenomena as occurring equally across a study area. As Fotheringham, et al. (2002) discussed, spatial data often exhibit what has been termed spatial nonstationarity, or the nonuniform distribution of spatial information. The benefit of GWR in geographical research is that it accounts for unique characteristics of spatial data by calculating the necessary statistical measures at each point in the study area, which provides individual level or point-unique statistical information, allowing a researcher to identify disparities in the spatial distribution of various phenomena. GWR served this research well given that previous research has demonstrated the spatial nonstationarity of disproportionate environmental risk (Mennis and Jordan 2005).

The model for this investigation included the following variables analyzed through OLS regression and GWR, as well as analysis using Pearson's Correlation between the Socioeconomic and environmental metrics.

The variables used were:

Median Household Income
Median Household Value
U.S. EPA Toxics Release Inventory
RCRA TSDFs
Superfund Sites
Child blood-Lead levels (BLL) by Zip code
NDVI Minimum
NDVI Maximum
NDVI Mean
NDVI Range
NDVI Standard Deviation
Population Density

11.5.1 Interaction terms

In addition to the variables listed above, an interaction term was created using the expansion method (Casetti 1972, Gatrell and Bierly 2002, Jensen, et al. 2005). The expansion method developed by Emilio Casetti was an early challenge to the existing statistical paradigm that assumed spatial relationships are constant across a study area (Gatrell, J., Chapter 5 of Jensen, et al. 2005). Casetti (1972) also attributed the nonstationarity of spatial phenomena to the interaction of terms across space. We relied upon this interaction of terms as we created a model containing a multiplied interaction of NDVI data with observed population density for each block group within the study area. Population density was used as it has been shown to be effective when modeling environmental parameters in an urban environment.

11.6 The Models

Below the models are presented. The models presented were subjected to OLS, stepwise, and GWR. The study models are:

$$Y = \beta_0 + \beta TR(u,v) + \beta TF(u,v) + \beta B(u,v) + \beta S(u,v) + \beta Sd\,(u,v) + \beta Min(u,v) + \beta Mx(u,v) + \beta A(u,v) + \beta R(u,v) + \beta I(u,v) + \varepsilon(u,v)$$

where:

Y is the dependent variable (socioeconomic status), in this

case, either median household income or median household value;
β_0 is the constant;
TR is the U.S. EPA's Toxics Release Inventory data for Vigo County;
TF is IDEM Treatment, Storage, and Disposal Facilities;
B is the blood-lead level in children for Vigo County during 2000-2005;
S is the Superfund facility data;
Sd is the standard deviation of the NDVI values;
Min represents the minimum NDVI value;
Mx is the maximum NDVI value;
A is the mean NDVI value ;
R is the range of NDVI values;
I refers to interaction terms using population density and A; and,
ε refers to the statistical noise assumed to be present in the calculation;

The formula for GWR is the following:

$$y_i = \beta_0(u_i,v_i) + \Sigma_k\beta_k(u_i,v_i)x_{ik} + \varepsilon_i$$

where:

y_i is the dependent variable at location I;
β_0 is an independent variable;
(u_i,v_i) is the coordinate location for the ith point;
$\beta_k(u_i,v_i)$ is the function continuously measuring parameter values at each point I; and,
ε_i is the noise associated with each point i
(Fotheringham, et al. 2002).

11.7 Results

A challenge to effective statistical analysis of geographical relationships is spatial nonstationarity, or the discontinuity of relationships among and between geographical cases or phenomena throughout a study area (Fotheringham et al. (2002). While the process of WLS does capture variability across space as driven by varying population densities, "global" WLS does so based on discrete points rather than across a continuous surface (Fotheringham, et al. 2002). For this reason, GWR calculates local statistics, spe-

cifically local r-square values, to determine the model performance in "place" and across "space". In this paper, we use local r-square values derived from GWR to visualize, or map, the spatial dynamics and model performance across the study area. WLS and GWR 3.0 yielded both global and local coefficients of determination. WLS regression was performed on the data, using population density as the weighting variable, in order to evaluate the Pearson's correlation values. We first examined the distribution of the relationships between socioeconomic conditions and environmental quality data using WLS regression. WLS indicated a very weak relationship between both median household income and median household value and environmental disamenities. WLS was able to discern variability in that relationship across space within Vigo County, but the overall relationships were quite weak. Local r-square values generated within the GWR software were mapped to provide a visual reflection of the data (Fig. 2 and 3). GWR was used to determine whether there was spatial nonstationarity among the relationship(s) between socioeconomic conditions and environmental disamenities.

