

HOTSPOTS, UNDERREPORTING, AND DYNAMIC SPACE-TIME INFLUENCES OF
WILDLIFE-VEHICLE COLLISIONS

By

Nathan P. Snow

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degrees of

Fisheries and Wildlife – Doctor of Philosophy
Ecology, Evolutionary Biology and Behavior – Dual Major

2014

ABSTRACT

HOTSPOTS, UNDERREPORTING, AND DYNAMIC SPACE-TIME INFLUENCES OF WILDLIFE-VEHICLE COLLISIONS

By

Nathan P. Snow

Vehicular collisions with wildlife are one of the most widespread and persistent human-wildlife conflicts that exist throughout the United States. An estimated 1–2 million wildlife-vehicle collisions (WVCs) occur each year, and that number is increasing annually. The total annual cost associated with WVCs is estimated to be >8.3 billion dollars, as well as the loss of millions of animals. Despite the magnitude of consequences from WVCs, relatively few options exist for reducing the frequency of collisions. A number of reasons can explain the shortage of options. First, identifying the most critical locations for mitigation is not straightforward. Hotspots of WVC locations are loosely defined, even though hotspots provide the best opportunity for cost-effective mitigation. Second, reporting of WVCs is inconsistent, resulting in incomplete information for studies that analyze where collisions occur. Inferences from these studies could be unreliable because of incomplete data. Finally, there is a lack of knowledge regarding large-scale and long-term trends in the frequencies of WVCs. Larger geographic and temporal studies are needed to understand the environmental influences for those trends. My overall objective was to enhance the current approaches for examining WVCs and provide more reliable inferences for reducing collisions. In Chapter 1, I address the issue of inconsistency and subjectivity in delineating hotspots of WVCs. In Chapter 2, I address the issue of sensitivity in statistical inferences from underreporting or studies of WVCs. In Chapter 3, I address the issue of understanding large-scale, dynamic processes that influence white-tailed deer (*Odocoileus virginianus*)-vehicle collisions (DVCs) through space and time.

The collective works in these chapters contribute 3 primary conclusions for better understanding the influences of WVCs. First, the landscape can be used to objectively delineate hotspots. This new approach indicates that hotspots are larger than previously reported. Second, analyses of the influences of WVCs are highly robust to underreporting likely because WVCs occur in highly predictable patterns (i.e., hotspots). Therefore, relatively few reports are required for reliably understanding the environmental influences on where hotspots occur. Third, the large-scale, ecological drivers of DVCs through are related to suburbanization. The suburb effect consists of a unique combination intermediate to high traffic volume, high abundances of deer, and a highly fragmented landscape with high proportions of croplands. The suburb effect did not change through time, indicating high spatiotemporal predictability for DVCs.

These collective works suggest large hotspots associated with suburban landscapes account for the highest frequencies of collisions, therefore these locations should be targeted for mitigation. Identifying the most critical locations to mitigate can be accomplished with relatively few reports of collisions if collected in a consistent manner. Managers should consider investing in long-term mitigation strategies (i.e., underpasses) to reduce WVCs for many years, because the ecological drivers of hotspots do not appear to change.

ACKNOWLEDGMENTS

I thank my advisor, Dr. Porter, for promoting self-discovery with timely guidance. I also thank my committee members, Drs. Finley, Roloff, Rudolph, and Winterstein. Dr. Finley patiently and graciously taught me space-time analyses. Dr. Roloff and Dr. Winterstein provided excellent suggestions regarding wildlife relationships and habitat. Dr. Rudolph helped focus the application of my research.

I thank the faculty and students in the Quantitative Wildlife Laboratory that provided insight and support, especially Dr. Williams and Ms. Jarzyna. This research was funded in part by the Boone and Crockett Club, the Michigan Department of Natural Resources, the Alces Journal, the Michigan Chapter of Safari Club International, and the Ambrose Pattullo Fund. I also thank Michigan State University for support.

Lastly, I am grateful for the steady love and interest from my family and friends. Their support was motivational.

TABLE OF CONTENTS

LIST OF TABLES.....	vii
LIST OF FIGURES.....	viii
PROLOGUE.....	1
A LANDSCAPE-BASED APPROACH FOR DELINEATING HOTSPOTS OF WILDLIFE- VEHICLE COLLISIONS.....	6
ABSTRACT.....	6
2.1 INTRODUCTION.....	7
2.2 METHODS.....	10
2.2.1 Study area.....	10
2.2.2 Data collection.....	13
2.2.3 Landscape metrics.....	14
2.2.4 Data analysis.....	17
2.3 RESULTS.....	21
2.4 DISCUSSION.....	27
2.5 ACKNOWLEDGEMENTS.....	30
UNDERREPORTING AND ASSESSING WILDLIFE-VEHICLE COLLISIONS WITH LARGE UNGULATES.....	31
ABSTRACT.....	31
3.1 INTRODUCTION.....	32
3.2 METHODS.....	34
3.2.1 Study area.....	34
3.2.2 Data collection.....	36
3.2.3 Study design and data analysis.....	39
3.2.4 Evaluating sensitivity in precision.....	40
3.2.5 Evaluating sensitivity to biased reporting.....	41
3.2.6 Evaluating sensitivity in model prediction.....	42
3.3 RESULTS.....	43
3.4 DISCUSSION.....	49
3.5 ACKNOWLEDGEMENTS.....	53
DYNAMIC SPACE-TIME ANALYSIS OF WHITE-TAILED DEER-VEHICLE COLLISIONS THROUGHOUT THE MIDWEST UNITED STATES.....	54
ABSTRACT.....	54
4.1 INTRODUCTION.....	55
4.2 METHODS.....	57
4.2.1 Study area.....	57

4.2.1 Data collection.....	60
4.2.1 Study design and data analysis.....	62
4.3 RESULTS.....	64
4.4 DISCUSSION.....	71
4.5 MANAGEMENT IMPLICATIONS.....	74
4.6 ACKNOWLEDGEMENTS.....	75
EPILOGUE.....	77
APPENDIX.....	80
LITERATURE CITED.....	87

LIST OF TABLES

Table 2.1 Reclassified land-cover and land-use types for 3 species of wildlife: A) island foxes on San Clemente Island, CA, USA (2006–2010), B) white-tailed deer in Onondaga County, NY, USA (2005–2006), and C) moose in western Maine, USA (1993–2010).....	14
Table 2.2 Metrics for data analysis used in kernel density estimation (KDE) for 3 species of wildlife: A) island foxes on San Clemente Island, CA, USA (2006–2010), B) white-tailed deer in Onondaga County, NY, USA (2005–2006), and C) moose in western Maine, USA (1993–2010).....	16
Table 2.3 Number of hotspots and length of roads delineated using a landscape-based approach and currently used methods for 3 species of wildlife: A) island foxes on San Clemente Island, CA, USA (2006–2010), B) white-tailed deer in Onondaga County, NY, USA (2005–2006), and C) moose in western Maine, USA (1993–2010 and 2008–2010).....	22
Table 3.1 Reclassified land-cover and land-use types for 2 study species: A) white-tailed deer in central Illinois, USA (2010), and B) moose in western Maine, USA (2000–2010).....	37
Table 3.2 Averaged values for environmental variables in counties used to examine influences on the probabilities of deer-vehicle collisions in central Illinois (2010) and moose-vehicle collisions in western Maine (2000–2010). Biased counties represented counties in which the rates of reporting for collisions were simulated to be lower than unbiased counties.....	44
Table 4.1 Reclassified land-cover and land-use types for 3 eco-zones in the Midwest United States from the 2001 and 2006 National Land Cover Databases.....	62
Table A.1 Description and sources of data used to examine the dynamic, space-time influences of deer-vehicle collisions (DVCs) throughout the Midwest, USA during 2000–2011.....	84

LIST OF FIGURES

Figure 1.1 Map generated by State Farm Mutual Automobile Insurance Company ©2010 showing the likelihood of a motorist being involved in a collision with a deer during 2011.....2

Figure 2.1 Study areas, roads, and locations of wildlife-vehicle collisions for: A) island foxes on San Clemente Island, CA, USA (2006–2010), B) white-tailed deer in Onondaga County, NY, USA (2005–2006), and C) moose in Western Mountains biophysical region, ME, USA (1993–2010).....12

Figure 2.2 Conceptual flowchart showing a process for delineating non-subjective hotspots of wildlife-vehicle collisions (WVC) using kernel-density estimation (KDEs). Step 1 is to gather accurate locations of WVCs. Step 2 is to generate a candidate set of grouping scenarios using KDEs with unique combinations of bandwidth and isopleth values. Additionally, calculate landscape metrics for each WVC with user-specified extent(s) of the landscape. Step 3 is to associate the landscape metrics to each grouping scenario, and then calculate the variance for each metric among groups of WVCs within each grouping scenario. Step 4 is to identify the grouping scenario with the greatest amount to variation.....19

Figure 2.3 Example delineations of San Clemente Island, CA, USA (2006–2010) fox-vehicle collision hotspots calculated with kernel density estimation. Each bandwidth and isopleth combination produced a unique delineation as part of an overall candidate set of potential delineations. This figure shows 10 of the 285 potential delineations that were calculated.....20

Figure 2.4 Maximum variance values (standardized) for each landscape metric among hotspots of wildlife-vehicle collisions for: A) island foxes on San Clemente Island, CA, USA (2006–2010), B) white-tailed deer in Onondaga County, NY, USA (2005–2006), C) moose in Western Mountains biophysical region, ME, USA (1993–2010), and D) moose 2-years subset (2008–2010).....25

Figure 2.5 Peaks in variation were identified at: A) 260 m bandwidth and 55% isopleth for the proportion of shrub landscape metric for island foxes on San Clemente Island, CA, USA (2006–2010), B) 1,600 m bandwidth and 50% isopleth for the proportion of forest landscape metric for white-tailed deer in Onondaga County, NY, USA (2005–2006), and C) 2,400 or 2,500 m and 20% isopleth for the interspersed-juxtaposition index landscape metric for the 2-years subset of moose in the Western Mountains biophysical region, ME, USA (2008–2010).....26

Figure 3.1 Study area locations of reported A) deer-vehicle collisions in central Illinois, USA during 2011, and B) moose-vehicle collisions in western Maine, USA during 2000–2011.....35

Figure 3.2 Average parameter estimates and 95% confidence intervals from 10,000 Monte Carlo simulations at different levels of underreporting for deer-vehicle collisions in central Illinois, USA, during 2011. The model with good predictive capability included: TRAFFIC = annual average daily traffic, ABUNDANCE = abundance of deer per county indexed by antlered

harvest, P_AG = proportion of agriculture land-cover within 800 m buffer, P_FOR = proportion of forest land-cover within 800 m buffer, CWED = contrast-weighted edge density within 800 m buffer, CONTAGION = contagion index within 800 m buffer. The model with poor predictive capability included: REGISTERED_VEH = the number of registered vehicles per county, ABUNDANCE, and P_WAT = proportion of water land-cover within 800 m buffer.....46

Figure 3.3 Average parameter estimates and 95% confidence intervals from 10,000 Monte Carlo simulations of logistic regression analyses including spatial biases in reporting for different levels of underreporting for deer-vehicle collisions in the most rural counties in central Illinois, USA, during 2011. The model with good predictive capability included: TRAFFIC = annual average daily traffic, ABUNDANCE = abundance of deer per county indexed by antlered harvest, P_AG = proportion of agriculture land-cover within 800 m buffer, P_FOR = proportion of forest land-cover within 800 m buffer, CWED = contrast-weighted edge density within 800 m buffer, CONTAGION = contagion index within 800 m buffer. The model with poor predictive capability included: REGISTERED_VEH = the number of registered vehicles per county, ABUNDANCE, and P_WAT = proportion of water land-cover within 800 m buffer.....47

Figure 3.4 Average receiver operating characteristic function (AUC) and 95% confidence intervals for logistic regression analyses including spatial biases in reporting for the least urban counties for deer-vehicle collisions in central Illinois, USA, during 2011 and moose-vehicle collisions in western Maine, USA during 2000–2011. The AUC values were calculated from 10,000 Monte Carlo simulations at different levels of underreporting.....48

Figure 4.1 Study area for examination of dynamic, space-time influences of deer-vehicle collisions at the county level throughout the Midwest, USA during 2000–2011.....59

Figure 4.2 Estimates of regression coefficients and 95% CIs from dynamic models for examining the influences of environmental variables on the frequencies of deer-vehicle collisions (DVCs) at a county level throughout the Midwest, USA during 2000–2011. CONTAGION = an index of fragmentation among land covers per county per year where lower values represent more fragmented landscapes.....67

Figure 4.3 Effect plots (medians and 95% CIs) for the interaction term of CONTAGION x ABUNDANCE for examining the influence on the frequency of deer-vehicle collisions (DVCs) in 3 eco-zones in the Midwest, USA during 2011. CONTAGION = an index of fragmentation among land covers per county per year where lower values represent more fragmented landscapes. The density of observed data points are plotted to indicate levels of confidence for the predicted surface. Red = high density of observed points. Blue = low density of observed points. No color = extrapolated effects.....68

Figure 4.4 Effect plots (medians and 95% CIs) for environmental variables for examining the influences on the frequency of deer-vehicle collisions (DVCs) in 3 eco-zones in the Midwest, USA during 2011. CONTAGION = an index of fragmentation among land covers per county per year where lower values represent more fragmented landscapes.....69

Figure 4.5 Predicted vs. observed number of deer-vehicle collisions (DVCs) for 10% of withheld data for model validation in 3 eco-zones in the Midwest, USA during 2011.....70

Figure A.1 Average parameter estimates and 95% confidence intervals from 10,000 Monte Carlo simulations at different levels of underreporting for moose-vehicle collisions in western Maine, USA, during 2000–2011. The model with good predictive capability included: TRAFFIC = annual average daily traffic, SPEED = speed limit (km/hr.), P_CUT = proportion of cutover forest land-cover within 2,500 m buffer, P_CONIF = proportion of conifer forest land-cover within 2,500 m buffer, IJI = interspersed-juxtaposition index within 5,000 m buffer, D_DEV = distance to development (m), D_SHBW = distance to shrub-wetland land-cover (m). The model with poor predictive capability included: SLOPE = degree of slope, D_STR = distance from stream (m), SIDI = Simpson’s diversity index within 5,000 m buffer.....81

Figure A.2 Average parameter estimates and 95% confidence intervals from 10,000 Monte Carlo simulations of logistic regression analyses including spatial biases in reporting for different levels of underreporting for moose-vehicle collisions in the most rural county in western Maine, USA, during 2000–2011. The model with good predictive capability included: TRAFFIC = annual average daily traffic, SPEED = speed limit (km/hr.), P_CUT = proportion of cutover forest land-cover within 2,500 m buffer, P_CONIF = proportion of conifer forest land-cover within 2,500 m buffer, IJI = interspersed-juxtaposition index within 5,000 m buffer, D_DEV = distance to development (m), D_SHBW = distance to shrub-wetland land-cover (m). The model with poor predictive capability included: SLOPE = degree of slope, D_STR = distance from stream (m), SIDI = Simpson’s diversity index within 5,000 m buffer.....82

Figure A.3 Average receiver operating characteristic function (AUC) and 95% confidence intervals at different levels of underreporting to depict the predictive capabilities of 4 models as fewer WVCs were reported. The AUC values were calculated from 10,000 Monte Carlo simulations at different levels of underreporting for deer-vehicle collisions in central Illinois, USA during 2001 and moose-vehicle collisions in western Maine, USA during 2000–2011.....83

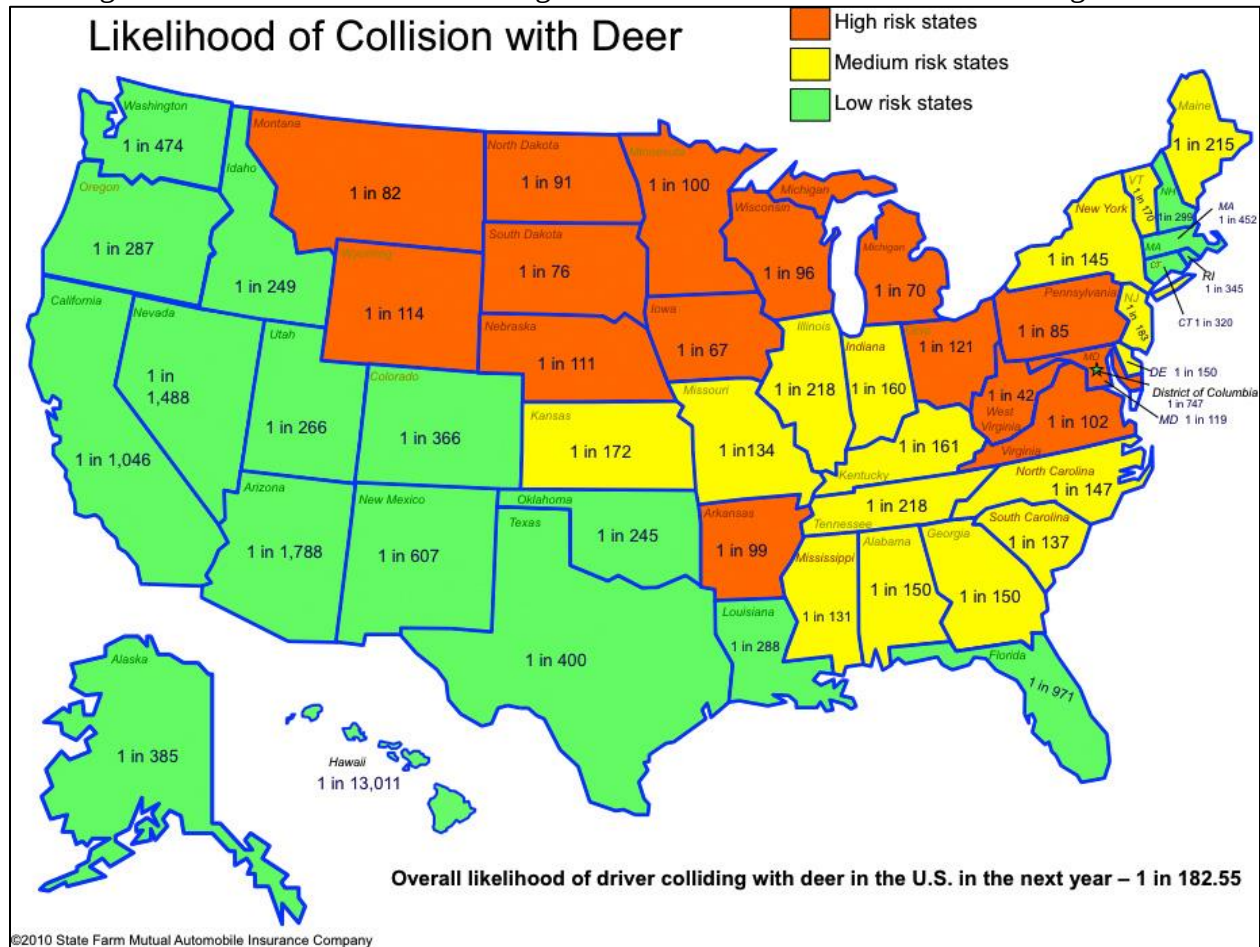
PROLOGUE

Vehicular collisions with wildlife represent only 5% of all reported crashes in North America, but their financial and ecological consequences are immense (e.g., Huijser et al. 2008). Wildlife-vehicle collisions (WVCs) represent one of the most widespread and persistent human-wildlife conflicts that exist throughout the United States. An estimated 1–2 million WVCs occur each year, and that number is increasing annually (Huijser et al. 2008). The total annual cost associated with WVCs is estimated to be >8.3 billion dollars. Approximately 26,000 injuries and 200 fatalities for humans can be attributed to WVCs each year. Since 1990, fatal collisions with wildlife have increased 104% (Sullivan 2011). Populations of wildlife suffer from WVCs, including reduced survival (e.g., Snow et al. 2012) and species endangerment (Huijser et al. 2008).

Despite the magnitude of consequences from WVCs, relatively few options exist for reducing the frequency of collisions. A number of reasons can explain the shortage of options. One reason is that identifying the most critical locations for mitigation is not straightforward. For example, hotspots of WVC locations are loosely defined, even though hotspots provide the best opportunity for cost-effective mitigation (Forman et al. 2003). Another reason for the shortage is inconsistent reporting of WVCs (Huijser et al. 2008), resulting in incomplete information for studies that analyze where collisions occur. Inferences from these studies could be unreliable because of incomplete data. Finally, there is a lack of understanding of the large-scale and long-term trends in the frequencies of WVCs. In particular, collisions with white-tailed deer (*Odocoileus virginianus*) occur throughout every state in the United States with high frequency

(Figure 1). Studies at larger geographic and temporal extents are needed to understand the environmental influences for those trends.

Figure 1.1 Map generated by State Farm Mutual Automobile Insurance Company ©2010 showing the likelihood of a motorist being involved in a collision with a deer during 2011.



To address the deficiencies in understanding the dynamics of WVCs, my goal was to bring together data sets spanning 3 species in 6 states, and draw on recent advances in statistical tools to enrich the current approaches for examining WVCs. Here I present 3 chapters that strengthen the analytical methods used to examine WVCs and enhance the exploration of their ecological relationships. Specifically, I focus on improving 3 current weaknesses that are limiting the ability to understand and manage WVCs. These weaknesses include 1) insufficient techniques for identifying the most critical locations for mitigating WVCs, 2) insufficient

knowledge about the number of WVC locations needed to inform research studies, and 3) insufficient knowledge about broad scale, ecological drivers of WVCs. The use of digital WVC records for all chapters were conducted with approval of the Michigan State University, Institutional Animal Care and Use Committee (approval date: 11 Feb 2013)

Chapter 1 focuses on delineating hotspots. The most cost-effective strategy for managing WVCs is to focus on the highest-risk locations (i.e., hotspots). However, defining a hotspot is not straightforward and has often been subjective. The definition of a hotspot varies among studies, therefore delineations of these high-risk locations are inconsistent (e.g., Openshaw and Taylor 1981, Gomes et al. 2009, Okabe and Sugihara 2012). This inconsistency leads to unreliable information for planning management strategies to reduce WVCs. I contend that these subjective choices may lead to non-reproducible results or pseudoreplicated observations from a lack of independence among delineated hotspots. I develop a non-subjective approach for delineating hotspots using variation in the surrounding landscape. Variation in the landscape provides an appropriate and convenient measure for delineating hotspots because attributes of landscapes are important ecological drivers of WVCs. I test this approach for a variety of species across different landscapes to ensure that hotspots were successfully delineated under dissimilar conditions.