When examining median household income using WLS, all four of the environmental quality variables received correlation values of .05 or lower, with two of the four having negative values. The most closely correlated variables to median household income were the NDVI maximum value (.412) and the NDVI mean value (.440). When regressing the median household value data against the independent variables, the standard deviation of the NDVI as well as the NDVI mean value displayed the strongest correlation to household value at .320 and .283, respectively (See Tables 1 and 2). With such a drastic disparity between the roles of environmental quality variables and NDVI variables, it is apparent that within Vigo County, the geographical distribution of traditional environmental quality indicators do not statistically account for the observed variation in and/or spatial arrangement of median household income and property valu**es.**

Table 1. Diagnostics (Enter Method)

	Household Income	Household Value
Constant	-38900.584	-40424.255
	(-2.166)**	(-.870)
TR	-.002	-.003
	(.718)	(-.418)
TF	-12504.292	-105772.454
	(-.366)	(-1.198)
B	475.390	1225.025
	(1.424)	(1.419)
S	324.068	66288.976
	(.023)	(1.802)*
Min	33476.614	---
	(1.346)	---
Mx	115678.429	30372.046
	(1.130)	(.105)
Sd	28544.582	818233.782
	(.163)	(1.805)*
A	133434.982	372417.923
	(1.495)	(1.613)
I	1.001	-30.097
	(.035)	(-.410)
R	---	-93881.855
	---	(-1.460)
R-Square	**.234**	**.195**
F-Statistic	**3.662**	**2.911**

***** Indicates the variable is significant at the .10 level.
****** Indicates the variable is significant at the .05 level.

The step-wise approach provided insight as to which model may be considered more “elegant”. In this case, it was the mean NDVI value which created a more elegant model when regressed against median household income. Regarding median household value, the standard deviation NDVI value contributed most to the performance of the model.

Table 2. Diagnostics (Stepwise Method)

	Household Income	Household Value
Constant	-10715.333	21478.996
	(-1.154)	(1.249)
Sd	---	783105.761
	---	(3.633)***
A	230062.271	---
	(5.274)***	---
R-Square	**.193**	**.102**
F-Statistic	**27.811**	**13.201**

******* Indicates the variable is significant at the .01 level.

As illustrated in Table 1, median household income was revealed as having the stronger model, with an r-square value of .234, as compared to that of median household value (.195). In this sense, income is more strongly related to the independent variables. Among the independent variables used here, NDVI variables possessed a stronger relationship to both income and household values, with a significantly weaker showing for the environmental quality variables. These figures provide for a discussion of the driving forces behind the phenomenon known as environmental justice. The environmental quality variables (TRI, TSDF, Superfund, and BLL) do not appear to be significantly related to the distribution of such indicators of socioeconomic status as income and household value. However, what this may indicate is less a case of environmental risk seeking out poor populations than wealthier populations seeking amenities, such as vegetation or "greenness".

When the local r-square values for median household income, generated by GWR 3.0, were mapped (Fig. 2), there was a clear relationship within block groups of the urban core of Terre Haute between median household income and the independent variables used. This was not surprising as the urban center of Terre Haute contains block groups with some of the lowest income population. In addition, the model performed strongly in the northwest and southeast block groups, which are predominantly rural areas.