Chapter 2 focuses on issues with underreporting. Many collisions are unreported, making it difficult to empirically identify hotspots that are in most critical need of management action. Only approximately 300,000 of the estimated 1–2 million WVCs in the United States are reported in national crash databases each year (Huijser et al. 2008). Many WVCs are not reported because they involve insufficient property damage to warrant reporting, motorists choose not to report them, or police, natural resource, and transportation agencies do not record

such information (Huijser et al. 2008). I test the usual caveat that these incomplete data produce reliable information about the ecological relationships on where collisions occur (i.e., influences from the environment). To my knowledge, no studies have examined this assumption, although many studies have assumed it is true. I simulate underreporting of wildlife-vehicle collisions for 2 species: white-tailed deer in central Illinois and moose in western Maine. These species cause the highest rate of monetary damage and human injury in the United States, respectively, from WVCs. They also represent 2 distinct datasets from different landscapes and road networks, thereby providing an objective comparison for how underreporting affects studies of wildlife-vehicle collisions in multiple situations.

Chapter 3 focuses on understanding dynamic, ecological drivers (e.g., from changes in traffic, abundance of deer, and the landscape) of white-tailed deer-vehicle collisions (DVCs) across large geographic and temporal extents. Deer-vehicle collisions cause the most human injuries and fatalities, and the highest amount of property damage than collisions with any other species (Huijser et al. 2008). Major changes have occurred throughout the North American landscape during the last half century (e.g., urbanization, parcelization of land, climate change, and wildlife management) that have affected wildlife habitat and abundances of wildlife populations (Roseberry and Woolf 1998, Thompson et al. 1998, Drzyzga and Brown 1999, Kling et al. 2003). These changes are predicted to continue, and it is unclear how they will affect the frequencies of DVCs in the future. Smaller scale studies have assumed that the ecological drivers of DVCs are constant through space (Finder et al. 1999, Hubbard et al. 2000, Nielsen et al. 2003, Ng et al. 2008, Hothorn et al. 2012). Other studies examined large geographic extents, but assumed the ecological drivers were constant through time (Finder et al. 1999, Sudharsan et al. 2005, Farrell and Tappe 2007, Hothorn et al. 2012). I apply a new technique to examine for

dynamic, environmental influences on the frequency of DVCs across 3 eco-zones (Northern Forest, Agriculture-Forest Matrix, and Agriculture) that intersect 4 Midwestern states ($n = 355$ counties throughout Illinois, Iowa, Michigan, and Wisconsin) during 12 years (2000–2011). I use a temporally dynamic model to estimate how traffic, abundance of deer, and attributes of the landscape interact to influence the frequency of DVCs for each eco-zone. The goal of this analysis is to determine how large-scale changes in the environment influence the frequencies of DVCs through time. Understanding these large-scale ecological relationships will help managers prioritize mitigation strategies for DVCs in the future.

A LANDSCAPE-BASED APPROACH FOR DELINEATING HOTSPOTS OF WILDLIFE-VEHICLE COLLISIONS

ABSTRACT

Imposing human perceptions about the scales of ecological processes can produce unreliable scientific inferences in wildlife research and possibly misinform mitigation strategies. An example of this disconnect occurs in studies of wildlife-vehicle collisions (WVCs). Subjective procedures are often used to delineate hotspots of WVCs, resulting in hotspots that are not spatially independent. I developed a new approach that identifies independent hotspots using attributes of the landscape to inform delineations instead of subjective measures. First, I generated a candidate set of grouping scenarios using unique combinations of kernel-density estimation (KDE) parameterization (i.e., bandwidth and isopleth values). Next, I associated the groups of WVCs with attributes of the surrounding landscape. Finally, I identified the grouping scenario with the highest amount of variation in the landscape among the groups. The highest variation corresponded to hotspots that were most distinguishable from each other (i.e., most independent) based on the surrounding landscape. I tested my approach on 3 species of wildlife (island foxes [*Urocyon littoralis*] on San Clemente Island, CA; white-tailed deer [*Odocoileus virginianus*] in Onondaga County, NY; and moose [*Alces alces*] in western Maine) that exemplified varying degrees of space-use in different landscapes. I found that the landscape-based approach was able to effectively delineate independent hotspots for each species without using subjective measures. The landscape-based approach delineated fewer or larger hotspots than currently used methods, suggesting a reduction in spatial dependency among hotspots. Variation in the landscape indicated that hotspots may be larger than previously identified; therefore current mitigation strategies should be adjusted to include larger areas of high risk.

2.1 INTRODUCTION

Human perceptions about scales of ecological processes may not closely match associated wildlife behaviors (Wiens 1976;1989). Examples of this disconnect occur in studies of wildlife that delineate hotspots of occurrence (i.e., areas of high incidence) using point-process data. Methods for delineating hotspots have evolved to rely on increasingly sophisticated quantitative tools, but most methods require assumptions that are built upon human perceptions of how animals respond to the environment. A variety of different assumptions are used, resulting in hotspots that are inconsistent and possibly pseudoreplicated. Accurately delineating hotspots is important for wildlife research because they often are indicative of influential processes that are affecting populations of wildlife.

Hotspots of wildlife-vehicle collisions (WVCs) are used to determine what environmental factors influence where the highest risk locations of WVCs exist. Typically, hotspots are the sample units in statistical models that examine how landscape, traffic, and abundance of wildlife influence the occurrence of WVCs (e.g., Malo et al. 2004, Ramp et al. 2005, Gomes et al. 2009, Danks and Porter 2010). These hotspots are treated as independent sample units, although the amount of dependency among them is usually unknown. If they are not independent, then they are pseudoreplicated. Pseudoreplication can mislead scientific inferences by identifying conflicting or invalid relationships, or underestimating the true variation in statistical models (Hurlbert 1984, Heffner et al. 1996).

A variety of methods are currently used for delineating hotspots of WVCs. One method is to ignore hotspots and treat all WVCs as independent observations (e.g., Snow et al. 2011). Another method uses predefined distances to group WVCs into hotspots (e.g., Ng et al. 2008). More sophisticated approaches use the counts of WVCs within predefined lengths of road

segments to identify hotspots (e.g., Malo et al. 2004, Ramp et al. 2005, Gomes et al. 2009). Even more sophisticated approaches use nearest-neighbor clustering with predefined threshold distances and the overall length of roads in the study area (e.g., Levine 2004, Clevenger et al. 2006), or use kernel-density estimators (KDEs) to identify hotspots (e.g., Xie and Yan 2008, Okabe et al. 2009, Danks and Porter 2010). Subjective choices are required for all these approaches and include decisions such as: 1) assuming every location is independent, 2) selecting the lengths of road segments, 3) selecting the length of threshold distances, or 4) defining the parameters for KDEs (i.e., bandwidths and isopleths).

Kernel-density estimators provide a promising, non-parametric, approach for objectively identifying independent groups of WVCs. For studies of wildlife, the application of KDEs has recently expanded from identifying boundaries of home ranges (e.g., Worton 1989, Seaman and Powell 1996, Seaman et al. 1999, Laver and Kelly 2008) to identifying hotspots of WVCs (e.g., Danks and Porter 2010). Kernel-density estimation uses a group of spatially-referenced points (i.e., observations) to generate a probability surface based on the concentration of observations across 2-dimensional space (Bailey and Gatrell 1995). Generating the probability surface depends on a user-specified, bandwidth smoothing parameter (Kernohan et al. 2001). Bandwidth parameters represent the amount of contribution each observation point contributes to the entire probability surface (Gitzen et al. 2006). A large bandwidth value specifies broad smoothing, and generates a smooth surface of mostly high probability (Kernohan et al. 2001). Whereas, a small bandwidth represents narrow smoothing, and generates a more fragmented surface of probability.

After a probability surface has been generated with KDE, isopleth lines are used to construct hard boundaries around user-specified volumes of the probability surface (Beyer 2012). For example, a 0.95 isopleth represents a boundary around 95% of the volume of probability. A

0.05 isopleth represents a more constricted boundary around 5% of the volume. Observations are grouped together within the boundaries of isopleths, or are not grouped and are considered single-occurrence events. The amount of grouping relies on the size of the bandwidth and percentage of isopleth used. Selection of these values has been highly scrutinized for studies of home ranges (Gitzen et al. 2006, Laver and Kelly 2008), but is not well understood for studies of WVCs.

I propose a new approach for parameterizing KDEs to delineate WVCs into hotspots without relying on subjective choices for bandwidths and isopleths. I suggest using measures of variation (i.e., variance) in the landscape surrounding locations of WVCs to inform non-subjective parameterization of KDEs. Attributes of the landscape provide a useful measure because WVCs are influenced by the landscape (Huijser et al. 2008). Specifically, variation of the landscape can inform how WVCs should be grouped based on the amount of dispersion (i.e., dissimilarity) identified in attributes of the landscape among proposed groups of WVCs. If variation among a set of proposed groups is low, then these groups are not easily distinguishable from each other. As variation increases, the groups become more distinguishable and independent, based on the landscape.

Examining for maximum variation is a concept developed for understanding scales of animal movement (i.e., first-passage time; Fauchald and Tveraa 2003, Williams et al. 2012), but can be expanded to identify independent hotspots of WVCs. I suggest using variation in attributes of the landscape as a means to objectively group WVCs into independent hotspots. Groups of WVCs that are associated with the greatest amount of variation in the landscape can be considered the most independently delineated groups possible. The amount of independence is informed by the landscape, and not by subjective measures of distance between WVCs.

The purpose of my paper is to explore an objective approach for grouping locations of WVCs into independent hotspots. Specifically, I used attributes of the landscape to inform KDE parameterization for grouping locations of WVCs into hotspots. I sought to explore the robustness of this approach by comparing it to previously used methods for delineating hotspots under a variety of conditions. Specifically, I compared my approach to 3 different methods that have been used for 3 species of wildlife, respectively, including 1) island foxes (*Urocyon littoralis*) on San Clemente Island, California, USA, 2) white-tailed deer (*Odocoileus virginianus*) in Onondaga County, New York, USA, and 3) moose (*Alces alces*) in the western region of Maine. This combination of species represented a gradient of animal space-use in a variety of landscape types.

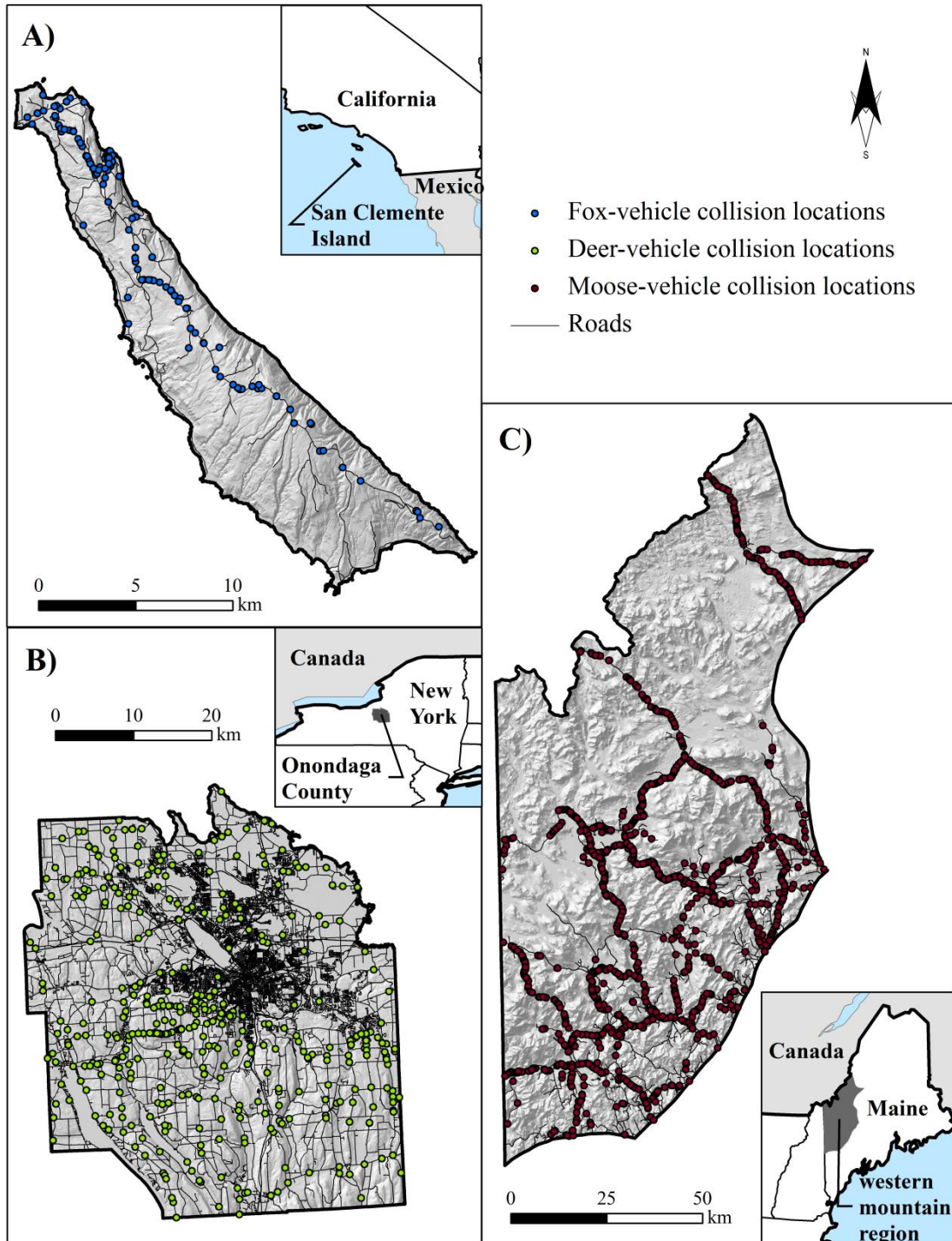
2.2 METHODS

2.2.1 Study area

The subspecies of island fox, (*U. l. clementae*), is found on San Clemente Island (146 km²). The island is the southernmost California Channel Island, located approximately 109 km west of San Diego, California (Figure 2.1A). Vegetation on the island was comprised primarily of 2 cover types: maritime desert scrub (54.4%) and grassland (32.8%; Thorne 1976, Sward and Cohen 1980). The island contained 613.5 km of roads for an overall road density of 4.2 km/km². White-tailed deer are found throughout Onondaga County, NY (2,085 km²). The county is located in the central region of New York State (Figure 2.1B). Vegetation throughout the county was comprised of a mix of forest (35%) and agriculture (33%) with small and large residential and commercial development (19%). The county contained 6,107 km of roads, for an overall road density of 2.9 km/km². Moose are found throughout the Western Mountains biophysical region

of Maine (10,721 km²). This region is located in the northern reach of the Appalachian Mountains (Figure 2.1C). Vegetation in Western Maine was mostly comprised of deciduous, conifer, or mixed forests (85%) with interspersed shrub wetlands (6%). Western Maine contained 2,474 km of roads, for an overall road density of 0.2 km/km².

Figure 2.1 Study areas, roads, and locations of wildlife-vehicle collisions for: A) island foxes on San Clemente Island, CA, USA (2006–2010), B) white-tailed deer in Onondaga County, NY, USA (2005–2006), and C) moose in Western Mountains biophysical region, ME, USA (1993–2010).



2.2.2 Data collection

I compiled records of island fox-vehicle collisions from 2006–2010, provided by the United States Navy and Colorado State University (Snow et al. 2011). Data were collected at accident sites using a handheld Global Positioning System device. I used a database of white-tailed deer-vehicle collisions from 2005–2006, provided by the State University of New York (Nystrom 2007). These data were compiled from law enforcement records and field observations. The deer-vehicle collision locations were verified and recorded using a handheld Global Positioning System device. Lastly, I used recorded locations of moose-vehicle collisions from 1993–2010, provided by the Maine Department of Transportation. These data were assimilated from law enforcement information at accident sites, and compiled with an estimated accuracy of 160 m (D. Brunell, Maine Department of Transportation, personal communication). *Post hoc*, I evaluated a 2-year subset of the moose-vehicle collision data (2008–2010) to represent the most recent collisions.

I used the 2006 Coastal Change Analysis Program for San Clemente Island, CA and western Maine to describe the land cover and land use (National Oceanic and Atmospheric Administration Coastal Services Center 2012). I used the 2001 National Land Cover Database for Onondaga County, NY (Homer et al. 2007). Land-cover and land-use maps were based on data collected with Landsat 7 Thematic Mapper with 30-m resolution with 85% overall classification accuracy for the Coastal Change Analysis program and 85.3% for the National Land Cover Database (Wickham et al. 2010, National Oceanic and Atmospheric Administration Coastal Services Center 2012). I reclassified land-cover and land-use types based on habitat requirements for each species (Table 2.1). For San Clemente Island, I used a 10-m digital elevation map from the United States Geological Survey, National Elevation Dataset (Gesch et

al. 2002, Gesch 2007), and shapefiles depicting urban areas (Gould and Andelt 2011). For western Maine, I used shapefiles depicting human development based on 1:24,000 quadrangles (Maine Office of GIS 2010).

Table 2.1 Reclassified land-cover and land-use types for 3 species of wildlife: **A)** island foxes on San Clemente Island, CA, USA (2006–2010), **B)** white-tailed deer in Onondaga County, NY, USA (2005–2006), and **C)** moose in western Maine, USA (1993–2010).

A) San Clemente Island, CA		B) Onondaga County, NY		C) Western Mountains, ME	
Class	%	Class	%	Class	%
Grassland	75.37	Forest	44.07	Deciduous-mixed forest	57.62
Scrub/shrub	20.61	Agriculture	29.73	Coniferous forest	27.70
Disturbed	3.04	Open water	12.20	Shrub wetland	6.14
Other	0.98	Rangeland	7.42	Open water	3.90
		Developed	5.81	Developed	3.55
		Wetland	0.66	Agriculture	0.88
		Barren	0.12	Other	0.21
				Cutover forest	0.0001

2.2.3 Landscape metrics

I characterized the landscape surrounding each WVC using multiple spatial extents based on the reported area requirements for each species (Leptich and Gilbert 1989, Peek 2007, Quinn et al. 2012, Resnik 2012). I used ArcGIS (v9.3, Environmental Systems Research Institute, Inc., Redlands, California) to construct 3 buffers around each WVC. The buffers corresponded to core-use areas, small home ranges, and large home ranges for each species (Table 2.2). I also included 1 additional buffer for moose to represent an extra-large home range size because they occasionally migrate (Hundertmark 2007).

I characterized the landscape surrounding each WVC location using a variety of landscape metrics (Table 2.2). I used a variety of different metrics for each species based on their reported habitat requirements. For island foxes, I focused on their reported use of grass and shrub land covers, edges between different types of land covers, urban areas, and canyons (Moore and

Collins 1995, Gould and Andelt 2011, Resnik 2012). For deer, I focused on their reported use of agriculture and forest land covers, edges between agriculture and forest land covers, and their use of fragmented and intermixed landscapes (Quinn et al. 2012). For moose, I focused on their use of forested and wetland land covers, their use of intermixed land covers, and their avoidance of urban areas (Allen et al. 1987;1988).

I calculated composition and configuration metrics using the Fragstatsbatch extension in ArcGIS (Mitchell 2005), and program FRAGSTATS v3.3 (McGarigal et al. 2002). Composition metrics represented the proportions of specific land-cover or land-use types inside each buffered area. Configuration metrics included edge density, contrast-weighted edge density (CWED), contagion and interspersions/juxtaposition index (IJI). Edge density was the sum of the length of borders between cover types divided by the area of the buffered area (km/km^2). Contrast-weighted edge density measured edges between agricultural and forested land covers. The CWED was the sum of the borders between cover types multiplied by a corresponding contrast-weight (i.e., *weight* = 1 for agriculture and forest cover types, and *weight* = 0 for all other cover types) divided by the buffered area (km/km^2). Contagion was an index of the spatial aggregation and interspersions of similar patch types. Interspersions/juxtaposition index is an index of the intermixing of different types of patches.

I used Topography Tools for ArcGIS (Dilts 2010) to calculate the average Topographic Position Index (TPI) value within buffered areas. Each TPI value was a measure of the ruggedness of the terrain, and represented the difference between the elevation of a central pixel and the mean of the surrounding cells. I also used ArcGIS to calculate the distances from each WVC to the nearest focal land-cover and land-use type(s).

Table 2.2 Metrics for data analysis used in kernel density estimation (KDE) for 3 species of wildlife: **A)** island foxes on San Clemente Island, CA, USA (2006–2010), **B)** white-tailed deer in Onondaga County, NY, USA (2005–2006), and **C)** moose in western Maine, USA (1993–2010).

	A) Island fox	B) White-tailed deer	C) Moose
Landscape areas (km ²) ^a	0.03, 0.28, and 1.13	0.50, 1.13, and 8.04	0.78, 3.14, 19.63, and 78.54
Bandwidth range (m) ^b	20–300	100–2,000	100–2,500
Isopleth range (%) ^c	5–95	5–95	5–95
No. KDE combinations ^d	285	380	475
Metrics examined	Proportion of grassland	Proportion of agriculture	Proportion conifer forest
	Proportion of shrub/scrub	Proportion of forest	Proportion forest
	Edge density	Contrast weighted edge density	Proportion shrub wetland
	Topographic position index	Contagion	Interspersion/juxtaposition index
	Distance to urban area ^e	Interspersion/juxtaposition index	Distance to development ^e
			Distance to shrub wetland ^e

^a Areas of the landscape examined around each WVC location, based on estimates of space-use (i.e., core-use and home range) for each species.

^b Bandwidth intervals were examined every 20 m for island foxes, and every 100 m for white-tailed deer and moose.

^c Isopleth intervals were examined every 5% for all species.

^d Represents the overall candidate set of potential delineations of WVCs.

^e Landscape metric was not associated with a landscape area.