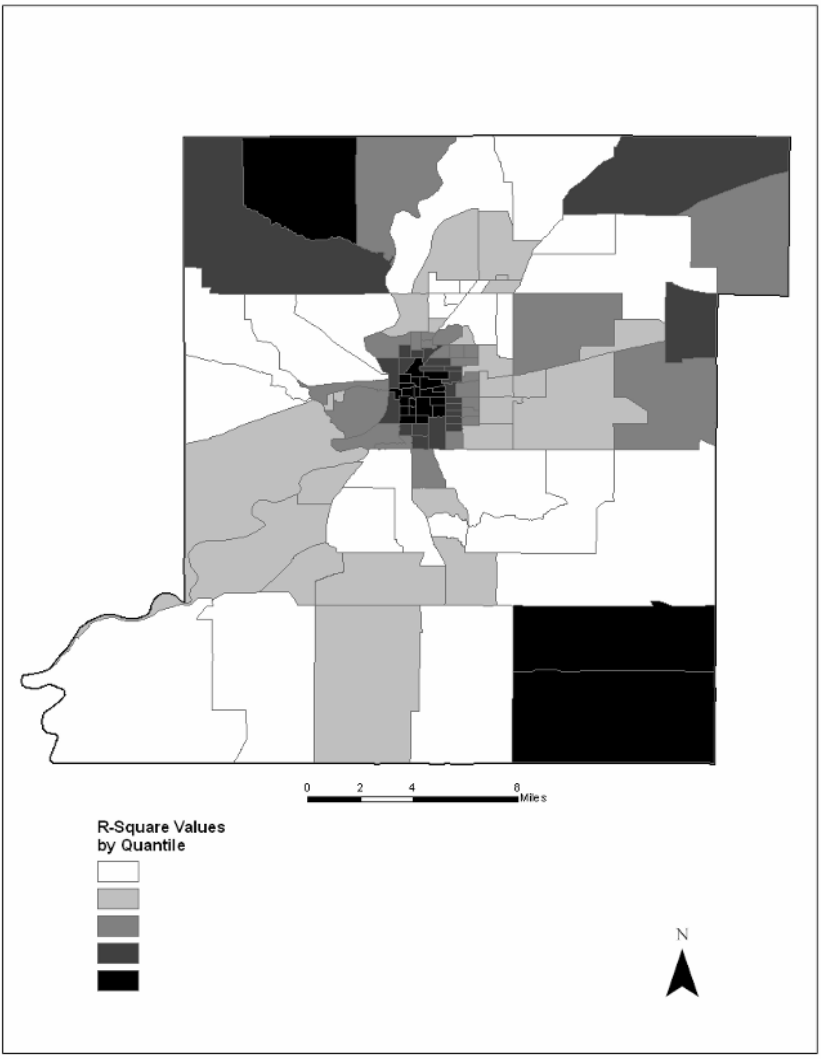

Fig. 2. Household income local r-square values by quantile.

A somewhat similar pattern emerged within the central (urban) and north central/northwest block groups when examining the r-square values for median household value. A strong relationship was observed in the central and northwest areas, but the difference between the mapped results of the two variables was the lack of any significant local r-square values for household value in the southeast corner of the study area (Fig. 3).

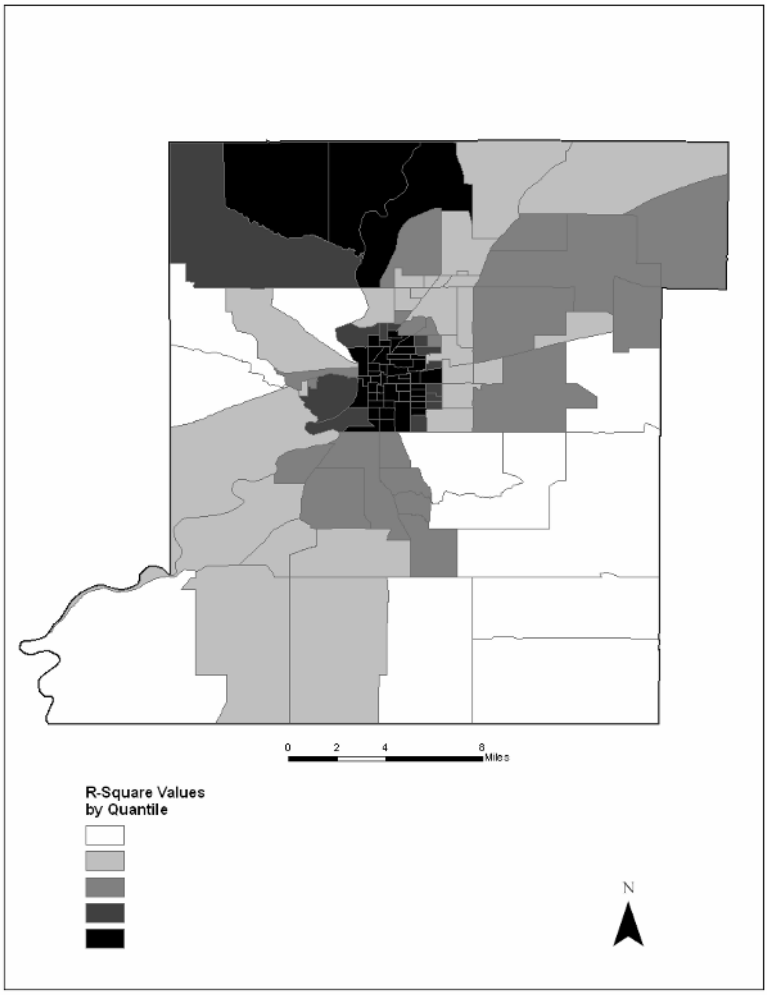

Fig. 3. Household value local r-square values by quantile.

11.8 Discussion

The results—while consistent with the earlier greenness research of Gatrell and Jensen and Jensen et al (2004, 2005)—suggest the environmental justice literature's focus on environmental disamenities may capture only part

of the complex interactions that occur within and between social and natural systems in urban environments. That is to say, the basic assertion that the co-location of marginalized groups and environmental disamenities, represents only part of the complete picture. Rather, as this study suggests, the geography of environmental disamenities and socio-economic variables does little to explain the implied relationships between negative externalities and class and race. Instead, the urban environmental geography of class—and perhaps race, too—may be better understood within the context of access to environmental amenities as determined by key proxy variables, such as NDVI. Moreover, the unique geography of Vigo County suggests the distribution of negative externalities are only co-incident—and not necessarily correlated—as "risky sites" occur in a wide range of socio-economic contexts. To that end, the study suggests further research is needed to understand the observed spatial disconnect between the urban geography of amenities and disamenities.

References

Anderton, Douglas L., A. Anderson, J. Oakes, and M. Fraser. 1994. "Environmental Equity: The Demographics of Dumping." *Demography* 31: 229-48.

Buzzelli, Michael, M. Jerrett, R. Burnett, and N. Finklestein. 2003. Spatiotemporal perspectives on air pollution and environmental justice in Hamilton, Canada, 1985-1996. *Annals of the AAG* 93(3): 557-73.

Casetti, E. 1972. Generating models by the expansion method: applications to geographic research. *Geographical Analysis* 4:81-91.

Cutter, S., M. E. Hodgson, and K. Dow, 2001. Subsidized inequities: The spatial patterning of environmental risks and federally assisted housing. *Urban Geography* 22 (1): 29-53.

Fotheringham, A. Stewart, C. Brunsdon, and M. Charlton, 2002. "Geographically Weighted Regression: The Analysis of Spatially Varying Relationships." John Wiley & Sons, Ltd. West Sussex.

Gatrell, J.D. and Bierly, G.D. 2002. Weather and voter turnout: Kentucky primary and general elections, 1990-2000. *Southeastern Geographer* Vol. XXXXII (1): 1-21.

Gatrell, J. D. and R. R. Jensen. 2002. Growth through greening: developing and assessing alternative economic development programs. *Applied Geography* 22:331-350.