2.2.4 Data analysis

I generated a candidate set of grouping scenarios that delineated hard boundaries around groups of WVCs (Figure 2.2B). The grouping scenarios represented all permutations of WVC groups identified using KDEs (see example Figure 2.3). To create these scenarios, I calculated multiple KDEs for each species using the Geospatial Modelling Environment (v0.7.1 RC1, Beyer, H. L.), ArcGIS (v10.0), and Program R (v2.12.1, R Development Core Team). Each KDE was comprised of a unique combination of bandwidth search area and isopleth percentage parameterization (Table 2.2). I examined comprehensive ranges of bandwidths and isopleths to ensure that all reasonable grouping scenarios were generated. The smallest bandwidths were representative of the core-area requirements for each species, whereas the largest bandwidths were the limit at which probability surfaces became overly smoothed (i.e., high probabilities of WVCs extended throughout the study areas). I examined all possible values for isopleth percentages, from 5–100% by 5% intervals. Within each grouping scenario, WVCs were either partitioned into groups or were occasionally solitary (i.e., isolated away from other WVCs). I considered WVCs that were not grouped with other WVCs as single-collision events (i.e., not hotspots).

I then associated groups of WVCs within each grouping scenario to corresponding values of landscape metrics (Figure 2.2C). For groups that were comprised of ≥ 2 WVCs, I averaged the corresponding landscape metrics from each WVC to obtain an overall value for the group. I scaled and centered the metric values (i.e., subtracted the mean and divided by the standard deviation) among all groups and grouping scenarios to allow standardized comparisons among metrics and across spatial scales (i.e., buffers).

Next, for each grouping scenario I calculated the variation in landscape metrics among groups using Program R (Figure 2.2D). I calculated the variation for each spatial scale (i.e., buffer size) of the landscape metrics. I examined for peaks in variation among grouping scenarios, and identified the bandwidth and isopleth parameterization that delineated groups of WVCs with the highest variance (i.e., most disparity) relative to the surrounding landscape. The grouping scenario with the most disparity represented groups that were most independent from each other, relative to the landscape. I considered the grouping scenario with maximum variance as the landscape-based delineation of WVC hotspots (Figure 2.2E).

Once the landscape-based delineation of hotspots was made, I compared the length of road and number of hotspots to those delineated using previous methods for each species. The previous methods considered were: 1) every location of a fox-vehicle collision as a unique hotspot for island foxes (Snow et al. 2011), 2) locations of collisions buffered with 300 m radii and dissolved for white-tailed deer (Ng et al. 2008), and 3) a KDE with 1 km bandwidth and 50% isopleth for moose (Danks and Porter 2010). I compared the lengths of roads (km) that were delineated as hotspots and examined the amount of overlap (km of roads) among methodologies. There was no length of roads associated with collision events for island foxes using the previous method; therefore I was unable to compare lengths of roads between methods for island foxes. Lastly, I examined the number of landscape-based hotspots required to account for 25, 50, and 75% quantiles of WVCs.

Figure 2.2 Conceptual flowchart showing a process for delineating non-subjective hotspots of wildlife-vehicle collisions (WVC) using kernel-density estimation (KDEs). Step 1 is to gather accurate locations of WVCs. Step 2 is to generate a candidate set of grouping scenarios using KDEs with unique combinations of bandwidth and isopleth values. Additionally, calculate landscape metrics for each WVC with user-specified extent(s) of the landscape. Step 3 is to associate the landscape metrics to each grouping scenario, and then calculate the variance for each metric among groups of WVCs within each grouping scenario. Step 4 is to identify the grouping scenario with the greatest amount to variation.

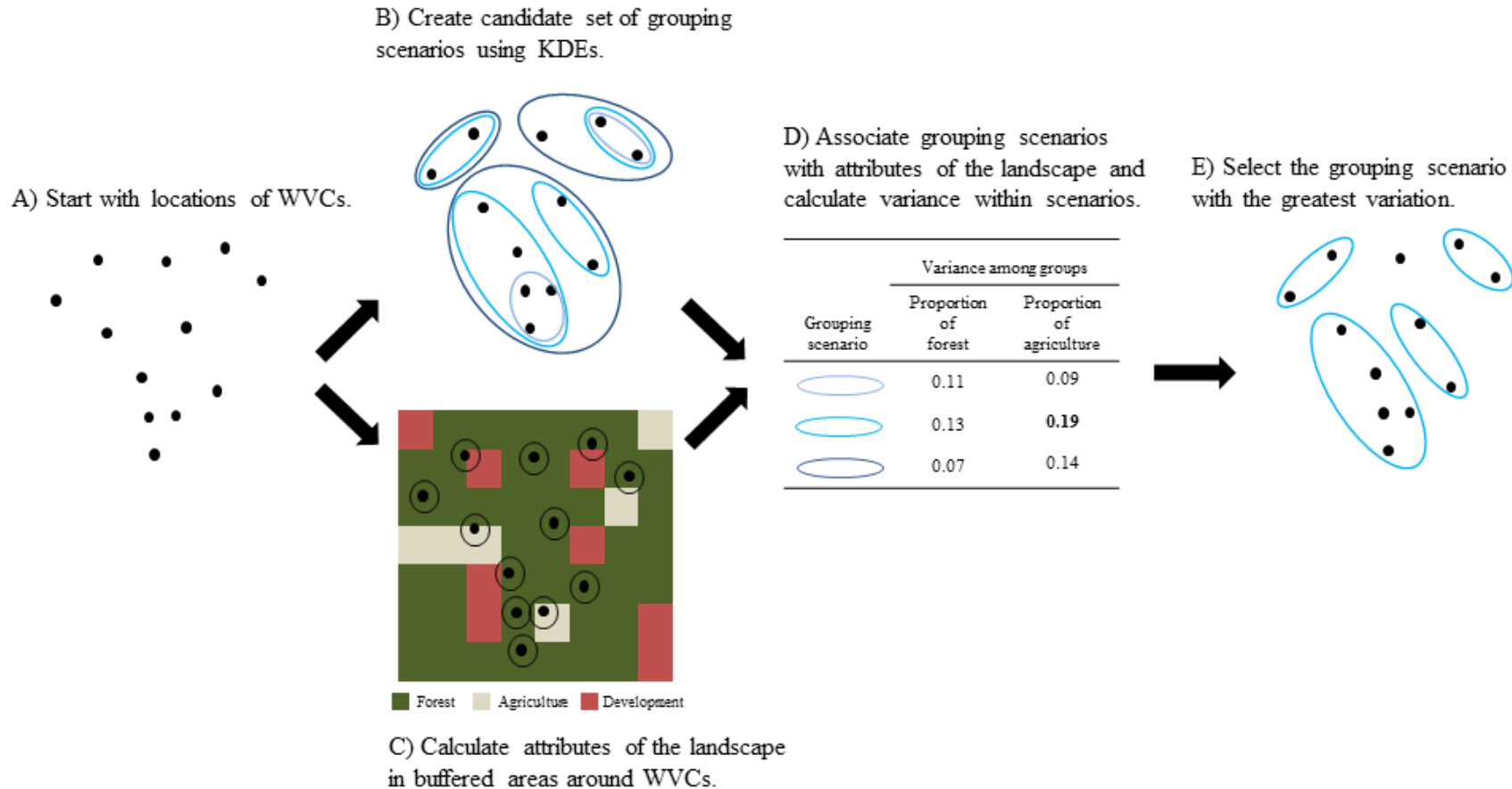
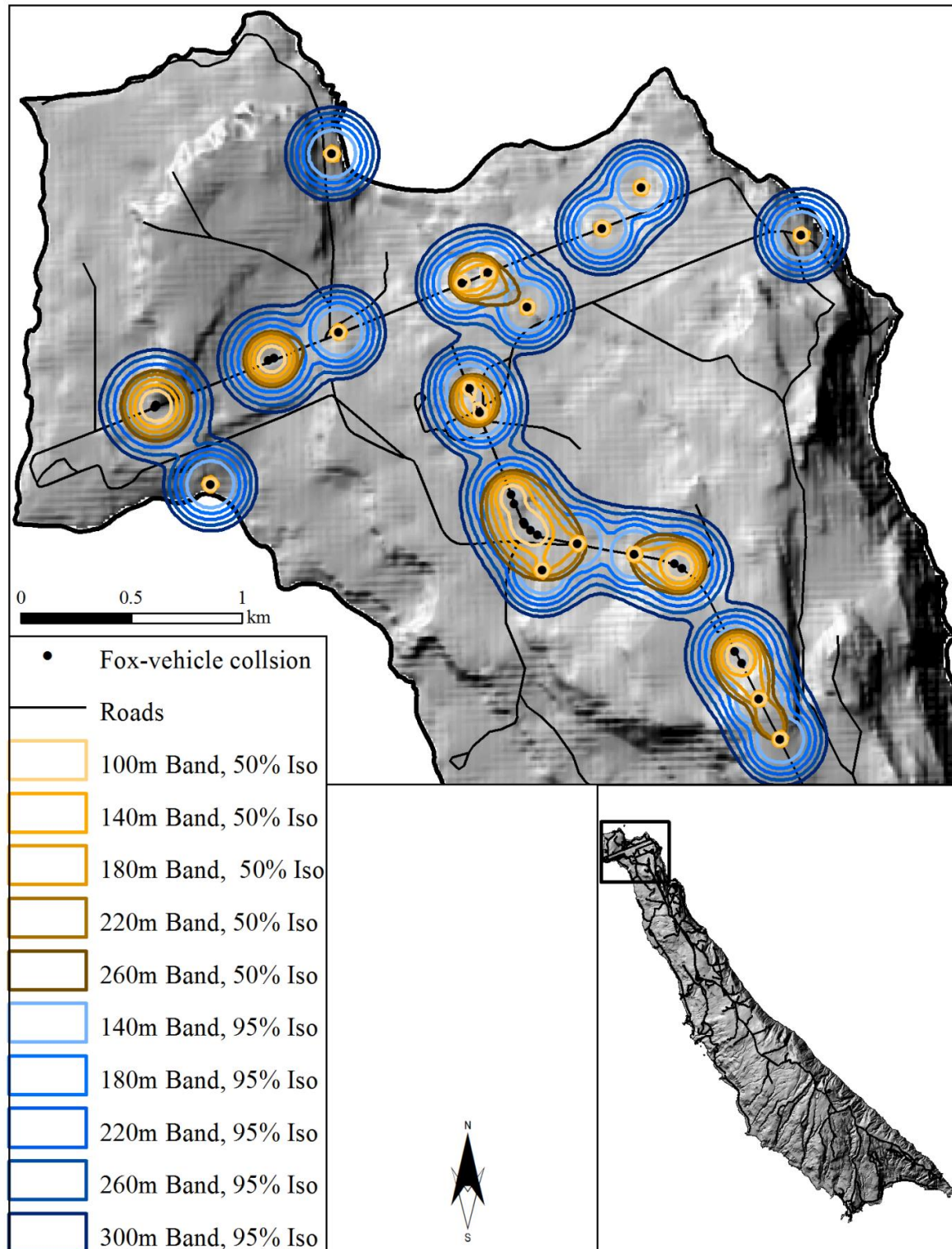


Figure 2.3 Example delineations of San Clemente Island, CA, USA (2006–2010) fox-vehicle collision hotspots calculated with kernel density estimation. Each bandwidth and isopleth combination produced a unique delineation as part of an overall candidate set of potential delineations. This figure shows 10 of the 285 potential delineations that were calculated.



2.3 RESULTS

I examined a total of 2,488 records of WVC locations and generated 1,615 grouping scenarios using KDEs with unique bandwidth and isopleth combinations for 3 species. I examined the variances of 16 landscape metrics (Table 2.2). The count of unique hotspots declined with increasing bandwidth and isopleth values as should be expected with KDE parameterization. I was able to successfully identify peaks in variation for all landscape metrics except 3. For those exceptions (i.e., TPI for island foxes, CWED for white-tailed deer, and distance to development for moose), the variance never reached a peak as the bandwidth and isopleth values increased. Thus, I considered the maximum variance as being undefined.

For island foxes, the proportion of shrub land cover at scale of large home ranges showed the maximum variance (standardized variance = 1.39; Figure 2.4), and therefore identified the most independent groups of WVCs based on a landscape attribute. The peak in variance was identified with a 260 m bandwidth and 55% isopleth (Figure 2.5). This combination delineated 21 hotspots that averaged 0.8 km of roads (SD = 1.4) and accounted for 72% of all collisions on San Clemente Island, CA (Table 2.3). The previous method did not delineate hotspots, thus I could not compare between the 2 approaches.

Table 2.3 Number of hotspots and length of roads delineated using a landscape-based approach and currently used methods for 3 species of wildlife: A) island foxes on San Clemente Island, CA, USA (2006–2010), B) white-tailed deer in Onondaga County, NY, USA (2005–2006), and C) moose in western Maine, USA (1993–2010 and 2008–2010).

Species	No. WVCs	Landscape-based				Previous method			
		No. hotspots	Avg. WVCs/ hotspot	Hotspot roads (km)	Single- collision events	No. hotspots	Avg. WVCs/ hotspot	Hotspot roads (km)	Single- collision events
A) Island fox	132	21	4.6	18.0	36	0	NA	NA	132
B) White-tailed deer	389	51	5.2	830.9	122	53	1.3	438.6	255
C) Moose (17 yr.)	1,927	42	43.1	990.5	161	99	13.0	258.5	635
C) Moose (2 yr. subset)	172	20	3.3	60.7	106	29	2.9	64.7	87

For white-tailed deer, the proportion of forest land cover at the scale of large home ranges was the most informative landscape attribute (standardized variance = 1.33; Figure 2.4) for delineating hotspots in Onondaga County, NY. A peak in variance was identified with a 1,600 m bandwidth and 50% isopleth (Figure 2.5). The landscape-based approach identified 51 hotspots that accounted for 69% of WVCs in Onondaga County, NY. The previous approach identified 53 hotspots that accounted for 34% of WVCs. The landscape-based approach delineated larger hotspots (mean = 13.8 km of road; SD = 34.4) than the previous method (mean = 1.4 km; SD = 0.9; Table 2.3). It overlapped the previous method along 293 km of roads, but also included 538 km more roads as hotspots.

For the full set of moose data (17 years), I found IJI at the landscape scale of core-use areas had the maximum variance (standardized variance = 1.33; Figure 2.4) and delineated the most independent groups of WVCs based on a landscape attribute. I identified a peak in variance at a bandwidth of 2,300 m and isopleth of 80% (Figure 2.5). The landscape-based approach identified 42 hotspots that accounted for 92% of WVCs in western Maine (Table 2.3). The previous approach identified 99 hotspots that accounted for 67% of WVCs. The mean length of a hotspot was 22.5 km (SD = 49.1), whereas those identified by the previous method were substantially shorter (mean = 2.6 km; SD = 3.6). The landscape-based approach overlapped all roads designated as hotspots by the previous method (i.e., 263 km), plus an additional 727 km.

For the subset of moose data (2 years), IJI had the maximum variance (standardized variance = 1.17) at the scale of large home ranges. The peak in variance occurred at bandwidths of 2,400 and 2,500 m and an isopleth of 20% (Figure 2.5). Using this combination I delineated 20 hotspots with a mean length of 2.8 km (SD = 1.8), that accounted for 38% of WVCs in western Maine. The previous approach identified 29 hotspots (mean = 1.2 km; SD = 0.8), that

accounted for 49% of WVCs. Overall, the landscape-based approach delineated 61 km of roads as hotspots, similar to the previous method that delineated 65 km.

Figure 2.4 Maximum variance values (standardized) for each landscape metric among hotspots of wildlife-vehicle collisions for: A) island foxes on San Clemente Island, CA, USA (2006–2010), B) white-tailed deer in Onondaga County, NY, USA (2005–2006), C) moose in Western Mountains biophysical region, ME, USA (1993–2010), and D) moose 2-years subset (2008–2010).

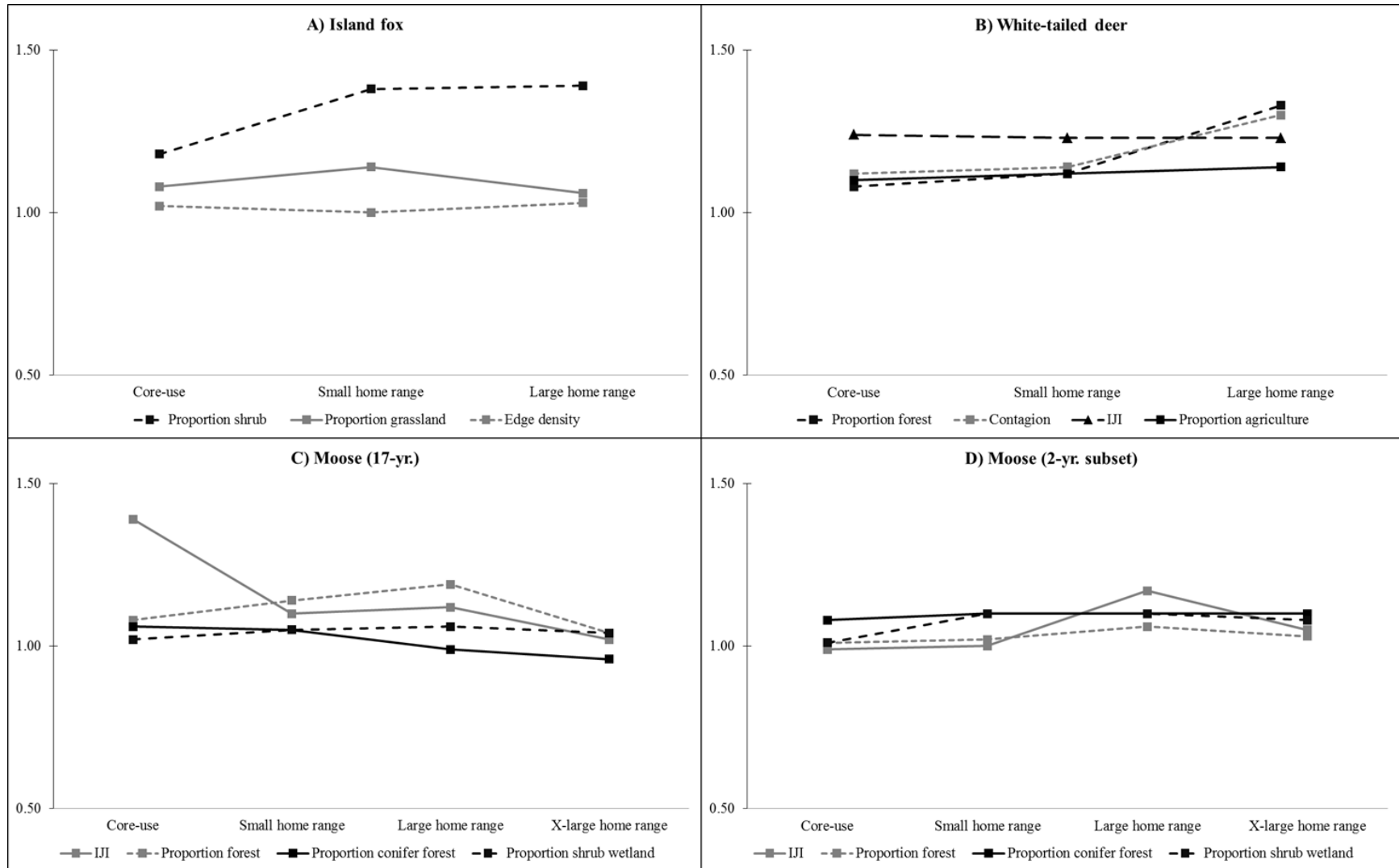
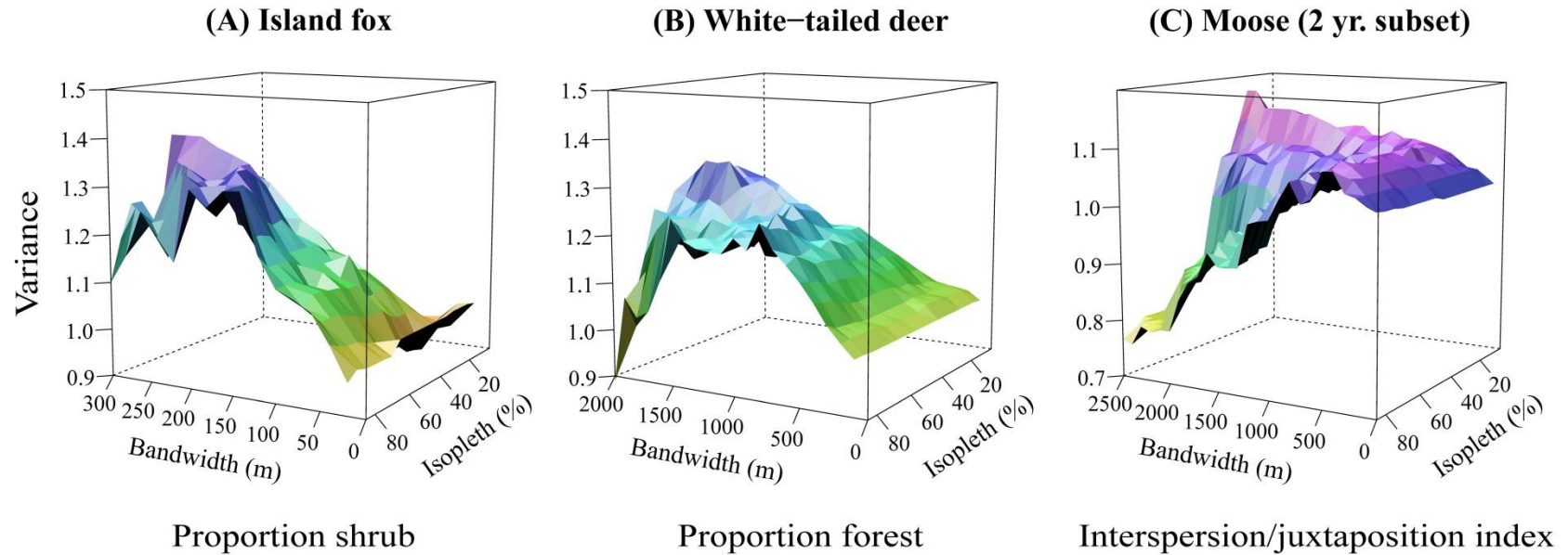


Figure 2.5 Peaks in variation were identified at: A) 260 m bandwidth and 55% isopleth for the proportion of shrub landscape metric for island foxes on San Clemente Island, CA, USA (2006–2010), B) 1,600 m bandwidth and 50% isopleth for the proportion of forest landscape metric for white-tailed deer in Onondaga County, NY, USA (2005–2006), and C) 2,400 or 2,500 m and 20% isopleth for the interspersed-juxtaposition index landscape metric for the 2-years subset of moose in the Western Mountains biophysical region, ME, USA (2008–2010).



2.4 DISCUSSION

The landscape-based approach provides some clear improvements over the previously used methods. First, by using variation in landscape metrics to inform the delineation of hotspots, the landscape-based approach ensures that the most independent groups of WVCs are identified relative to the surrounding landscape. Previous methods disregard issues with pseudoreplication among groups of WVCs by only considering subjective measures of spatial proximity to delineate hotspots. Using variation in the landscape as measures of independence provides an objective approach for avoiding pseudoreplication in statistical models of WVCs. The previous strategies further ignore the ecological processes that influence the arrangement of WVCs, and thereby provide little information for reducing pseudoreplication among delineated hotspots. Ensuring independent observations is important for studies that use statistical models to examine for influences on hotspots of WVCs. Otherwise, the true variation in parameter estimates will be underestimated by pseudoreplicated samples (Hurlbert 1984, Heffner et al. 1996).

Second, the landscape-based approach performs well in a variety of situations and thereby provides a flexible, but consistent, methodology for delineating hotspots. Comparatively, the previous methods yielded inconsistent delineations of hotspots because of the variety of methodologies used (Openshaw and Taylor 1981, Gomes et al. 2009, Okabe and Sugihara 2012). Consistent approaches will afford more reliable comparisons among species and environments. The landscape-based approach allows for differing degrees of space-use by animals and differing complexities of landscapes by incorporating multiple spatial scales and landscape metrics. Using multiple scales and metrics also reduces the chances of biasing the delineation hotspots based on the researcher's perceptions.

Third, the landscape-based approach uses the landscape to inform hotspots and can be expanded to include other influences that affect hotspots. An obvious expansion includes examining variation in volume and speed of traffic for delineating hotspots (e.g., Forman et al. 2003). A critical requirement will be that data on the volume and speed of traffic are available at sufficient resolution to calculate variation. To my knowledge, the landscape-based approach is the first, flexible approach for using the ecological processes to help determine how hotspots are delineated.

Fourth, the landscape-based approach provided a less subjective and easily identifiable means for delineating hotspots. I avoided subjective choices in 3 ways. First, I used variation of the landscape metrics as non-subjective criteria for selecting grouping scenarios that represented the most unique delineation of hotspots. Second, I used the biology of each species to inform the landscape metrics and the scales at which I examined them. These metrics quantified important landscape variables for the habitat requirements of each species based on previous literature. I examined multiple metrics at multiple scales to avoid bias from human perceptions. Third, I examined the entire ranges of bandwidth and isopleth values that could be used in KDE analyses to group WVCs, and therefore assured that all possible combinations were tested. By combining these techniques, I developed the first-known, landscape-based approach that successfully informed the delineation of the most independent hotspots without imposing human perceptions about the ecological processes involved.

Although the landscape-based approach improves upon previous methods, some constraints still exist. For most landscape metrics I tested, I easily detected peaks in variance. For 3 metrics, however, a peak in variance could not be identified because variation appeared to be driven by sample sizes of the delineated hotspots (i.e., variance increased linearly with

decreasing numbers of hotspots). This suggests that not all landscape metrics are useful for informing hotspots. Particularly, metrics that contain very little variation throughout the landscape are less useful. I recommend comparing multiple landscape metrics that represent varying degrees of heterogeneity within the study area to inform the delineation of hotspots.

The temporal scale of WVCs may influence the delineation of hotspots. For moose in Maine, the 17-year dataset contained a large sample size of WVCs that occurred on most sections of roads, thereby resulting in hotspots that encompassed most roads. However, the more sparsely distributed 2-year subset indicated much fewer and smaller hotspots, suggesting that temporal scales are important considerations for delineating hotspots. If landscape metrics change through time, then delineating hotspots without considering the temporal dynamics of the landscape may not be useful. However, if landscape metrics are relatively stable, then using locations of WVCs over longer timeframes should delineate more accurate hotspots.

The landscape-based approach identified fewer or larger hotspots than the previously used methods, providing some important implications for mitigating WVCs. My approach suggests that larger contiguous areas may need to be targeted for mitigating WVCs. For example, fencing may need to be extended over larger areas to exclude wildlife from roads for some hotspots. My approach also indicated that fewer hotspots may need to be targeted to reduce WVCs. For example, >20 hotspots accounted for 50% of all WVCs for each species, respectively. Managers can use this information to target mitigation efforts in a more cost-effective way (e.g., Clevenger et al. 2006, Huijser et al. 2009, Conover 2010). My results indicate that previous methods may not consider large enough spatial scales for delineating hotspots. This finding is similar to recent evidence that scales of effect should be measured at larger scales than are previously used for predicting population responses to landscape structure

(Jackson and Fahrig 2012). Finally, my approach indicated that hotspots included more WVCs, on average, than the previously used methods. Grouping more WVCs into hotspots will reduce the chances of analyzing pseudoreplicated collision sites in exploratory models.

Lastly, my landscape-based approach can be extended beyond hotspots of WVCs. Many other ecological studies require objective delineations of hotspots, such as hotspots of bird nests (e.g., Hatchwell et al. 1996), insect infestations (e.g., Nelson and Boots 2008), or species distributions (e.g., Stohlgren et al. 2001). These hotspots can be delineated without relying on human perceptions about the ecological processes to provide the most unbiased estimates and inferences. Researchers can examine for peaks in variation from a variety of inputs (i.e., not just landscapes) that might inform how hotspots are delineated.

2.5 ACKNOWLEDGEMENTS

I thank W. F. Andelt from Colorado State University and M. A. Booker from the United States Department of the Navy for providing records of island fox-vehicle collisions. I thank H. B. Underwood and S. Nystrom from the State University of New York College of Environmental Science and Forestry for providing records of white-tailed deer-vehicle collisions. I thank D. Brunell from the Maine Department of Transportation for providing records of moose-vehicle collisions. I appreciate helpful reviews of this manuscript from Z. D. Danks and anonymous reviewers. This research was funded in part by the Boone and Crockett Club, the Michigan Department of Natural Resources, the Alces Journal, and the Michigan Involvement Committee of Safari Club International. I thank Michigan State University, the Department of Fisheries and Wildlife, and the Boone and Crockett Quantitative Wildlife Laboratory for support. The final publication is available at <http://link.springer.com/article/10.1007/s10980-014-0018-y/fulltext.html>.

UNDERREPORTING AND ASSESSING WILDLIFE-VEHICLE COLLISIONS WITH LARGE UNGULATES

ABSTRACT

Conflicts from wildlife-vehicle collisions (WVCs) pose serious challenges for managing and conserving large ungulates throughout the world. Identifying research-based management strategies to reduce these conflicts is problematic because up to 60% of WVCs are not reported, and are subsequently excluded from study. Our objective was to examine the sensitivity of studies to underreporting for exemplary species of large ungulates in different environmental settings: white-tailed deer (*Odocoileus virginianus*) in open landscapes of central Illinois and moose (*Alces alces*) in forested areas of western Maine, USA. We simulated underreporting and evaluated for changes in precision, parameter estimates, and prediction in identifying relationships between the environment and location of collisions. We found similar effects for both species, indicating our findings are widely applicable. Relatively few reports (30%) were needed to generate reliable assessments of WVC incidence, likely because WVCs tend to occur in non-random patterns (i.e., hotspots) and variability in their influences is low. Shifts in parameter estimates were detected only for environmental variables that were unevenly distributed in locations biased with less reporting. The predictive capabilities of models were stable with few reports (~25%). These findings suggest that reliable inferences are produced where the rates of reporting are consistent throughout the study region, even with low rates of reporting. To avoid unreliable inferences, studies of WVCs should adopt study designs that primarily ensure consistent rates of reporting and secondarily ensure large enough sample sizes. Increased reporting will reduce shifts in parameter estimates from biased reporting, but will not overcome it. These principles can extend to other studies that experience underreporting, such as survey studies.

3.1 INTRODUCTION

High degrees of underreporting for wildlife-vehicle collisions (WVCs) are disconcerting for natural resource and transportation managers that strategically plan for reducing collisions. These strategies often rely on information about the ecological drivers of collisions to inform cost-effective solutions (Forman et al. 2003). Obtaining reliable information is challenging because only 300,000 of the estimated 1–2 million WVCs are reported in national crash databases each year (Huijser et al. 2008). Large amounts of underreporting can reduce the ability to distinguish ecological drivers of collisions, or shift the estimated relationships if reporting is not evenly distributed (i.e., spatially biased; Groves 2004, Lavrakas 2008).

Reasons WVCs are not reported include insufficient property damage to warrant reporting, motorist decision not to report, or police, natural resource, and transportation agency conclusion that the accident does not merit reporting (Huijser et al. 2008). Even if concerted actions are taken to locate WVCs (e.g., Clevenger et al. 2001, Sullivan et al. 2004, Ramp et al. 2005, Snow et al. 2011), underreporting persists because injured animals move away from roads following collisions (Snow et al. 2012), are scavenged or decomposed, or are covered by roadside vegetation and not detected. Consequently, it is important to determine whether underreporting is affecting statistical inferences about WVCs, and what proportion of reports is necessary for obtaining reliable inferences.

The central question prompted by underreporting is whether the many studies conducted to date are affected by lack of precision or bias. Studies of WVCs explore environmental conditions associated with high probabilities of collisions (e.g., Ng et al. 2008, Grilo et al. 2009, Danks and Porter 2010, Snow et al. 2011, Hothorn et al. 2012). Generally, these studies compare sites of collisions and non-collisions using logistic regression models and information theoretic

procedures to evaluate how the landscape, traffic, and abundance of wildlife influence the probability of WVCs. Regression coefficients from these models and associated 95% confidence intervals (CIs) are used to determine which variables influence the probability of WVCs (e.g., Finder et al. 1999, Danks and Porter 2010, Snow et al. 2011). Then, inferred relationships from these variables inform strategies for mitigating collisions.

Collisions with white-tailed deer (*Odocoileus virginianus*; Zimmermann, 1780) are among the most frequent collisions, estimated at >1 million each year in the United States (Conover et al. 1995). Deer-vehicle collisions generate the highest amount of monetary damage from WVCs, averaging \$6,717 per collision (Huijser et al. 2008). Collisions with moose (*Alces alces*; Linnaeus, 1758) generate the highest rate of human injuries and death. Up to 10% of collisions with moose result in human injury or fatality (Huijser et al. 2008). Reporting for these collisions is not consistent throughout the United States, with some states prioritizing better than others. The Departments of Transportation in Illinois and Maine prioritize compiling accurate reports of ungulate-vehicle collisions, respectively. These databases are among the most comprehensive reporting of WVCs in the United States. As such, they provide an opportunity to independently assess sensitivity to underreporting for 2 species of wildlife with differing population abundances, and that lived in environments with differing landscapes and traffic.

Our objective was to evaluate the sensitivity of statistical models for detecting influences of landscapes, traffic, and abundance of wildlife on the probability of WVCs to varying degrees of underreporting. Specifically, we examined the potential impacts for 1) reduction in precision of regression coefficients, 2) shifts in regression coefficients evidenced by significant changes in point estimates, and 3) reduction in the predictive power of models as underreporting increased. We sought to identify thresholds in reporting rates where precision, shifts in coefficients, and

prediction became unstable and generated unreliable inferences. Our intent was to evaluate whether effects of underreporting were generalizable across dramatically different conditions associated with different ecoregions, traffic, movement behaviors, and population abundances using deer in Illinois and moose in Maine.

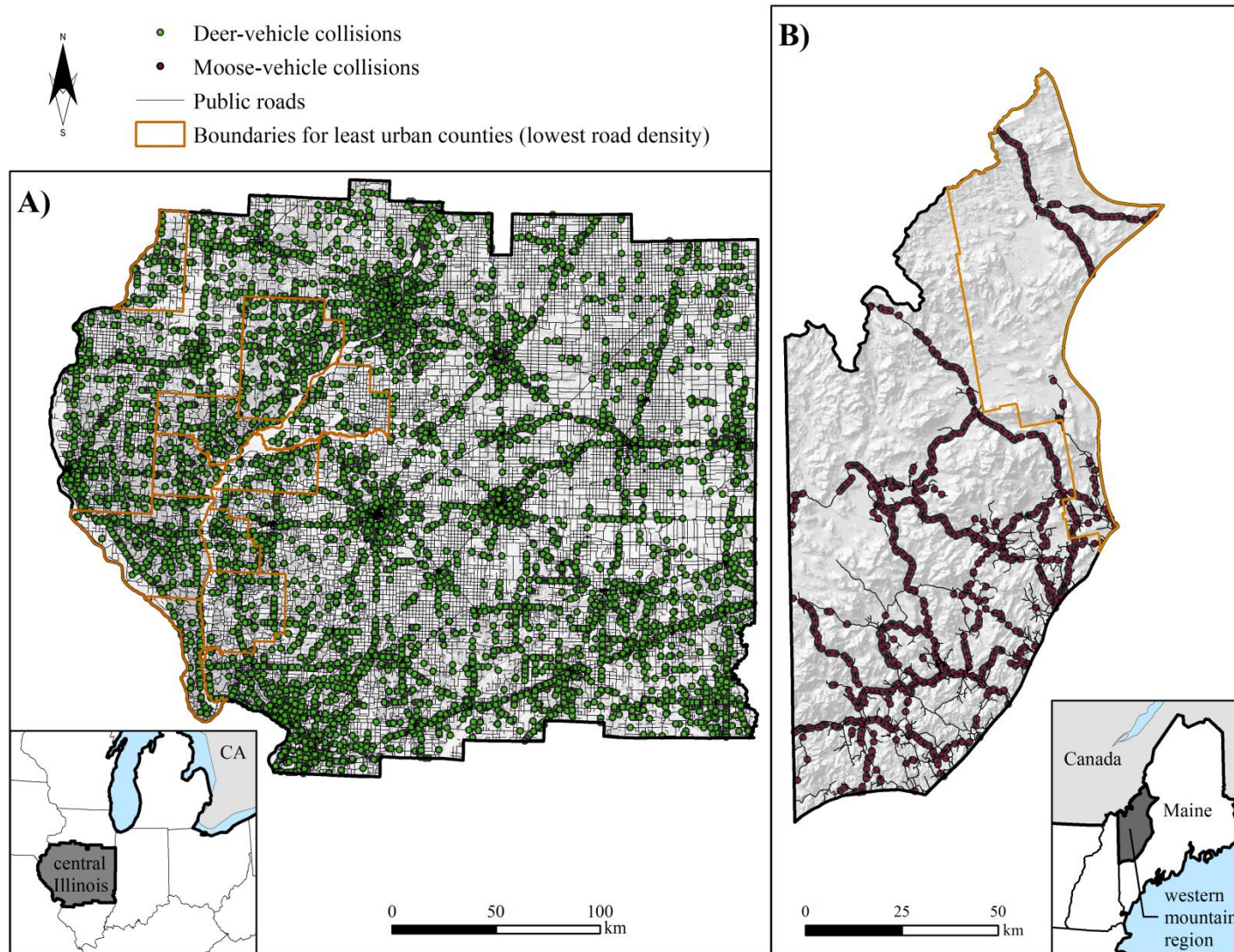
3.2 METHODS

3.2.1 Study area

My study area included 50 counties in central Illinois ($77,655 \text{ km}^2$) and portions of 3 counties in western Maine ($10,721 \text{ km}^2$; Figure 3.1). The vegetation in central Illinois was characteristic of the temperate, Prairie Parkland ecosystem province (Bailey 1980;1995). The landscape contained agriculture (74%), development (9%), intermixed deciduous trees (1.5%), and prairies and groves (<1%). Row crops are comprised primarily of a corn and soybean matrix (Rosenblatt et al. 1999). Central Illinois contains 71,498 km of public roads, for an overall road density of 0.9 km/km^2 . In 1992, the overall densities of deer were estimated to be 4–5 deer/ km^2 , but were much higher in forested areas at 30–37 deer/ km^2 (Roseberry and Woolf 1998).

Vegetation in western Maine was characteristic of the Adirondack–New England Mixed Forest– Coniferous Forest–Alpine Meadow ecosystem province (Bailey 1980;1995, Maine Office of GIS 2010). Vegetation in western Maine was composed of deciduous, conifer, or mixed forests (85%), interspersed shrub wetlands (6%), and development (3.5%). Western Maine contains 2,474 km of public roads, for an overall road density of 0.2 km/km^2 . In 1999, the densities of moose were estimated to be approximately 2.8 moose/ km^2 (Morris 1999).

Figure 3.1 Study area locations of reported A) deer-vehicle collisions in central Illinois, USA during 2011, and B) moose-vehicle collisions in western Maine, USA during 2000–2011.



3.2.2 Data collection

I used a database of 8,060 deer-vehicle collisions in central Illinois that occurred during 2011, provided by the Illinois Department of Transportation. These data were compiled from law enforcement officials where $\geq \$1,500$ in property damage or human injury occurred, with an estimated location accuracy of ± 400 m (C. Adams, Illinois Department of Transportation, personal communication). For moose-vehicle collisions, I used a database of 1,067 recorded collisions in western Maine during 2000–2010, provided by the Maine Department of Transportation. These data were compiled from law enforcement officials at accident sites where $\geq \$1,000$ in property damage or a human injury occurred, with an estimated location accuracy of ± 160 m (D. Brunell, Maine Department of Transportation, personal communication). For non-collision sites, I generated 1 set of independent, random sites separately for deer and moose that were each ≥ 1.5 times the number of reported collisions for each species using intervals of distance along the networks of roads. I generated sets of random points at intervals of 5,000 m for deer ($n = 14,306$ random points) and 500 m for moose ($n = 4,877$ random points) using ArcGIS (v10.1; Environmental Systems Research Institute, Inc., Redlands, CA). Later, I sampled from these sets of random points.

I used the 2006 National Land Cover Database to represent land cover and land use throughout central Illinois (Fry et al. 2011). For western Maine, I used the National Gap Analysis Program (GAP) Land Cover Data-Version 2 (U.S. Department of the Interior | U.S. Geological Survey 2012). Both land-cover and land-use databases were based on data collected with Landsat 7 Thematic Mapper with 30 m resolution. I reclassified the land-cover and land-use databases to 7 classes for deer (Anderson et al. 1976, Williams et al. 2012) and 11 classes for moose (Allen et

al. 1987, Koitzsch 2002, Danks and Porter 2010) based on the reported habitat requirements for each species (Table 3.1).

Table 3.1 Reclassified land-cover and land-use types for 2 study species: A) white-tailed deer in central Illinois, USA (2010), and B) moose in western Maine, USA (2000–2010).

A) Central Illinois		B) Western Maine	
Class	%	Class	%
Agriculture	74.5	Deciduous-mixed forest	57.6
Forest	15.0	Coniferous forest	27.7
Developed	8.6	Shrub wetland	6.1
Water	1.3	Open water	3.9
Wetlands	0.3	Developed	3.6
Rangeland	0.2	Agriculture	0.9
Barren	0.03	Other	0.2
		Cutover forest	0.0001
		Forested wetland	0.0
		Nonwoody wetland	0.0

Previous research identified that 800 m buffers around observations were useful for explaining influences on deer-vehicle collisions (Finder et al. 1999, Ng et al. 2008). Therefore, I used 800 m buffers to calculate composition and configuration metrics of the landscapes using FRAGSTATS (v4.1, University of Massachusetts, Amherst) for deer in Illinois. I calculated 3 composition metrics using the proportions of land-cover types, including the proportion of agriculture (P_AG), proportion of forest (P_FOR), and proportion of water (P_WAT). I calculated 2 configuration metrics, contrast-weighted edge density (CWED) and a contagion index (CONTAGION). These metrics represented measures of edge and fragmentation in the agro-forested landscape, because deer have higher densities in landscapes with these characteristics (Campa et al. 2011, Lovely et al. 2013). Contrast-weighted edge density was a measure of the length of edges between agriculture and forested, and rangeland and forested classes within each county (km/km^2). Contagion served as an index of the aggregation and

interspersed among all land-cover and land-use patches. A contagion value of 0 represented a highly fragmented and intermixed landscape, whereas a value of 100 represented a landscape comprised of a single patch.

Previous research identified that 2,500 and 5,000 m buffers around observations were most useful for explaining influences on moose-vehicle collisions (Danks and Porter 2010). Therefore, I used these buffer sizes to calculate metrics of the landscape in Maine. I calculated 4 composition metrics that influenced landscape-level habitat suitability for moose (Allen et al. 1987, Koitzsch 2002, Dussault et al. 2006). These included proportion of conifer forest (P_CONIF), proportion of cutover forest (P_CUT), proportion of nonwoody wetlands (P_NWW), and the Simpson's diversity index (SIDI) of land cover and land use within the 2,500 m buffer. The SIDI was a measure of land-cover and land-use richness on a 0–1 scale, where 1 represents the richest landscape. I calculated 1 configuration metric to represent a measure of interspersed of land-cover and land-use patches, which is also important for habitat suitability for moose (Dussault et al. 2006). I used an interspersed-juxtaposition index (IJI) to examine the complexity of the landscape within the 5,000 m buffer on a 0–100 scale, where a value of 100 represents high interspersed of patches.

For moose, I used some other measures of the landscape that were associated with moose-vehicle collisions (Danks and Porter 2010). I used ArcGIS to calculate the nearest distances to 3 landscape features: distance to the nearest shrub-wetland land-cover patch (D_SHBW), distance to the nearest stream (D_STR), and distance to the nearest developed area (D_DEV). I used shapefiles depicting streams and human development based on 1:24,000 quadrangles (Maine Office of GIS 2010). I also calculated the degree slope (SLOPE) using a 10-

m digital elevation map from the United States Geological Survey, National Elevation Dataset (Gesch et al. 2002, Gesch 2007).

I used characteristics of traffic and the abundance of wildlife to examine for influences on WVCs. I used estimates of annual average daily traffic (TRAFFIC), provided by Illinois and Maine Departments of Transportation for deer and moose, respectively, at each observation. For deer, I used estimates of the numbers of registered vehicles per county (REGISTERED_VEH) provided by the Illinois Secretary of State. I also estimated the relative abundance of deer using estimates of antlered deer harvested by county provided by the Illinois Department of Natural Resources (ABUNDANCE). The numbers of harvested moose did not influence moose-vehicle collisions in Maine (Danks and Porter 2010), thus I did not include this variable in my analyses. I used speed limits (SPEED) at each observation point for moose.