Jensen, R., J. Gatrell, J. Boulton, and B. Harper. 2004. Using Remote Sensing and Geographic Information Systems to Study Urban Quality of Life and Urban Forest Amenities. Ecology and Society 9(5): 5. http://www.ecologyandsociety.org/vol9/iss5/art5/

Jensen, R., J. Gatrell, and D. McLean (eds) 2005. "Geo-Spatial Technologies in Urban Environments." Heidelberg, Germany: Springer-Verlag.

Margai, Florence L., 2001. Health Risks and environmental inequity: a geographical analysis of accidental releases of hazardous materials. *The Professional Geographer* 53(3): 422-34.

Mennis, J. 2002. Using geographic information systems to create and analyze statistical surfaces of population and risk for environmental justice analysis. *Social Science Quarterly* 84: 281-297.

Mennis, J. and Lisa Jordan. 2005. The Distribution of Environmental Equity; Exploring Spatial Nonstationarity in Multivariate Models of Air Toxic Releases. *Annals of the Association of American Geographers*, 95(2), 2005, pp. 249-268.

Mohai, Paul, and Bunyan Bryant, eds., *Race and the Incidence of Environmental Hazards: A Time for Discourse*. Boulder, Colo.; Westview.

Pastor, M., J. Sadd, and J. Hipp. 2001. Which came first? Toxic facilities, minority move-in, and environmental justice. *Journal of Urban Affairs* 23: 1-21.

Pulido, L. 2000. Rethinking environmental racism: white privilege and urban development in southern California. *Annals of the Association of American Geographers* 90: 12-40.

Tucker, C.J. 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of the Environment* 8: 127-150.

United Church of Christ Commission for Racial Justice. 1987. *Toxic Wastes and Race in the United States: A National Report on the Racial and Socioeconomic Characteristics of Communities with Hazardous Waste Sites.* New York: Public Data Access.

www.epa.gov/tri/ U.S. EPA Toxics Release Inventory Program web page. 2006.

www.epa.gov/compliance/basics/ej.html U.S. EPA Office of Environmental Justice. 2006.

12 Geotechnologies, Public Policy, & Practical Applications

Jay D. Gatrell, Department of Geography, Geology & Anthropology, Indiana State University, Terre Haute, IN

Ryan R. Jensen, Department of Geography, Geology & Anthropology, Indiana State University, Terre Haute, IN

Daniel D. McLean, Department of Recreation and Sport Management, Indiana State University, Terre Haute, IN

12.1 Technologies & Methods

The collection of papers in this volume includes standard urban applications of GIS and multi-spectral remote sensing—as well as applications that demonstrate the utility of emerging (and increasingly accessible) technologies, such as LIDAR (J. Jensen et al.). The inclusion of LIDAR examples reflects the growing use of LIDAR technologies in urban planning and civil engineering. The precision of this technology makes it ideal for use in complex urban systems. Similarly, hyper-spectral sensors offer high resolution options to assess a variety of urban surfaces. With respect to both technologies, falling costs associated with both LIDAR and hyper-spectral data are making these increasingly viable methods for urban data collection.

In addition to new technologies, this collection also introduced alternate methodological frameworks. Like the first collection, several papers used the expansion method and in the case of Fuller et al. included it within a robust Geographically Weighted Regression (GWR) analysis. The use of GWR in complex urban systems may enable students to understand not only the interaction between human and physical systems—but also how these interactions explicitly vary across space. More over, the use of

GWR in a policy space has the potential to be especially useful insofar as it demonstrates that public policy can be (or perhaps should be) tailored to various local conditions. Similarly, Spiker and Warner examine one of the ways in which geography complicates data acquisition and analysis. Their paper on spatial autocorrelation underscores one of the continuing methodological issues that confront urban analysis and thereby frustrates the ability of research to more fully inform the decisions of policy makers.

Finally, several papers including R. Jensen and Hardin demonstrate the efficacy of using artificial neural networks in conjunction with geo-technical applications. In the case of assessing the urban canopy, regression analysis demonstrates the utility of the methods. More importantly, the paper underscores the potential for geo-technologies to be used as effective—more importantly cost effective—assessment tools.