3.2.3 Study design and data analysis

To mimic common situations in which model certainty and underreporting are unknown (e.g., Ng et al. 2008, Danks and Porter 2010, Snow et al. 2011), and to evaluate how underreporting affected models with differing predictive capabilities, I examined 2 predictive models for deer and moose. One model was considered to have good predictive capabilities, and one model had poor capabilities. My criteria for the predictive capability was based on area under the receiver operating characteristic function (AUC) where good models had a value of ≥ 0.7 and poor models had a value of <0.7 (Hosmer et al. 2013). The models I evaluated were:

Deer model (good): $p = \text{TRAFFIC} + \text{ABUNDANCE} + \text{CWED} + \text{CONTAGION} + \text{P_AG} + \text{P_FOR}$

Deer model (poor): $p = \text{REGISTERED_VEH} + \text{ABUNDANCE} + \text{P_WAT}$

Moose model (good): $p = \text{TRAFFIC} + \text{D_DEV} + \text{D_SHBW} + \text{IJI} + \text{P_CUT} + \text{P_CONIF} +$

SPEED

Moose model (poor): $p = D_STR + SIDI + SLOPE$

where p was probability of a site being a WVC.

I conducted an intercorrelation analysis of the data and excluded the less biologically interpretable explanatory variable(s) from any correlated pair (i.e., $|r| \geq 0.70$; Program R v2.15.1; R Development Core Team). I scaled and centered the remaining variables (i.e., subtracted the mean and divided by the standard deviation) to allow standardized comparisons among regression coefficients. I examined for influences on locations of deer- and moose-vehicle collisions by comparing attributes of reported collisions to sites of non-collisions using a maximum-likelihood approach with generalized linear models following a binomial error term and logit-link function in Program R. I examined the logistic regression coefficients ($\hat{\beta}$) and 95% confidence intervals (CIs) to ascertain the strength and directionality of the potential influences from each environmental variable on the probability of a location reportedly being a deer- or moose-vehicle collision.

3.2.4 Evaluating sensitivity in precision

I considered the total number of reports for deer- and moose-vehicle collisions to represent 100% reporting rates. I used these 100% reporting scenarios to estimate the true relationships between the environmental variables and the probabilities of deer- and moose-vehicle collisions. Then, I evaluated the effects of underreporting by randomly excluding reports of deer- and moose-vehicle collisions, and examining for deviation from the true relationships. I randomly excluded points by sampling collision locations without replacement for deer- and moose-vehicle collisions using the sample function in Program R. I sampled collisions to represent levels of 0–95% underreporting in increments of 5% ($n = 20$ levels of underreporting). At each level, I

resampled and evaluated 10,000 Monte Carlo simulations for each of the 2 models for deer- and moose-vehicle collisions.

I randomly sampled new sets of non-collision sites for every simulation because sites of non-collisions are dependent on where the sites of deer- and moose-vehicle collisions were sampled. I used a 2-step process to sample non-collision sites for deer and moose. First, I excluded any random locations that were within 500 m from the reported locations of collisions to avoid confounding sites of collisions and non-collisions. Then, I sampled without replacement equivalent numbers of non-collision sites as numbers of collision sites that were sampled to represent the different levels of underreporting.

I plotted the mean regression coefficients and 95% CIs from the 10,000 simulations to evaluate their sensitivity to different levels of underreporting. I evaluated the precision of the regression coefficients by examining the spread of the 95% CIs and examining for any changes in statistical significance (i.e., lost or gained significance) as underreporting increased. I considered any reduction of the true precision by a factor of 2 (i.e., doubling of the spread of 95% CIs) as a conservative cutoff to identify undesirable situations for making statistical inferences.

3.2.5 Evaluating sensitivity to biased reporting

Reporting of WVCs may not be consistent among geographic jurisdictions (Knapp et al. 2005). Therefore, I conducted similar analyses as described above, but I included a spatial bias in reporting for deer- and moose-vehicle collisions based on jurisdictions of county boundaries. I biased the reporting rates to be lower throughout 10 of the 50 counties in central Illinois, and 1 of the 3 counties in western Maine (Figure 3.1). These counties represented approximately 20% of each study area. I selected these counties because they had the lowest densities of roads and

thereby represented the least urban counties, respectively, for each study area. Urbanization has been identified as an important, county-level, predictor for the incidence of WVCs with ungulates (Finder et al. 1999, Iverson and Iverson 1999, Farrell and Tappe 2007). For these counties only, I sampled deer- and moose-vehicle collisions to represent levels of 0–100% underreporting in 5% increments. The remaining counties were held constant at 100% reporting. I randomly sampled non-collision sites as described above. I conducted 10,000 Monte Carlo simulations of the logistic regression models for each level of underreporting. I plotted the mean parameters estimates and 95% CIs to evaluate their sensitivity to underreporting and determine whether reporting bias changed or weakened the statistical significance of each explanatory variable.

3.2.6 Evaluating sensitivity in model prediction

For each of the simulations I used to evaluate precision, I randomly withheld 10% of the sampled deer- and moose-vehicle collisions and 10% of the sampled non-collision sites to predict and validate the models. I input these data into the logistic regression models and examined how well the models correctly classified the withheld data into sites of collisions or non-collisions. I used the pROC package in Program R (Robin et al. 2011) to calculate the AUC values and their 95% CIs using for determining the predictive capabilities of the models. I plotted the mean AUC values and 95% CIs from the 10,000 Monte Carlo simulations for each analysis.

For all the simulations used to evaluate bias, I randomly withheld 2 sets of data to compare how well the biased models predicted: 1) new-biased data, and 2) new-true data. The first set of withheld data included 10% of the sampled deer- and moose-vehicle collisions that were spatially biased, and an equivalent number of non-collision sites. The second set included randomly sampled deer- and moose-vehicle collisions from the true dataset (i.e., not spatially biased), and an equivalent number of non-collision sites. Both sets were constrained to have the

same sample sizes at each level of underreporting. I plotted the mean AUC values and 95% CIs from the 10,000 Monte Carlo simulations for both sets of data.

3.3 RESULTS

Overall, I conducted 1,640,000 Monte Carlo simulations of logistic regression models to evaluate the sensitivity of precision, shifts in regression coefficients, and prediction of statistical models to underreporting of deer- and moose-vehicle collisions. When underreporting occurred randomly for deer-vehicle collisions, I found that the precision of regression coefficients remained relatively stable until $\geq 70\%$ of collisions were unreported (Figure 3.2). After this, the amount of uncertainty doubled. For 2 of 9 variables (CONTAGION and REGISTERED_VEH), the 95% CIs began overlapping zero (i.e., lost statistical significance) when $\geq 90\%$ of deer-vehicle collisions were unreported. The point estimates of the coefficients remained stable up to 95% of deer-vehicle collisions being unreported.

I found similar results for moose-vehicle collisions when underreporting occurred randomly. The precision of regression coefficients was stable until $\geq 70\%$ of collisions were unreported (Figure A.1). For 1 of 10 variables (D_SHBW), the 95% CIs began overlapping zero when $\geq 60\%$ of moose-vehicle collisions were unreported. Five other variables (P_CONIF, IJI, D_DEV, D_STR, and SIDI) lost significance after $\geq 85\%$ of moose-vehicle collisions were unreported. The point estimates of the coefficients remained stable up to 95% of moose-vehicle collisions being unreported.

The 10 least urban counties in central Illinois contained 1,458 deer-vehicle collisions, representing 18% of the total records. When underreporting was spatially biased, I found the point estimates of regression coefficients for 2 of 9 variables (ABUNDANCE in both models) shifted at 40% of deer-vehicle collisions being unreported (Figure 3.3). This shift became

stronger as fewer collisions were reported. The average values of ABUNDANCE were substantially higher in the less urban counties compared to the 40 more urban counties (Table 3.2), suggesting that the abundance of deer was unevenly distributed throughout the study area. Point estimates for 5 other variables also changed slightly as fewer collisions were reported (P_AG, P_FOR, CWED, CONTAGION, REGISTERED_VEH), although these shifts did not change statistical inferences. The average values for these 5 variables were also unevenly distributed throughout the study area (Table 3.2).

Table 3.2 Averaged values for environmental variables in counties used to examine influences on the probabilities of deer-vehicle collisions in central Illinois (2010) and moose-vehicle collisions in western Maine (2000–2010). Biased counties represented counties in which the rates of reporting for collisions were simulated to be lower than unbiased counties.

Central Illinois			Western Maine		
Environmental variable (deer)	Biased counties	Unbiased counties	Environmental variable (moose)	Biased counties	Unbiased counties
TRAFFIC	956	3,301	TRAFFIC	1,357	1,684
CONTAGION	65.5	74.0	D_DEV (m)	3170	2493
CWED (km/km ²)	34.2	18.1	D_SHBW (m)	289	408
ABUNDANCE ^a	1,043	633	D_STR (m)	444	412
P_AG	0.60	0.77	IJI	50.0	54.3
P_FOR	0.29	0.13	P_CONIF	0.27	0.28
P_WAT	0.03	0.01	P_CUT	0.02	0.02
REGISTERED_VEH	16,810	59,291	SIDI	0.59	0.58
			SLOPE (degree) ^a	5.9	8.2
			SPEED (km/hr)	75.8	72.9

^a Represents variables with shifting regression coefficients from logistic regression models as fewer collisions were reported.

TRAFFIC = annual average daily traffic, CONTAGION = contagion index of land-cover and land-use types, CWED = contrast-weighted edge density among land-cover and land-use types, ABUNDANCE = index for abundance of deer from estimated harvest of antlered deer, P_AG = proportion of agriculture, P_FOR = proportion of forest, P_WAT = Proportion of water, REGISTERED_VEH = number of registered vehicles, D_DEV = distance to nearest developed area, D_SHBW = distance to nearest shrub wetland, D_STR = distance to nearest stream, IJI = interspersed/juxtaposition index of land-cover and land-use types, P_CONIF = proportion of conifer forest, P_CUT = proportion of cutover (harvested) forest, SIDI = Simpson's index of

diversity for land-cover and land-use types, SLOPE = degree of slope, SPEED = posted speed limit.

I found similar results for moose-vehicle collisions when underreporting was spatially biased. The least urban county in western Maine contained 229 moose-vehicle collisions, 21% of the total records. Point estimates for 1 of 10 variables (SLOPE) shifted as fewer collisions were reported (Figure A.2). The average degree of slope was substantially lower in the least urban county (Table 3.2), indicating an uneven distribution throughout the study area. The other variables were evenly distributed, thus resulting in point estimates that remained stable.

The predictive capabilities of both good and poor models for deer-vehicle collisions were similarly affected by underreporting of collisions. When underreporting occurred randomly, the estimated AUC values were not affected by underreporting (Figure A.3). The AUC values remained constant for predicting deer-vehicle collisions (i.e., ~0.80 for the good model and ~0.60 for the poor model), and for predicting moose-vehicle collisions (i.e., ~0.87 for the good model and ~0.67 for the poor model). However, imprecision around the AUC values doubled after approximately $\geq 75\%$ of collisions were unreported for deer and moose. When underreporting occurred with spatial bias, the predictive capabilities of the biased models were mostly similar between the new-biased and new-true reports of deer- and moose-vehicle collisions (Figure 3.4). The only simulation in which prediction differed was for the poor model at 100% underreporting of deer-vehicle collisions in the least urban counties. Here, the ability to predict true data was reduced, indicating that the biased model was substantially affected by underreporting.

Figure 3.2 Average parameter estimates and 95% confidence intervals from 10,000 Monte Carlo simulations at different levels of underreporting for deer-vehicle collisions in central Illinois, USA, during 2011. The model with good predictive capability included: TRAFFIC = annual average daily traffic, ABUNDANCE = abundance of deer per county indexed by antlered harvest, P_AG = proportion of agriculture land-cover within 800 m buffer, P_FOR = proportion of forest land-cover within 800 m buffer, CWED = contrast-weighted edge density within 800 m buffer, CONTAGION = contagion index within 800 m buffer. The model with poor predictive capability included: REGISTERED_VEH = the number of registered vehicles per county, ABUNDANCE, and P_WAT = proportion of water land-cover within 800 m buffer.

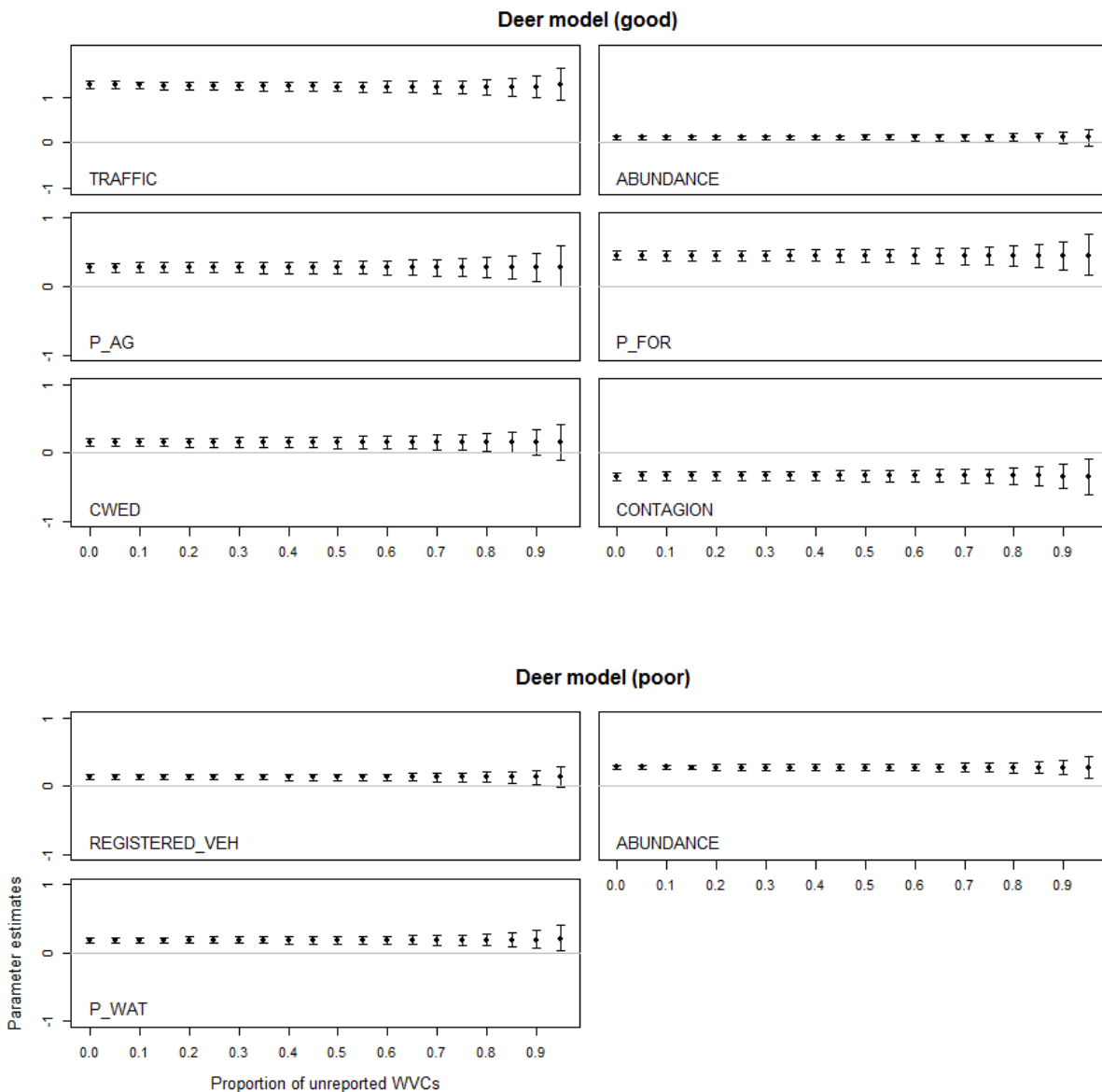


Figure 3.3 Average parameter estimates and 95% confidence intervals from 10,000 Monte Carlo simulations of logistic regression analyses including spatial biases in reporting for different levels of underreporting for deer-vehicle collisions in the most rural counties in central Illinois, USA, during 2011. The model with good predictive capability included: TRAFFIC = annual average daily traffic, ABUNDANCE = abundance of deer per county indexed by antlered harvest, P_AG = proportion of agriculture land-cover within 800 m buffer, P_FOR = proportion of forest land-cover within 800 m buffer, CWED = contrast-weighted edge density within 800 m buffer, CONTAGION = contagion index within 800 m buffer. The model with poor predictive capability included: REGISTERED_VEH = the number of registered vehicles per county, ABUNDANCE, and P_WAT = proportion of water land-cover within 800 m buffer.

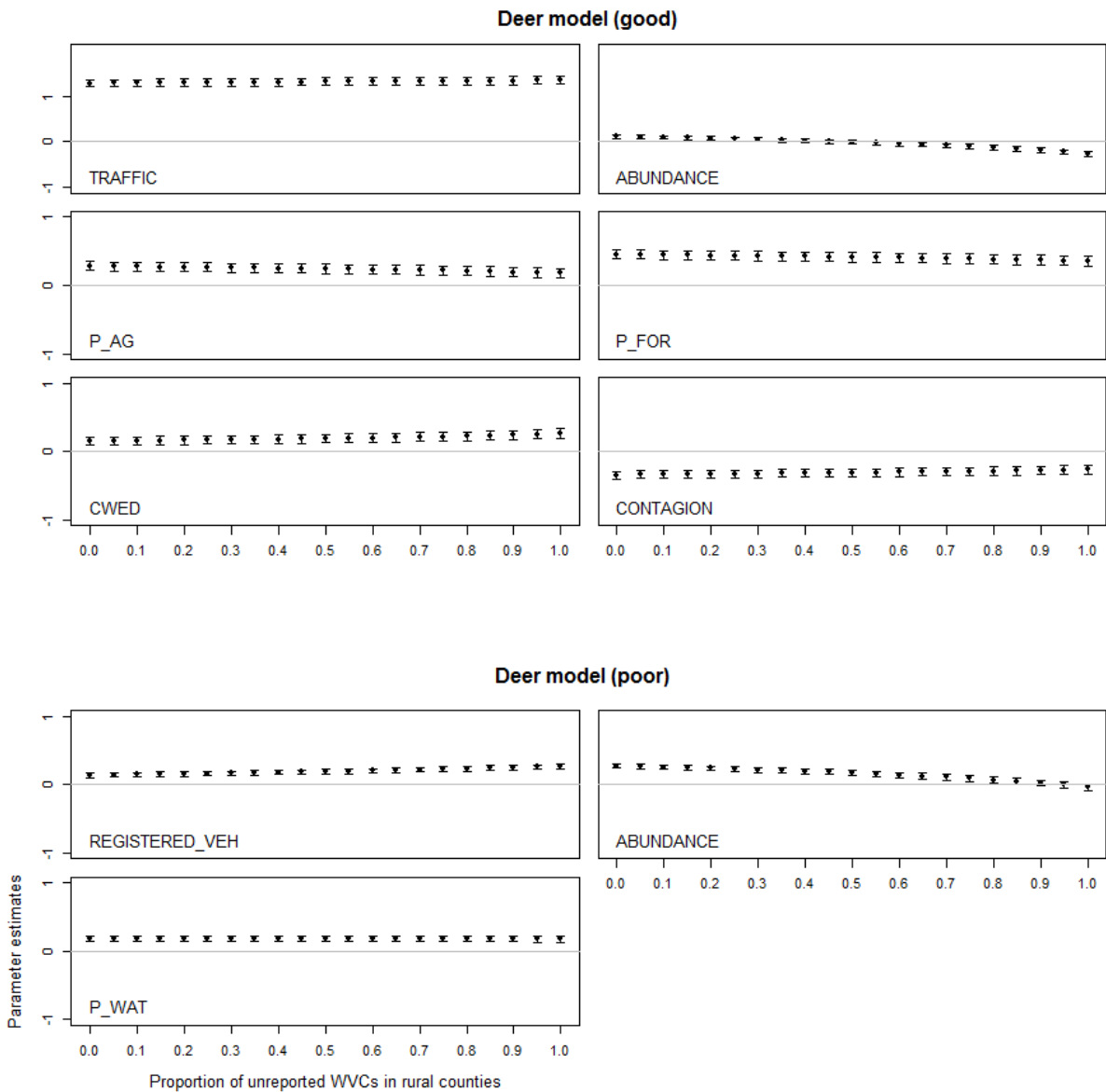
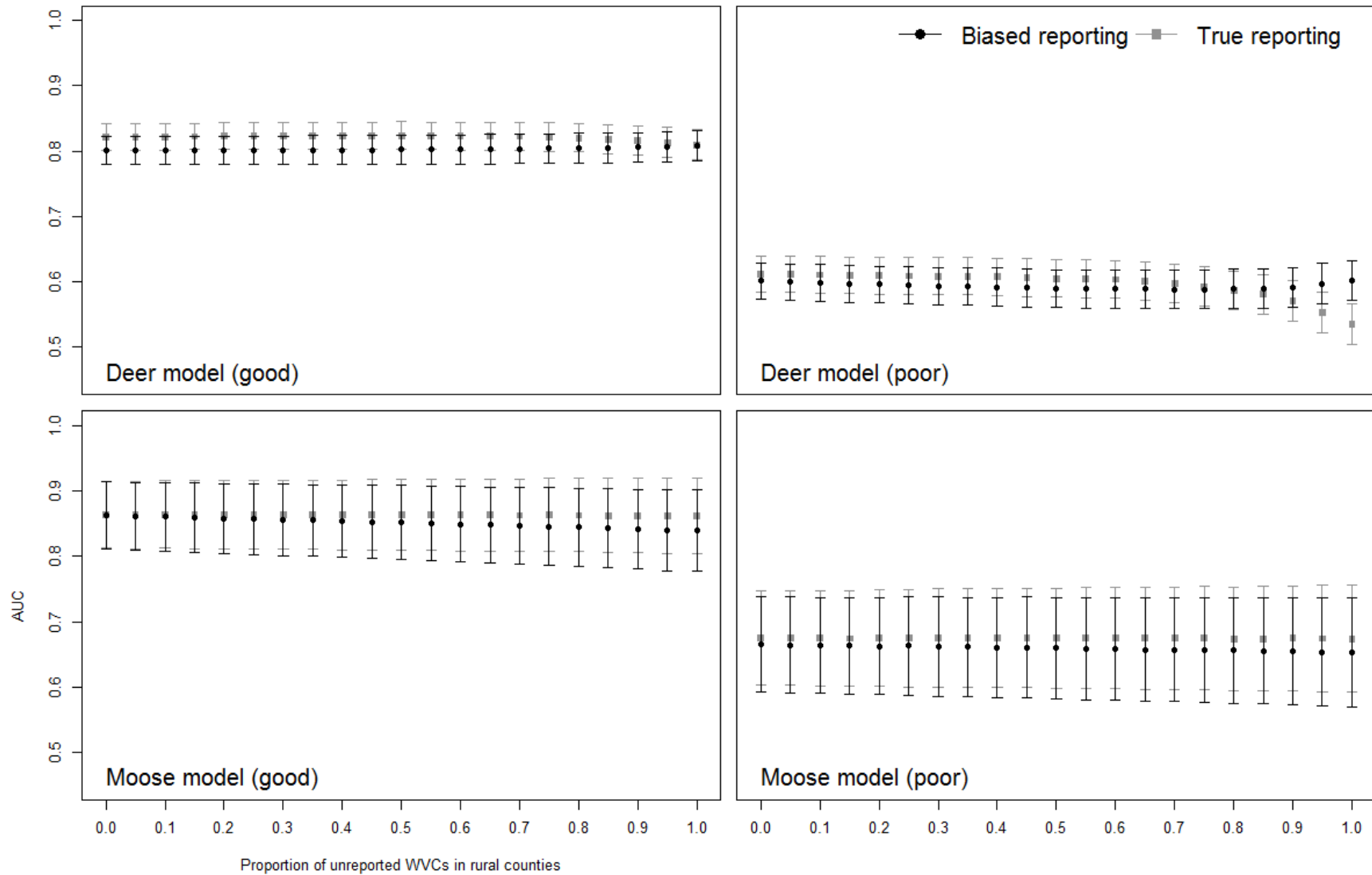


Figure 3.4 Average receiver operating characteristic function (AUC) and 95% confidence intervals for logistic regression analyses including spatial biases in reporting for the least urban counties for deer-vehicle collisions in central Illinois, USA, during 2011 and moose-vehicle collisions in western Maine, USA during 2000–2011. The AUC values were calculated from 10,000 Monte Carlo simulations at different levels of underreporting.



3.4 DISCUSSION

I sought to determine the proportion of WVCs that should be reported in order to support statistical analyses for investigating the relationships between environmental variables and the probabilities of WVCs. I identified similar effects from underreporting for deer-vehicle collisions in central Illinois and moose-vehicle collisions in western Maine. These similarities in results across such large differences in landscape, traffic, movement behavior, and population abundance suggest that the effects of underreporting can be generalized across studies of WVCs. Consequently, I suggest my results are widely applicable. A common feature of WVCs, their tendency to occur in non-random patterns (i.e., hotspots; Huijser et al. 2008, Ng et al. 2008, Danks and Porter 2010), can likely explain these similarities. The existence of hotspots indicates that variation in how environmental variables influence WVCs is low. Therefore, underreporting of WVCs is mostly unproblematic until identification of the influences of hotspots is not possible for most studies of WVCs.

When underreporting occurs randomly, my findings indicate that studies of WVCs are robust to underreporting. Estimates of precision for the regression coefficients were stable until $\geq 70\%$ of WVCs were unreported. If more WVCs were unreported, the small sample sizes caused reduced statistical power (Krebs 1999) and difficulty in detecting some relationships between the environmental variables and the probability of WVCs. Relatively few reports are necessary for identifying the influences of high incidence locations ($\sim 30\%$) because WVCs occur in easily identifiable patterns. For instance, hotspots occur on roads with intense traffic that intersect animal movement corridors (e.g., Ramp et al. 2005, Brockie 2007, Gomes et al. 2009, Grilo et al. 2009), and identifying the environmental variables that influence these hotspots does not require vast amounts of reports. Reporting rates for deer-vehicle collisions are estimated to be 42–50%

(Decker et al. 1990, Romin and Bissonette 1996, Marcoux and Riley 2010), exceeding the 30% threshold I identified. Reporting rates for moose are unknown, but I expect > 30% are reported because collisions with such large animals result in high property damage or human injury.

When reporting rates were simulated to be spatially biased, my findings indicate that models of WVCs are less robust to underreporting. I detected shifts in 2 of 9 estimates of regression coefficients for deer and 1 of 10 estimates for moose, only for variables that were unevenly distributed throughout the study areas. For example, ABUNDANCE was substantially higher (1,043) in the 10 counties I selected to have lower reporting than in the other 40 counties (633). This discrepancy changed the statistical influence of ABUNDANCE, because fewer collisions were reported in regions that had higher abundances of deer. My analysis showed that uneven rates of reporting combined with uneven distribution of environmental variables in $\leq 20\%$ of the study area were enough to result in shifted inferences. However, the shifts were less noticeable at higher rates of reporting. Higher reporting can reduce shifts in regression coefficients, but not eliminate it. Similar types of unreliable inferences have been identified in survey studies from measuring subsets of reports (Groves 2004, Lavrakas 2008). Unreliable inferences are observed when some groups or locations are sampled less, and therefore incomplete information is used in data analyses. Study areas with uneven landscapes (i.e., variable land cover and land use, topography, etc.) are at risk of shifted regression coefficients for WVCs if reporting is also uneven, because uneven reports may not accurately characterize the true environmental influences of collisions.

The performances for good- and poor-predictive models were similarly affected by underreporting. When underreporting was random, the precision of AUC values remained stable until $\geq 75\%$ of WVCs were unreported. When underreporting was spatially biased, the predictive

capabilities of the models were stable (i.e., predicted new-biased and new-true data similarly), indicating a high degree of robustness to spatially biased underreporting. The one exception in which true data were not accurately predicted suggests that the robustness of models is related to the relative importance of each variable. For instance, the relative importance was low for ABUNDANCE in the good model for deer-vehicle collisions, but was high in the poor model. In both cases, the strength and directionality of effects for ABUNDANCE were increasingly shifted as fewer deer-vehicle collisions were reported. Yet, the good model was capable of predicting true data and the poor model was not, when reporting was low. This is because the low relative importance for ABUNDANCE in the good model suggested that ABUNDANCE had little to no impact on the probability of WVCs, even with a high degree of bias in reporting. Therefore, ABUNDANCE had little to no influence for predicting collisions in the good model, relative to the other variables.

I recognize that the biological inferences from my logistic regression models may be afflicted by the same underreporting I evaluated in this study. The data I defined as true for my purposes here (i.e., 100% reporting), were not actually true. I used the best datasets (to my knowledge) for accurately representing the majority of WVCs, and therefore my analysis represents the best relative effects of underreporting. Using this method identified conservative estimates of thresholds because I based the thresholds on the number of reported collisions as opposed to the unknown larger number of true collisions. I also recognize that the WVCs used in this study were relatively common events, thus produced large sample sizes. Underreporting for studies of rare events will likely generate more uncertain inferences, and should be analyzed with caution.

Findings from this research provide useful implications for studies of WVCs. If underreporting is random, I found only 30% of reports were necessary for reliable inferences. Underreporting may be random if reports are generated based on non-space-related criteria. For instance, reports of deer- and moose-vehicle collisions were generated for collisions involving certain amounts of property damage or injury, not by where the collisions occurred. For studies that systematically search roads for WVCs, searches should aim to record $\geq 30\%$ of collisions. If reporting of WVCs is spatially biased, or if the degree of spatial bias is unknown, researchers should examine the distribution of environmental variables to determine whether shifts in regression coefficients might be generated. For unevenly distributed variables, the estimated relationships with WVCs may be unreliable. Increasing the rate of reporting would be helpful to reduce shifts in regression coefficients, but cannot overcome it. Therefore, even rates of reporting throughout the study area are the most efficient methods for generating reliable estimates. This principle should be considered in studies that use reports of WVCs across different municipalities. Some municipalities may not prioritize collecting reports with the same rigor, thus spatial biases in underreporting may exist. Discerning whether reports of WVCs are spatially biased requires in-depth understanding of how each municipality collects the reports.

The current rates of reporting for deer-vehicle collisions in Illinois and moose-vehicle collisions in Maine are sufficient for producing highly reliable findings. For other species, the maximum cost-benefit ratio for producing reliable findings will be achieved by evenly collecting reports at rates of $\geq 30\%$ of collisions. This study design may also be applied to other studies (e.g., survey studies) that use incidence reports to assess risk, particularly if the incidences occur in hotspots. Even reporting, even if rates are low, is sufficient for reliable findings.

3.5 ACKNOWLEDGEMENTS

I thank L. Midden from the Illinois Department of Transportation and D. Brunell from the Maine Department of Transportation for providing databases of deer- and moose-vehicle collisions.

This research was funded in part by the Boone and Crockett Club, the Michigan Department of Natural Resources, the Alces Journal, and the Michigan Involvement Committee of Safari Club International. I thank Michigan State University, the Department of Fisheries and Wildlife, and the Boone and Crockett Quantitative Wildlife Laboratory for support.

DYNAMIC SPACE-TIME ANALYSIS OF WHITE-TAILED DEER-VEHICLE COLLISIONS THROUGHOUT THE MIDWEST UNITED STATES

ABSTRACT

During the last half-century, human changes to landscape structure and wildlife management throughout the United States have dramatically altered densities of white-tailed deer (*Odocoileus virginianus*) and traffic patterns. Today, deer-vehicle collisions (DVCs) represent one of the most persistent, widespread, costly, and dangerous human-wildlife conflicts. Yet, it remains unclear how specific factors influence the frequency of DVCs across the range of deer and through time. I examined how traffic volume, abundance of deer, and land cover influenced DVCs through a 12 year period (2000–2011) across 3 eco-zones (Northern Forest, Forest-Agriculture Matrix, and Agriculture) that intersect 4 Midwestern states ($n = 355$ counties from Iowa, Illinois, Michigan, and Wisconsin). I applied a new, temporally dynamic Bayesian model to assess whether the influences of DVCs changed through time and space. I found that the frequencies of DVCs were influenced by the suburb effect— a unique combination of features generated in suburban environments that coalesce to increase the frequencies of DVCs. These include: 1) intermediate to high traffic, 2) high abundances of deer, and 3) fragmented landscapes (i.e., intermixed and low aggregation of land covers) containing high proportions of agriculture. The suburb effect was strongest in the Forest-Agriculture Matrix where large cities with adjacent outlying suburbs occur. The Agriculture and Northern Forest eco-zones, where suburbanization is less prevalent, experienced similar but weaker influences. I found no changes in factors that influenced DVCs over time, indicating long durations of DVC predictability. DVCs should be expected to increase in all eco-zones where suburbanization occurs unless effective mitigation can be implemented. These findings should be considered in urban and road development planning to minimize the substantial impacts associated with DVCs. This research

provides new justification for incorporating long-term mitigation measures (e.g., underpasses) in infrastructure development or redevelopment in suburban environments to reduce future DVC occurrence for many years. Harvesting more deer in suburban environments will also reduce collisions.

4.1 INTRODUCTION

Vehicular collisions with white-tailed deer (*Odocoileus virginianus*) are one of the most widespread and persistent human-wildlife conflicts throughout the United States (Conover 2010). An estimated >1 million deer-vehicle collisions (DVCs) occur each year in the United States (Conover et al. 1995) and are rising (Huijser et al. 2008). Since 1990, fatal collisions with wildlife (mostly deer) have increased 104% (Sullivan 2011) and DVCs comprised one of the largest sources of economic loss from wildlife, averaging \$6,717 per DVC (Huijser et al. 2008). The Midwest region of the United States records the highest reports of DVCs compared to other regions. Deer-vehicle collisions have become prolific enough that states are prioritizing keeping the frequency of DVCs below publicly tolerable levels. In 2008, Illinois implemented a deer-management objective to keep the rate of DVCs at ≤ 207 DVCs per billion of miles traveled (University of Illinois Extension 2013).

Changes in landscape structure and wildlife management during the past several decades have led to historic levels of population abundance for deer (Coulson 1999) and increases in DVCs. Meanwhile, human populations have increased and dispersed away from centralized cities (i.e., suburbanization; Jordan et al. 1998, Alig et al. 2004, Baum-Snow 2007). Suburban environments reportedly maintain high abundances of deer by providing refuge from hunting (Harden et al. 2005, Lovely et al. 2013). Agricultural activities throughout the Midwest also maintain high densities of deer (Roseberry and Woolf 1998) especially when coupled with

interconnected patches of forests (Walter et al. 2009). Additionally, future changes in the environment from increasing suburbanization (Alig et al. 2004) and climate change (i.e., lower winter severity and reduction in old growth forests; Thompson et al. 1998) may continue to increase densities of deer and DVCs throughout the Midwest region.

I hypothesized that large-scale, environmental characteristics associated with the volume of traffic, abundance of deer, and configuration and composition of the landscape influence the frequencies of DVCs throughout the Midwest. New developments in dynamic process, Bayesian modeling (Gelfand et al. 2005, Finley et al. 2012) have unlocked the ability to examine for these spatiotemporal processes. Dynamic models incorporate temporal processes by using previous information to inform current and future events. An important complement to these models has been the long-term reporting of DVCs, traffic, indices for abundance of deer, and mapping of land use and land cover across the Midwest during the last 2 decades. This combination of dynamic modeling with large datasets provides new opportunities for comprehensive studies of DVCs over large spatial and temporal extents.

Currently, a number of unanswered questions about the influences of DVCs limit our ability to understand their frequencies throughout the range of deer and through time. My objective was to examine the dynamic influences of DVCs and answer 2 important questions: i) is the frequency of DVCs influenced by the dynamic nature of environmental variables through time, and ii) do environmental variables influence the frequency of DVCs differently across space? I attempted to answer these questions by examining the spatiotemporal influences of DVCs in 3 eco-zones (i.e., Northern Forests, Forest-Agriculture Matrix, and Agriculture) throughout 12 years in the Midwest.

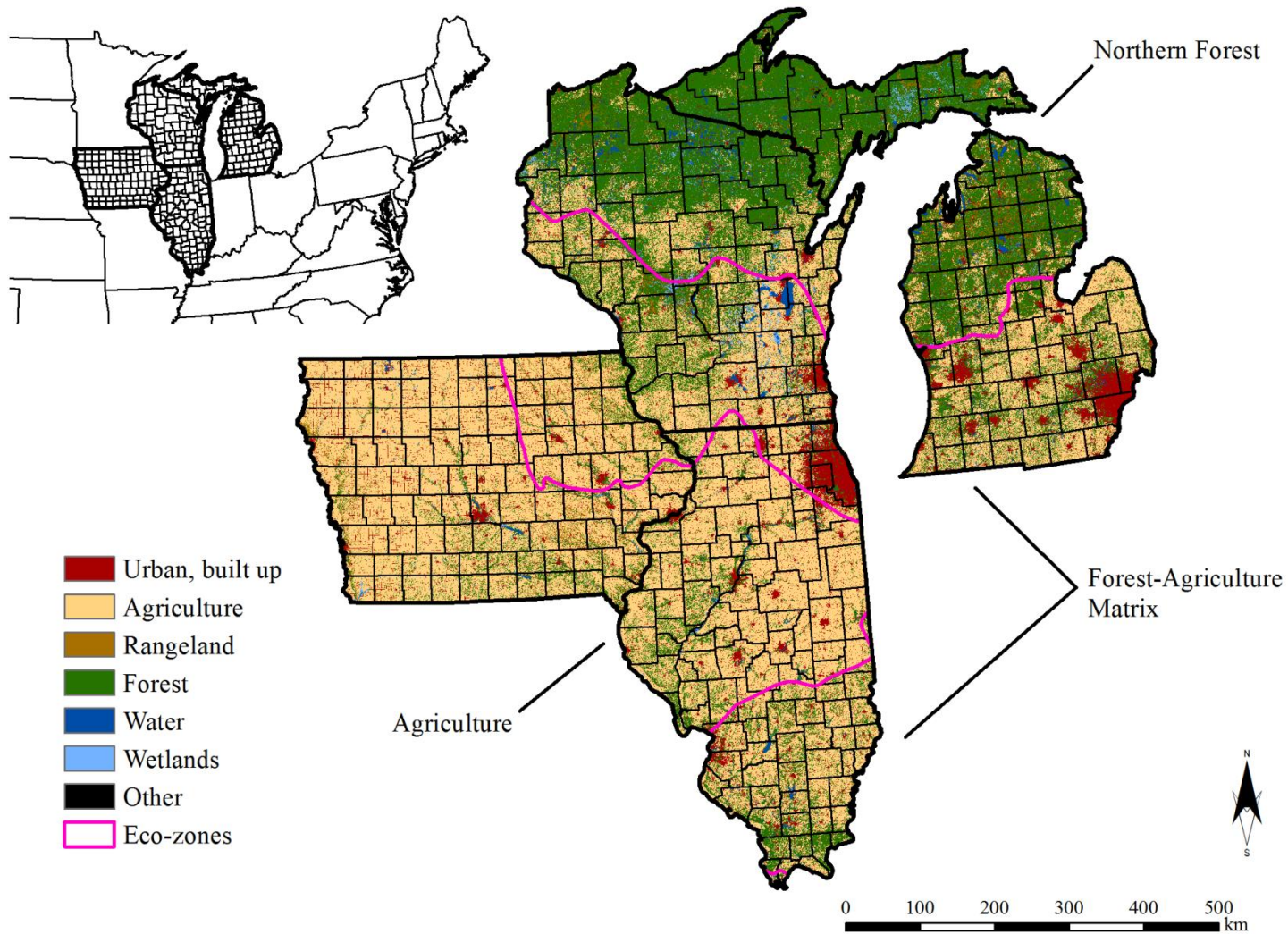
4.2 METHODS

4.2.1 Study area

My study area ($584,493 \text{ km}^2$) was comprised of 355 counties from Illinois, Iowa, Michigan, and Wisconsin (Figure 4.1). Overall, this region included approximately 981,765 km of roads for an overall road density of 1.68 km/km^2 . I divided the region into 3 eco-zones: Northern Forests, Forest-Agriculture Matrix, and Agriculture based on the province divisions of ecoregions for the United States (Bailey 1983;1995). The Northern Forests lies within the Laurentian mixed forest ecosystem province. The land cover of this province was dominated by stands of conifers, deciduous, and mixed conifer-deciduous trees; such as balsam fir (*Abies balsamea*), pines (*Pinus* spp.), spruce (*Picea* spp.), eastern white cedar (*Thuja occidentalis*), maples (*Acer* spp.), yellow birch (*Betula alleghaniensis*), and American beech (*Fagus grandifolia*). The average human density during 2000–2011 was estimated to be 16.5 people/km^2 (United States Census Bureau 2012). The average road density per county was 1.47 km/km^2 . The average density of deer harvested per county during 2000–2011 was 1.30 deer/km^2 (e.g., Fawley 2012). The Forest-Agriculture Matrix ecoregion lies within the eastern broadleaf forest (continental) ecosystem province. The land cover and land use of this province was a mix of oak (*Quercus* spp.)–hickory (*Carya* spp.) stands with maples and American beech trees; and corn, soybean, wheat, and livestock production. The average human density was estimated to be $104.7 \text{ people/km}^2$. The average road density per county was 2.19 km/km^2 . The average density of deer harvested per county was 1.74 deer/km^2 . The Agricultural eco-zone lies within the prairie parkland (temperate)

ecosystem province. This province was dominated by corn, soybean, alfalfa, and livestock production; and intermixed patches of native grasslands. The average human density was estimated to be 27.7 people/km^2 . The average road density per county was 1.33 km/km^2 . The average density of deer harvested per county was 0.58 deer/km^2 .

Figure 4.1 Study area for examination of dynamic, space-time influences of deer-vehicle collisions at the county level throughout the Midwest, USA during 2000–2011.



4.2.1 Data collection

Descriptions and sources of data are provided in Table A.1. All data were compiled annually at the county level for the states of Illinois, Iowa, Michigan, and Wisconsin during 2000–2011. I excluded Menominee County, WI because no reports of DVCs were obtained. Deer-vehicle collisions were considered as the reported numbers of traffic accidents involving deer during each year. Counts of DVCs were biased low because not all collisions were reported. Reporting rates of DVCs have been estimated to be 42–50% (Decker et al. 1990, Romin and Bissonette 1996, Marcoux and Riley 2010). Snow et al. (unpublished report) determined that models regarding the influences of WVCs were robust to random underreporting if >30% of collisions were reported. Therefore, I used the reported number of DVCs per county as relative indices for the true frequency of collisions. Collisions with deer were reported when they resulted in human injury or death, or exceeded a certain amount of property damage. The amounts of property damage varied by state and year (Table A.1), but the average amount of damage from DVCs exceeded the minimum amounts required for reporting in all cases (Huijser et al. 2008).

I compiled records of antlered and antlerless deer harvested in each county each year. No restrictions were placed on the number of individuals that may purchase licenses to hunt antlered deer, which the majority of hunters preferred to harvest over antlerless deer (Fawley and Rudolph 2014). Therefore, in the absence of consistently developed annual deer populations at the county level, the number of antlered deer harvested during year t was used as an index to represent the abundance of deer during year t (ABUNDUNCE). Antlerless permits were more regulated with intent to remove a number of female deer to achieve desired impacts on population abundance (Brown et al. 2000). Therefore, I used the number of antlerless deer

harvested during year $t-1$ as an index for the effect of management on the population of deer during year t (ANTLERLESS_LAG).

I compiled shapefiles depicting locations of all roads in each state to identify all public roads using ArcGIS (v10.1; Environmental Systems Research Institute, Inc., Redlands, CA). I used National Functional Classification categories to exclude private roads from this analysis (Michigan Department of Transportation 2013). Using the Geospatial Modeling Environment program (v0.7.2.1, Spatial Ecology LLC), I calculated the length of roads inside each county, and then calculated the road density for each county as the total length of roads divided by the area of the county (km/km^2). Estimates of annual vehicle miles traveled (TRAFFIC) were used to measure the volume of traffic in each county during each year. Because traffic may be highest in heavily urbanized areas that are mostly devoid of suitable deer habitat, I also examined TRAFFIC^2 to identify any non-linear effects from traffic. The interaction (TRAFFIC x ABUNDANCE) was also considered to examine the interacting relationships between the volume of traffic and the abundance of deer on the frequency of DVCs. Lastly, I compiled the numbers of registered vehicles in each county (VEH_REGISTERED) to provide an index of the number of vehicles on roads.