12.2 Risk

In the previous collection, the relationship between environmental amenities and socio-demographics was explored using a variety of case studies. In this book, we have included a basic exploratory paper on intra-urban population (Hardin, Jackson and Shumway)—as well as contributions that explore more targeted questions associated with specific public policy. In Fuller et al., the authors examined the socio-spatial implications of poverty and race relative to the environmental quality of communities. Liu et al. use GIS and remote sensing technologies to unlock the geography of child physical activity within Indianapolis. In concert, these two papers begin to chart the important role geo-technologies can play in improving the everyday lives (and health) of urban citizens. Similarly, Hipple's real world analysis of urban risk and his unique perspective as a practitioner within a policy agency demonstrates the potential impact geo-technologies can and do have on shaping public policy. Similarly, Johnson explores the potential us-e of geo-technologies to develop strategies to mitigate urban medical risks. Finally, Gonser and Horn examine the risk of deer collisions on the urban fringe.

12.3 Planning

Several papers investigate the role of GIS and Remote Sensing as planning tools in the public sector or industry. These applications demonstrate that

GIS has become an important tool in both the public and private sectors. In particular, GIS and Remote Sensing have become essential tools relative to the mapping, assessment and management of urban change (Hardin, Jackson, and Otterstrom). Likewise, the exploratory analysis of census data illustrates the potential uses of GIS in the public sector (Hardin, Jackson and Shumway). Finally, Johnson's paper on investigating urban health considers the efficacy of developing participatory planning frameworks that utilize GIS. As these examples suggest, GIS/RS technologies will continue to evolve and impact decision making in urban settings.

In conclusion, the objective of this collection was to expand the catalog of real world applications, introduce new or novel methodological approaches, and to explore the full range of urban applications that exist. Insofar as this is a second collection in a series on this topic, the two combined collections represent only a small fraction of the many exciting possibilities that exist for the geo-technologies.

Index

A

Active Fire Mapping Program, 37
Advanced Spaceborne Thermal Emission and Reflection Radiometer
  ASTER, 97, 221
allometric population growth models, 51
ArcGIS, 16
ArcMap, 13, 19, 26
Artificial neural networks
  Applied to change detection, 157
AVHRR, 35

B

Biophyllia, 113
body mass indices, 123

C

cellular automata, 164, 165
Cellular automata (CA), 164
Census, 47–92
  Decennial Census of the US, 87
Central place theory, 51
Change detection, 145, 147
  Artificial neural networks and, 157
  Change vector analysis, 155
  Comparative studies of methods, 152
  Decision trees and, 159
  Econometric panel techniques and, 161
  IHS transform and, 160
  Image overlay and, 145
  Image ratioing and, 148
  Image regression and, 150
  Post-classification and, 146
  Principal components analysis and, 150
  Simple image differencing and, 146
  Sources of error in, 152
  Spectral mixture analysis, 156
  Spectral-temporal classification and, 146
charge couple device
  ccd, 9
color infrared (CIR) photography, 53
Component-method II procedures, 50
Culex pipiens, 110, 112

D

Decennial Census of the United States, 87
Deer-Vehicle Collisions, 177, 181
digital frame cameras, 9
DVC, 177, 178, 181, 183, 184, 186, 188, 192, 193, *See* Deer-Vehicle Collisions

E

environmental justice, 215, 217, 218, 219, 220, 222, 227, 229, 230, 231
Environmental justice, 217
Enhanced Thematic Mapper + ETM+, 35, 41

F

FireLine, 38, 44
flood
  flooding, 34
floodplain, 34, 40, 41, 42, 43, 44

G

Game of Life, 164
Geary's *c*, 186, 199, 200, 201, 202, 206, 211
geocoding, 125
Geographically weighted regression
  gwr, 218, 219, 222, 223, 224, 227
Getis-Ord G_i, 202
GIScience, 215, 218
Global Positioning System, 35