I used ArcGIS to reclassify the 2001 and 2006 National Land Cover Database (Homer et al. 2007, Fry et al. 2011) from 16 to 7 classes (agriculture, forest, developed, rangeland, wetlands, water, and other; Table 4.1) that represented important land-cover and land-use classes for deer (Anderson et al. 1976). I used the 2001 database to represent the time period 2000–2005, and the 2006 database to represent 2006–2011. I calculated 5 landscape metrics using program FRAGSTATS (v4.1, University of Massachusetts, Amherst) to quantify important land covers and configurations for deer in each county. I calculated the proportions of agriculture (P_AG),

forest (P_FOR), and developed (P_DEV) land covers in each county. The contrast-weighted edge density (CWED) represented the sum of the borders between cover types multiplied by a corresponding contrast-weight (i.e., *weight* = 1 for agriculture, rangeland, and forest cover types, and *weight* = 0 for all other cover types) divided by the area of the county (km/km^2). Contagion was an index of the spatial aggregation and interspersions of similar patch types. A contagion value of 0 represented a highly fragmented and intermixed landscape, whereas a value of 100 represented a landscape comprised of a single patch.

Table 4.1 Proportion of reclassified land-cover and land-use types for 3 eco-zones in the Midwest United States from the 2001 and 2006 National Land Cover Databases.

Class	Northern Forest		Forest-Agriculture Matrix		Agriculture	
	2001	2006	2001	2006	2001	2006
Agriculture	0.16	0.16	0.56	0.56	0.78	0.78
Forest	0.68	0.67	0.24	0.24	0.10	0.10
Developed	0.05	0.05	0.13	0.13	0.08	0.08
Rangeland	0.05	0.05	0.02	0.02	0.03	0.03
Wetlands	0.02	0.02	0.01	0.01	0.00	0.01
Water	0.03	0.03	0.02	0.02	0.01	0.01
Other	0.00	0.00	0.00	0.00	0.00	0.00

4.2.1 Study design and data analysis

I used counties as my spatial units of observation and years as my temporal units of observation. Each county was associated with an annual response variable (count of DVCs) and 10 annual explanatory variables (TRAFFIC, TRAFFIC^2 , ABUNDANCE, ANTLERLESS_LAG, TRAFFIC x ABUNDANCE, VEH_REGISTERED, P_AG, P_FOR, P_DEV, CWED, and CONTAGION). *Post hoc*, I considered the interaction of CONTAGION x ABUNDANCE. I conducted an intercorrelation analysis of the explanatory variables to examine for correlation among variables. I excluded the less biologically interpretable variable(s) from any correlated

pair (i.e., $|r| \geq 0.60$; Program R v2.15.1; R Development Core Team). The remaining variables were used in subsequent models.

I used a Bayesian modeling framework to develop a dynamic, space-time model for each eco-zone to examine the influences of DVCs at the county-level. I used the OpenBUGs program (v3.2.2, Members of the OpenBUGs Project Management Group; Medical Research Council, UK; and Imperial College, UK) and Program R to construct dynamic models that varied by year (Finley et al. 2012). I used the road density (km/km^2) as an offset for each county to account for differences in the total lengths of roads and areas of counties. I incorporated normally distributed random intercepts for each state to account for the differences in reporting among states. I also included normally distributed, temporal random effect to account for the higher than expected variability (i.e., overdispersion) in the distribution of reported DVCs for my Poisson model (Kéry 2010). The full dynamic, Poisson generalized linear mixed model with a log-link was used to estimate the count of DVCs at year (t) in county (c), as:

$$DVC_t(c) = offset + \alpha_{s,t} + x_t\beta_t + \varepsilon_t, \quad t = 1, 2, \dots, N_t;$$

$$s = 1, 2, \dots, N_s$$

where $\alpha_{s,t}$ represented the random intercept at year (t) for state (s). The random time intercept included a Markovian time-dependent process in the specification of the normally distributed priors, so that the prior of the intercept for year (t) was informed by the posterior of year ($t-1$). The x_t were vectors of the environmental variables at year (t). The regression coefficients β_t included a Markovian time-dependent process in the specification of normally distributed priors, so that the prior for the predictor of year (t) was informed by the posterior of year ($t-1$). The ε_t were the temporal-random effects to account for overdispersion for year (t). I examined 4 Monte

Carlo Markov chains of 100,000 iterations with burn-ins of 95,000 iterations. Combined, I made inferences based on 20,000 iterations from 4 converged chains.

I examined the median and 95% credible intervals (CIs) from the distributions of the estimated regression coefficients to identify influences from the environmental variables on the frequency of DVCs. Specifically, I examined the CIs for overlap of zero to ascertain which environmental variables had clear effects on the count of DVCs. I used the final year of the analysis (2011) to examine effect plots for each of the environmental variables. The final year included the most information from the time-evolving priors, and was most pertinent for current management strategies. Effect plots were constructed for the range of data values for each environmental variable in each eco-zone. I used models without the interaction terms to construct effect plots for the main effects of each variable.

I validated the models by randomly withholding 10% of the reported number of DVCs during 2011 from the datasets for each eco-zone. I predicted the number of collisions, and compared to the number of observed DVCs. Finally, I examined residuals from the models for each eco-zone for evidence of spatial autocorrelation using a semivariogram (package *geoR* v 1.7-4; Ribeiro and Diggle 2001).

4.3 RESULTS

Overall, 1,387,948 DVCs were reported during my 12-year study averaging approximately 115,662 DVCs per year within the 3 eco-zones. The average number reported per county was $\bar{x} = 385.3$ (SD = 275.5) in the Northern Forest, $\bar{x} = 471.0$ (SD = 458.6) in the Forest-Agriculture Matrix, and $\bar{x} = 158.9$ (SD = 122.8) Agriculture eco-zone. I found that ABUNDANCE and ANTLERLESS_LAG were highly correlated ($r = 0.83$), therefore I excluded ANTLERLESS_LAG from further analysis. TRAFFIC was highly correlated with

VEH_REGISTERED ($r = 0.83$) and P_DEV ($r = 0.84$), therefore I excluded VEH_REGISTERED and P_DEV from further analysis. CONTAGION and CWED were highly correlated ($r = -0.65$), therefore I excluded CWED from further analysis. Also, P_AG and P_FOR were highly correlated ($r = -0.91$), therefore I excluded P_FOR from further analysis. The remaining explanatory variables I examined were: TRAFFIC, TRAFFIC², ABUNDANCE, TRAFFIC x ABUNDANCE, P_AG, CONTAGION, and CONTAGION x ABUNDANCE).

The estimates of regression coefficients and their 95% CIs for each variable were mostly stable throughout the 12 years for each eco-zone (Figure 4.2). The interaction (TRAFFIC x ABUNDANCE) did not influence the count of DVCs in any eco-zones during any years. The interaction (CONTAGION x ABUNDANCE) had a positive influence on the frequency of DVCs for 1 year in the Northern Forest, 5 years in the Forest-Agriculture Matrix, and 7 years in the Agriculture eco-zone. The 3D effect plots for CONTAGION x ABUNDANCE indicated that the frequency of DVCs increased most strongly in the Forest-Agriculture Matrix where counties had highly fragmented landscapes and high abundances of deer (Figure 4.3). In the other 2 eco-zones, DVCs increased as abundances of deer increased.

I found a negative quadratic effect from TRAFFIC² during 9 years in the Northern Forest, 2 years in the Forest-Agriculture Matrix, and 12 years in the Agriculture eco-zone (Figure 4.2), suggesting that intermediate levels of traffic increased the frequency of DVCs during those years. Otherwise, the frequency of DVCs increased in counties that had higher levels of traffic. The frequencies of DVCs also increased in counties with higher abundances of deer during all years in all eco-zones. The effect plots indicated that the Forest-Agriculture Matrix was most strongly influenced by TRAFFIC and ABUNDANCE, but this eco-zone also experienced the highest values from both variables (Figure 4.4).

The proportion of agriculture had a positive effect on the frequency of DVCs during only 1 year in the Northern Forest, during all 12 years in the Forest-Agriculture Matrix, and during no years in the Agriculture eco-zone (Figure 4.2). CONTAGION had a positive effect during all years in the Northern Forest; and a negative effect during all years in the Forest-Agriculture Matrix and during 1 year in the Agriculture eco-zone. A negative effect from CONTAGION indicated that the frequency of DVCs increased in more fragmented habitats. The effect plots showed that the Forest-Agriculture Matrix was most strongly influenced by the characteristics of the landscapes, although lesser influences were noticeable for the other eco-zones (Figure 4.4).

Model validation indicated that the models predicted the counts of DVCs with reasonable accuracy for all eco-zones, especially for counties with lower counts of DVCs (e.g., <500 DVCs/year; Figure 4.5). Prediction became more variable with higher counts of DVCs, particularly in the Forest-Agriculture Matrix. The residuals from the models for each eco-zone showed no indication of spatial autocorrelation among counties; therefore, no correction for autocorrelation in my models was required.

Figure 4.2 Estimates of regression coefficients and 95% CIs from dynamic models for examining the influences of environmental variables on the frequencies of deer-vehicle collisions (DVCs) at a county level throughout the Midwest, USA during 2000–2011. CONTAGION = an index of fragmentation among land covers per county per year where lower values represent more fragmented landscapes.

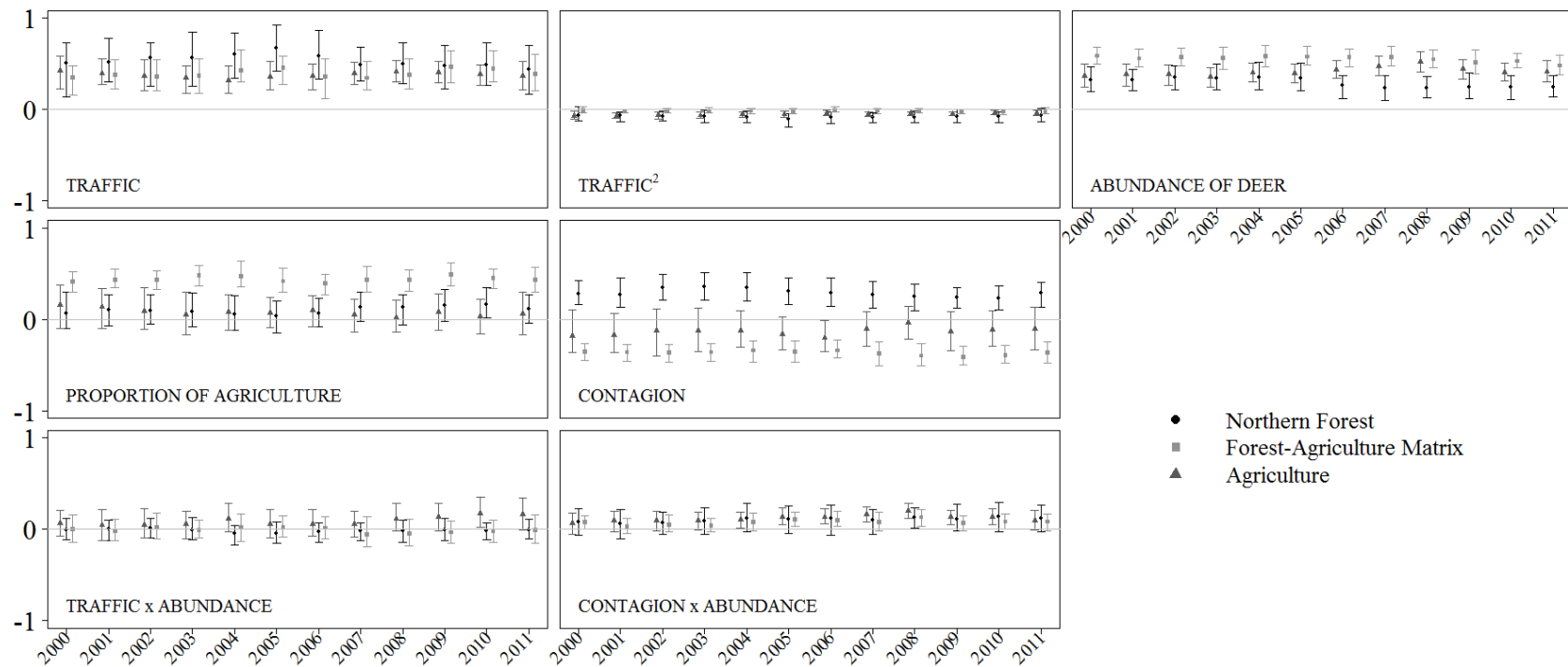


Figure 4.3 Effect plots (medians and 95% CIs) for the interaction term of CONTAGION x ABUNDANCE for examining the influence on the frequency of deer-vehicle collisions (DVCs) in 3 eco-zones in the Midwest, USA during 2011. CONTAGION = an index of fragmentation among land covers per county per year where lower values represent more fragmented landscapes. The density of observed data points are plotted to indicate levels of confidence for the predicted surface. Red = high density of observed points. Blue = low density of observed points. No color = extrapolated effects.

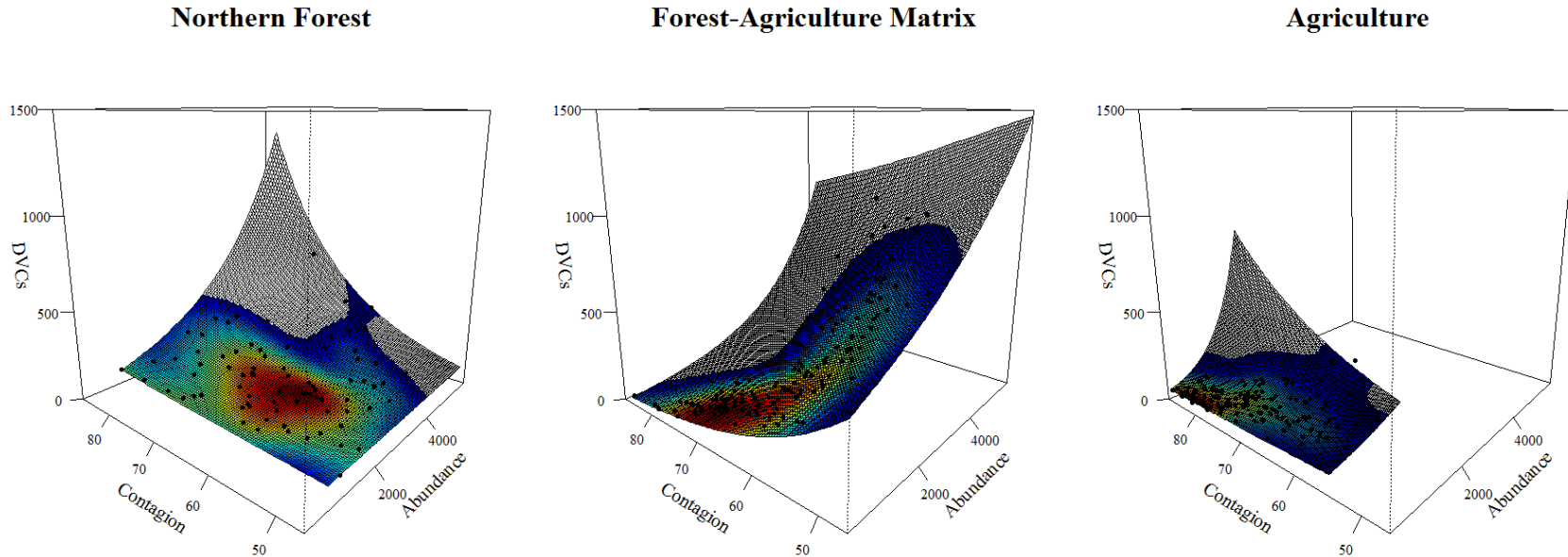


Figure 4.4 Effect plots (medians and 95% CIs) for environmental variables for examining the influences on the frequency of deer-vehicle collisions (DVCs) in 3 eco-zones in the Midwest, USA during 2011. CONTAGION = an index of fragmentation among land covers per county per year where lower values represent more fragmented landscapes.

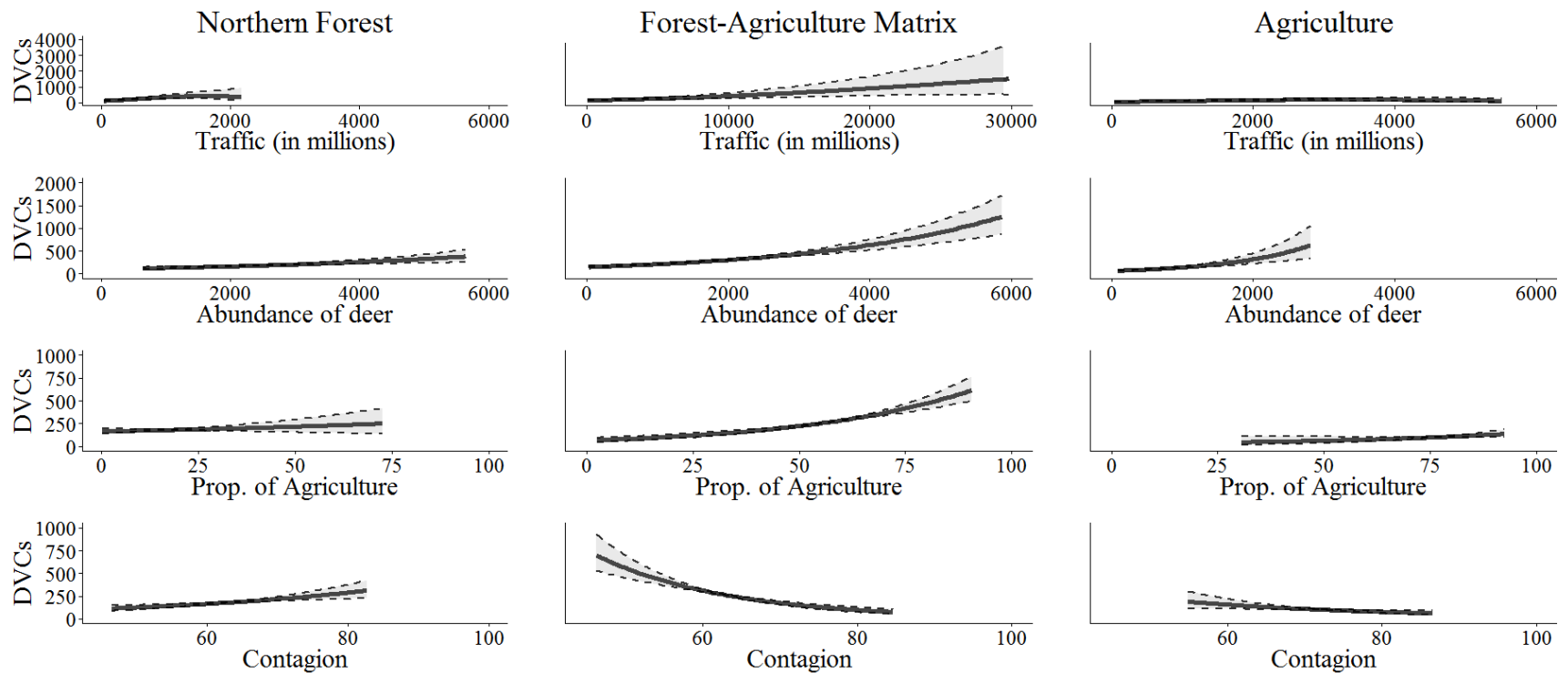
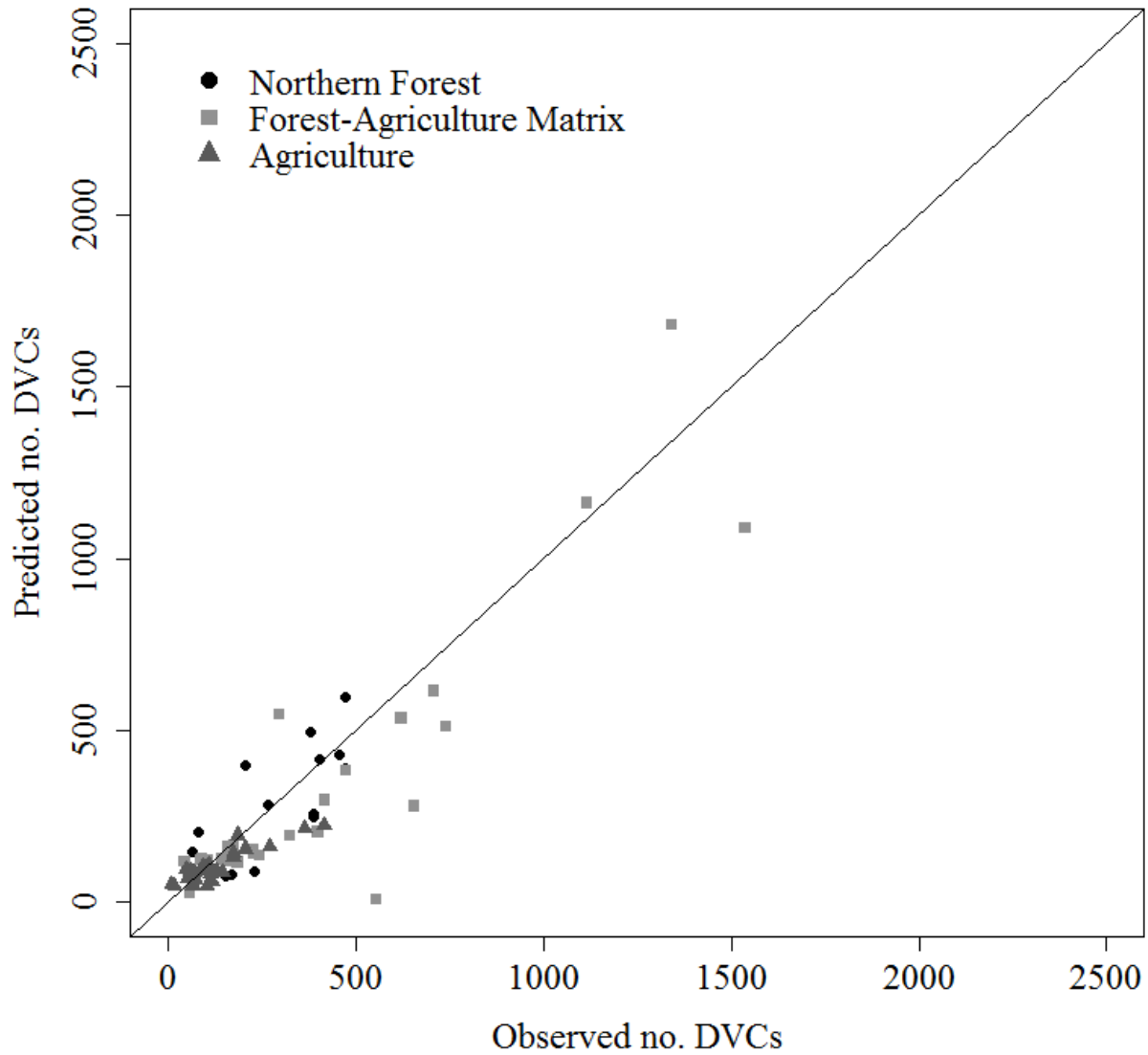


Figure 4.5 Predicted vs. observed number of deer-vehicle collisions (DVCs) for 10% of withheld data for model validation in 3 eco-zones in the Midwest, USA during 2011.