H

hazards
  hazard, 33
health, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118
Heat related illness, 115
Housing
  Aerial photography, 56, 60, 61, 63, 64, 66
  Dwelling count estimates, 56
  Dwelling identification, 52
  Photographic keys, 52, 55, 56
housing types, 54
housing-unit method, 50
Hurricane Katrina, 35
hyperthermia, 115

I

IDW. See Inverse Distance Weighting,
IKONOS, 34
Indian Resourcesat AwIFS, 35
Indiana, 215, 216, 219, 220, 221
Indianapolis Mapping and Geographic Information Service (IMAGIS), 125
Intensity-Hue-Saturation transform
  Applied to change detection, 160
interaction term, 223
Intraurban population estimation, 57, 68
Inverse Distance Weighting, 186

K

Kansas City, MO, 41
Kauth-Thomas, 147, 149

L

Landsat TM, 35, 38, 41
Landscape metrics, 162
Landtype
  and population estimation, 64, 65, 66
Leaf area index
  Urban remote sensing and, 94
Leaf area index (LAI), 93, 106
  Ceptometers and, 94
  Estimating with neural networks, 98
  Field measurement and, 95
  Gap-fraction analysis and, 95

LIDAR, 8, 10, 11, 22, 23, 24, 25, 26, 27, 30, 31, 35
local indicators of spatial autocorrelation, 201

M

Medical geography, 110
MODIS, 35, 36, 37, 45
Moran's *I*, 186, 187, 191, 199, 200, 201, 202, 206, 211
Morgantown, West Virginia, 202

N

NCI See neighborhood correlation
NDVI, 219, 221, 223, 224, 225, 226, 227, 230,
neighborhood correlation image analysis
  nci, 11
Neural Networks (artificial)
  Estimating LAI with, 98
No Free Lunch Theorem, 152
Normalized Difference Vegetation Index (NDVI)
  Formula for, 97
Normalized Differenced Vegetation Index (NDVI), 149
normalized differential vegetation index (NDVI), 129

O

obesity, 122
OrbView, 34

P

physical activity, 121, 122, 123, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139
planimetric base
  planimetric base map, 34
population density estimating, 51, 52, 60, 61, 62, 63, 64, 66, 68, 69, 70, 71, 72, 73, 74, 75, 76, 78, 79, 81, 82, 83, 84, 85
population estimation, 48
  dwelling identification and counting, 52
  landtype surrogates, 60, 64, 66
  Landtype surrogates, 60
  regression modeling of, 62, 65, 68, 70, 84
  thermal temperature and, 73, 74, 75, 81, 82, 83, 84
Population Estimation
  Pixel-based approaches, 66
PPGIS, 109, 110, 114, 116, 117, 118
Principal Component Analysis, 149

Q

quality of life, 89
QuickBird, 34, 203, 205, 212

R

Ratio-correlation procedures, 49
risk, 33

S

Salt Lake City
  population estimation study, 57, 58, 76, 77, 78
semivariance, 199, 200, 201, 205
semivariogram, 200
SLEUTH, 164
spatial autocorrelation, 198, 199, 200, 201, 211, 213
spectral-temporal classification, 146, 155
St. Louis, MO, 41
suburban sprawl, 141
Summer Health Assessment Program Education (SHAPE), 123
supervised classification, 157, 158

T

Terre Haute, Indiana, 215
  LAI study and, 95

U

U.S. Census, 41
Ugly Duckling Theorem, 153
Upper Mississippi and Missouri River, 39, 40, 44
Urban forest, 93
Urban growth, 141
  Modeling with cellular automata, 164
  Planning policy and, 144,
  Spatial metric growth signatures and, 163
Urban growth measurement
  Landscape metrics and, 162
urbanization, 47

V

Vigo County, IN, 188, 192, 196
VIS model of urban surfaces, 77, 78

W

White-tailed deer, 179, 181, 183, 193, 195
wildfire mapping, 37