4.4 DISCUSSION

The results from this research signify the first known, empirical explanation for the large-scale processes driving the frequencies of DVCs throughout the Midwest United States. Throughout the Midwest, the highest annual reports of DVCs are in the Forest-Agriculture Matrix eco-zone. My research shows that the frequencies of DVCs are highest where unique combinations of traffic volumes, abundance of deer, and landscape characteristics coalesce to increase the frequencies of DVCs— the suburb effect. A smaller-scale study in Michigan found similar influences on DVCs, but did not connect the process to the suburb effect (Sudharsan et al. 2005). Specifically, I describe the suburb effect as such:

1. Counties with intermediate to high levels of traffic, indicative of the roadways that facilitated suburbanization extending away from cities (Baum-Snow 2007). These roadways connect suburban and urban developments by traversing through mixtures of urban and rural landscapes.
2. Counties with high abundances of deer, consistent with suburbs providing refugia from hunters (Roseberry and Woolf 1998, Lovely et al. 2013).
3. Counties with relatively high proportions of croplands that are not highly aggregated and are intermixed with other land-covers (i.e., fragmented landscape). These attributes of the landscape are relics of converting croplands and forests into developed areas for suburbanization (Harris 1943, Alig et al. 2004). These landscapes provide high quality forage and shelter habitats for deer (Roseberry and Woolf 1998, Walter et al. 2009).

The influence of suburban environments was strongest in the Forest-Agriculture Matrix because suburbanization was most prevalent here. The largest cities were located in this eco-zone (i.e., Chicago, Detroit, Flint, Grand Rapids, Lansing, Madison, Milwaukee) and suburbanization

extend out from city centers (Baum-Snow 2007). According to my findings, areas where DVCs are most common are within these suburban environments. The effect plots indicated that the highest frequencies of DVCs were observed in the Forest-Agriculture Matrix where counties had >10,000 million vehicle miles traveled in a year, in counties with >4,000 antlered deer harvest in a year, in counties with >50% agriculture, and in counties with a contagion index of <60. These empirical results provide new information for urban planners, transportation managers, and natural resource managers to identify regions that experience the highest frequencies of DVCs and identify the root causes of those frequencies.

The Agriculture eco-zone had fewer cities and a lower density of humans because most of the landscape was designated as working farms. The influences of DVCs in this eco-zone were similarly indicative of suburban environments, but with less strong effects than the Forest-Agriculture Matrix. The Agriculture eco-zone had similar human populations and traffic as a prior study in Arkansas that identified urban components as increasing DVCs, but did not connect the process to the suburb effect (Farrell and Tappe 2007). The Northern Forest eco-zone had the lowest density of humans, indicating the least amount of suburbanization. The proportion of agriculture was not important for increasing DVCs in this eco-zone because it contained the least amount of agriculture, mostly located in counties in the southern portion of the eco-zone (Figure 3.2). In the Northern Forest eco-zone, counties with more fragmented landscapes had lower counts of DVCs which was contrary to the relationships found for the other eco-zones. This was likely the case because fragmentation was more closely tied to the presence of wetlands and water in the Northern Forest eco-zone. Fewer roads and deer exist in environments with more wetlands and water.

Higher abundances of deer represented the strongest relationship for increasing DVCs among all eco-zones. Similar trends have been identified elsewhere (Hothorn et al. 2012). A study in Illinois found that lethal control (i.e., sharpshooting) of white-tailed deer reduced DVCs in a localized area (DeNicola and Williams 2006). This research indicates that reducing deer abundance throughout suburban environments may reduce DVCs on larger scales (e.g., DeNicola et al. 2000). Hunting deer in suburban environments could present opportunities for reduction to occur on larger scales.

My finding suggests that emphasis should also be placed on other components of the suburb effect to reduce DVCs. Traffic, agriculture, and partitioning of the landscape influenced the frequencies of DVCs. Counties where the frequencies of collisions were highest would benefit from other types of mitigation that reduce collisions, such as fences and underpasses (Clevenger et al. 2001, Glista et al. 2009). The need to consider these components together may represent a primary reason why few simple options exist for reducing DVCs on large scales (e.g., Huijser et al. 2008).

The influences of collisions were stable throughout the Midwest United States during the last 12 years. This demonstrates the existence of spatiotemporal hotspots for DVCs. Other studies have identified spatial hotspots of wildlife-vehicle collisions, demonstrated by their tendency to occur in non-random, spatial patterns (Huijser et al. 2008, Ng et al. 2008, Danks and Porter 2010, Snow et al. 2014). However, no studies have identified this predictability through both space and time. Likely, my 12-year study was too short to notice new effects from suburbanization. Also, the available land-cover and land-use data for this region were limited to 2 time periods (2001 and 2006), and may have limited the ability to detect influences from metrics of the landscape. Suburbanization has occurred throughout the United States during

much of the last century (Harris 1943), and the process likely takes >12 years to generate noticeable effects at the county scale. Also, the economic recession of 2007-2008 restricted new construction (Stock and Watson 2012), limiting my ability to detect changes from new development. These findings indicate that DVCs should be expected to increase in all eco-zones where suburbanization occurs unless effective mitigation can be implemented.

Likely, the suburb effect is not exclusively associated with deer. Other species of wildlife exist in suburban and urban environments and are affected by collision with vehicles. Coyotes (*Canis latrans*) inhabit urban environments and are commonly killed by vehicles (Grinder and Krausman 2001, Gehrt 2004). Turtles and other reptiles and amphibians inhabit ponds and drainages from suburban developments (Spinks et al. 2003, Conner et al. 2005). The distribution of Virginia opossums (*Didelphis virginiana*) has been linked to the characteristics of suburbia (Kanda et al. 2009). Collisions with these species are typically unreported, but anecdotal evidence suggests that they occur regularly.

A limitation of the current statistical approach is the inability to explicitly account for spatial structure among the counties for each eco-zone. Accounting for spatial autocorrelation would improve the performance of this model for predicting the counts of DVCs (Banerjee et al. 2004, Finley et al. 2009). Spatial autocorrelation can be captured most effectively using hierarchical models (e.g., Finley et al. 2007) and represents an important line of future research for understanding DVCs.

4.5 MANAGEMENT IMPLICATIONS

The multiple components of the suburb effect in counties (i.e., higher traffic, higher abundance of deer, higher proportion of agriculture, more fragmentation) imply that reducing DVCs on large scales requires multifaceted tactics. Reducing the abundance of deer in suburban

environments will be an effective option because fewer deer will be on heavily trafficked roads. Any means of reducing deer (e.g., hunting or sharpshooting) should reduce collisions. However, the long term continuity in spatiotemporal hotspots provides opportunities for more targeted mitigation. Underpasses with high-fencing along the roadsides is an effective option (e.g., Clevenger et al. 2001), but requires substantial financial investment (Reed et al. 1982, Huijser et al. 2009). The spatiotemporal predictability provides justification for constructing more of these permanent structures to reduce collisions for the long-term. A cost-effective approach will be strategically placing these structures on roads that transport commuters between their suburban homes and urban destinations (i.e., locations where traffic and fragmentation are highest). Urban planning for new development and road construction (or reconstruction) projects should consider mitigating future collisions that result from new development.

4.6 ACKNOWLEDGEMENTS

I thank T. Litchfield from the Iowa Department of Natural Resources, T. Micetich from the Illinois Department of Natural Resources, J. Sumners from the Missouri Department of Conservation, and R. Rolley and D. Storm from the Wisconsin Department of Natural Resources for cooperation and providing data for this project. I also thank personnel from the Illinois Department of Transportation, Iowa Department of Transportation, Michigan Department of Transportation, Missouri Department of Transportation, Wisconsin Department of Transportation, Illinois Secretary of State, Michigan State Police, and Missouri State Highway Patrol for compiling and providing data. This research was funded in part by the Michigan Department of Natural Resources, the Boone and Crockett Club, the Alces Journal, and the Michigan Involvement Committee of Safari Club International. I thank Michigan State

University, the Department of Fisheries and Wildlife, and the Boone and Crockett Quantitative Wildlife Laboratory for support.

EPILOGUE

I sought to address 3 deficiencies that exist in current research studies for understanding the influences of WVCs. My goal was to provide more objective and realistic approaches for understanding the influences on WVCs, and thereby enhance our ability to effectively manage them. In Chapter 1, I addressed the issue of inconsistency and subjectivity in delineating hotspots of WVCs. In Chapter 2, I addressed the issue of underreporting and its effects on inferences made from logistic regression modeling regarding influences of WVCs. In Chapter 3, I addressed the issue of understanding the dynamic influences of WVCs over large spatial and temporal extents.

The collective works in these chapters contribute 3 primary conclusions for better understanding ecological relationships for WVCs. First, attributes of the landscape surrounding locations of WVCs can be used to objectively delineate hotspots. This new approach indicates that hotspots are larger than previously reported. Second, analyses of WVCs are highly robust to underreporting likely because WVCs occur in highly predictable patterns (i.e., hotspots). Therefore, relatively few reports are required for reliably understanding the environmental influences on where hotspots occur. Third, the large geographic and temporal drivers that increase the frequency of DVCs are driven by suburbanization. The suburb effect consists of a unique combination of intermediate to high traffic volume, high abundances of deer, and a highly fragmented landscape with high proportions of croplands. These influences did not change through the last 12 years, indicating high spatiotemporal predictability for DVCs.

Chapter 1 addressed the issue of subjectivity in delineating hotspots of wildlife-vehicle collisions. The current methods for delineating hotspots are inconsistent and rely on subjective choices by the researchers. This inconsistency can generate non-reproducible results and possibly

pseudoreplicated hotspots. I developed an objective approach for delineating hotspots that can be used across regions and species, using variation in the landscape. The landscape-based approach provides an improvement over currently used methods because it is informed by the ecological processes that influence where hotspots exist, and not by human perceptions about the scales of hotspots.

Chapter 2 addressed the issue of underreporting in studies of WVCs. Large proportions of WVCs are unreported; therefore studies of WVCs typically rely on incomplete data. I identify thresholds in the numbers of WVCs that need to be reported to obtain reliable information about the environmental influences of collisions. My simulations indicated that studies of WVCs are highly robust to underreporting. Only >30% of WVCs are needed to produce reliable, statistical inferences. Shifts in inferences were detected only if the rate of reporting and the environmental variables were unevenly distributed throughout the study area. I speculated that robust assessments of WVC incidence are generated with relatively few reports because they tend to occur in non-random patterns (i.e., hotspots) and variability in their influences is low. However, the most reliable inferences will be produced where the rates of reporting are consistent throughout the study region, even with low rates of reporting.

Chapter 3 addressed the lack of understanding in large geographic and temporal influences on the frequencies of DVCs. This lack of understanding reduces our ability to broadly manage DVCs across large extents and plan for future changes in their frequencies. My findings suggested that DVCs are associated with a suburb effect— a unique combination of features generated in suburban environments that coalesce to increase the frequencies of DVCs. The influences of collisions were relatively stable throughout the Midwest United States during the last 12 years. These findings indicate that DVCs should be expected to increase in all eco-zones

where suburbanization occurs unless effective mitigation can be implemented. The suburb effect provides the first empirical explanation for why the frequencies of DVCs vary spatially across the Midwest.

The findings from these chapters satisfy the impetus for this dissertation. I sought to improve our ability to understand where wildlife-vehicle collisions occur. The collective works suggest that understanding hotspots is critical for effectively managing WVCs. The landscape-based approach for delineating hotspots indicates that they are larger than previously known. The existence of these hotspots allows for analyzing fewer reports of collisions to determine the ecological relationships associated with the locations of collisions. Environmental influences on collisions are unlikely to change through time, but examining them across large geographic and temporal scales can inform expected trends for the future. Increases in suburbanization will likely expand the size and intensity of hotspots, particularly for deer-vehicle collisions.

The implications of my results suggest that identifying the most critical locations to mitigate can be accomplished with relatively few reports of collisions if collected in a consistent manner. Large hotspots associated with suburban landscapes account for the highest frequencies of collisions, therefore these locations should be targeted for mitigation. Managers should consider investing in long-term mitigation strategies (i.e., underpasses) to reduce WVCs for many years, because the influences of hotspots do not change.

APPENDIX

Figure A.1 Average parameter estimates and 95% confidence intervals from 10,000 Monte Carlo simulations at different levels of underreporting for moose-vehicle collisions in western Maine, USA, during 2000–2011. The model with good predictive capability included: TRAFFIC = annual average daily traffic, SPEED = speed limit (km/hr.), P_CUT = proportion of cutover forest land-cover within 2,500 m buffer, P_CONIF = proportion of conifer forest land-cover within 2,500 m buffer, IJI = interspersed-juxtaposition index within 5,000 m buffer, D_DEV = distance to development (m), D_SHBW = distance to shrub-wetland land-cover (m). The model with poor predictive capability included: SLOPE = degree of slope, D_STR = distance from stream (m), SIDI = Simpson’s diversity index within 5,000 m buffer.

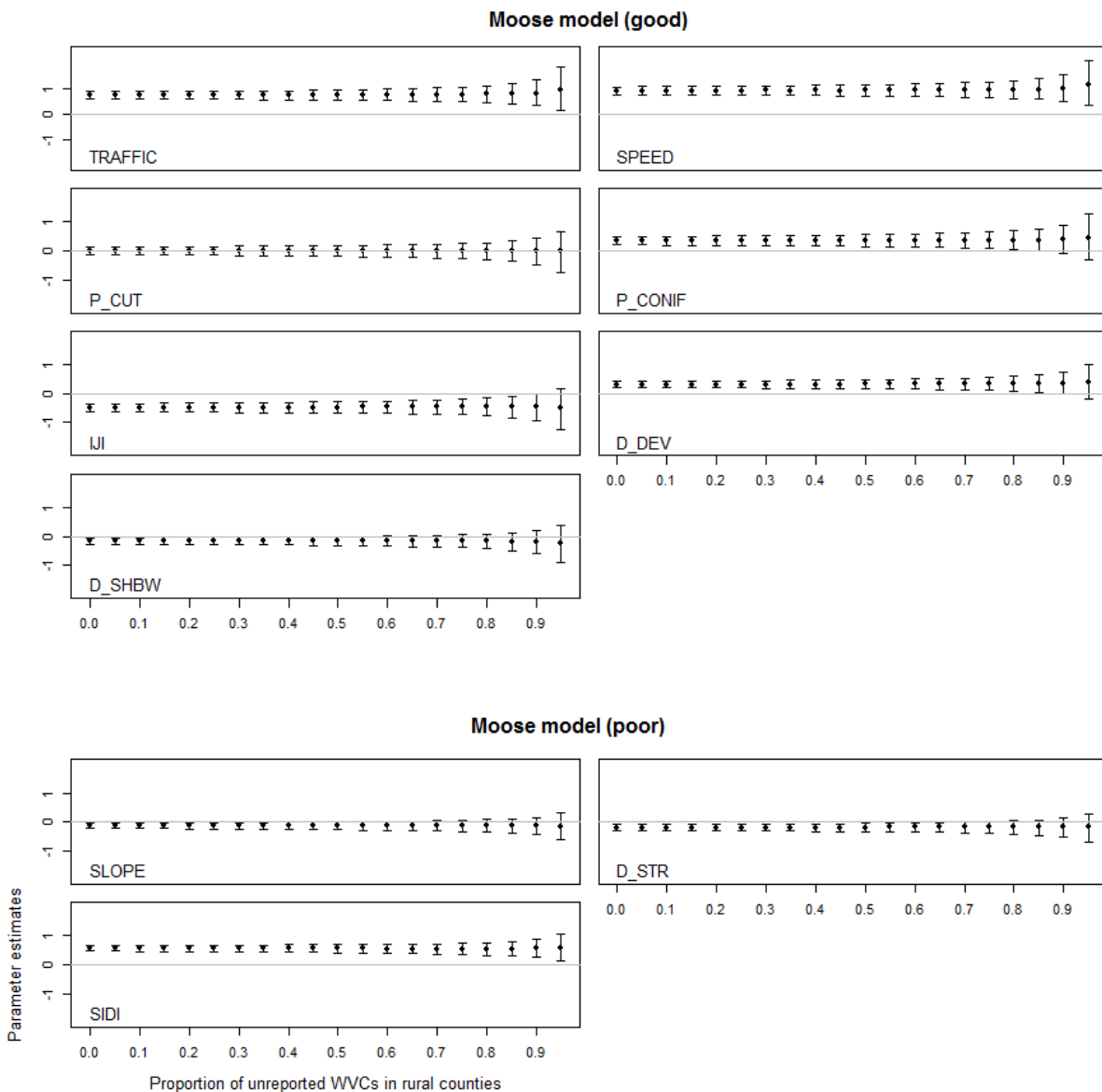


Figure A.2 Average parameter estimates and 95% confidence intervals from 10,000 Monte Carlo simulations of logistic regression analyses including spatial biases in reporting for different levels of underreporting for moose-vehicle collisions in the most rural county in western Maine, USA, during 2000–2011. The model with good predictive capability included: TRAFFIC = annual average daily traffic, SPEED = speed limit (km/hr.), P_CUT = proportion of cutover forest land-cover within 2,500 m buffer, P_CONIF = proportion of conifer forest land-cover within 2,500 m buffer, IJI = interspersed-juxtaposition index within 5,000 m buffer, D_DEV = distance to development (m), D_SHBW = distance to shrub-wetland land-cover (m). The model with poor predictive capability included: SLOPE = degree of slope, D_STR = distance from stream (m), SIDI = Simpson’s diversity index within 5,000 m buffer.

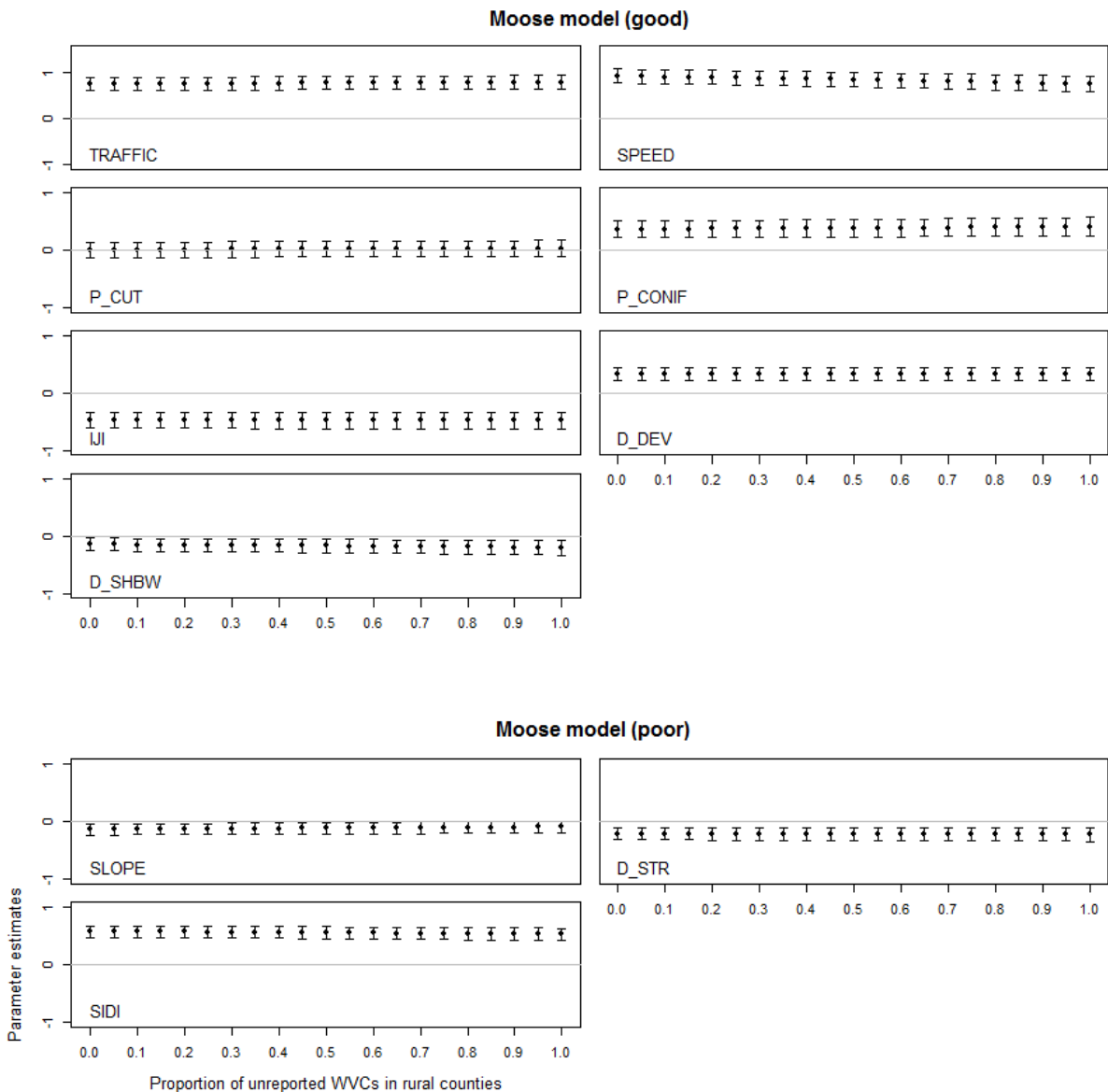


Figure A.3 Average receiver operating characteristic function (AUC) and 95% confidence intervals at different levels of underreporting to depict the predictive capabilities of 4 models as fewer WVCs were reported. The AUC values were calculated from 10,000 Monte Carlo simulations at different levels of underreporting for deer-vehicle collisions in central Illinois, USA during 2001 and moose-vehicle collisions in western Maine, USA during 2000–2011.

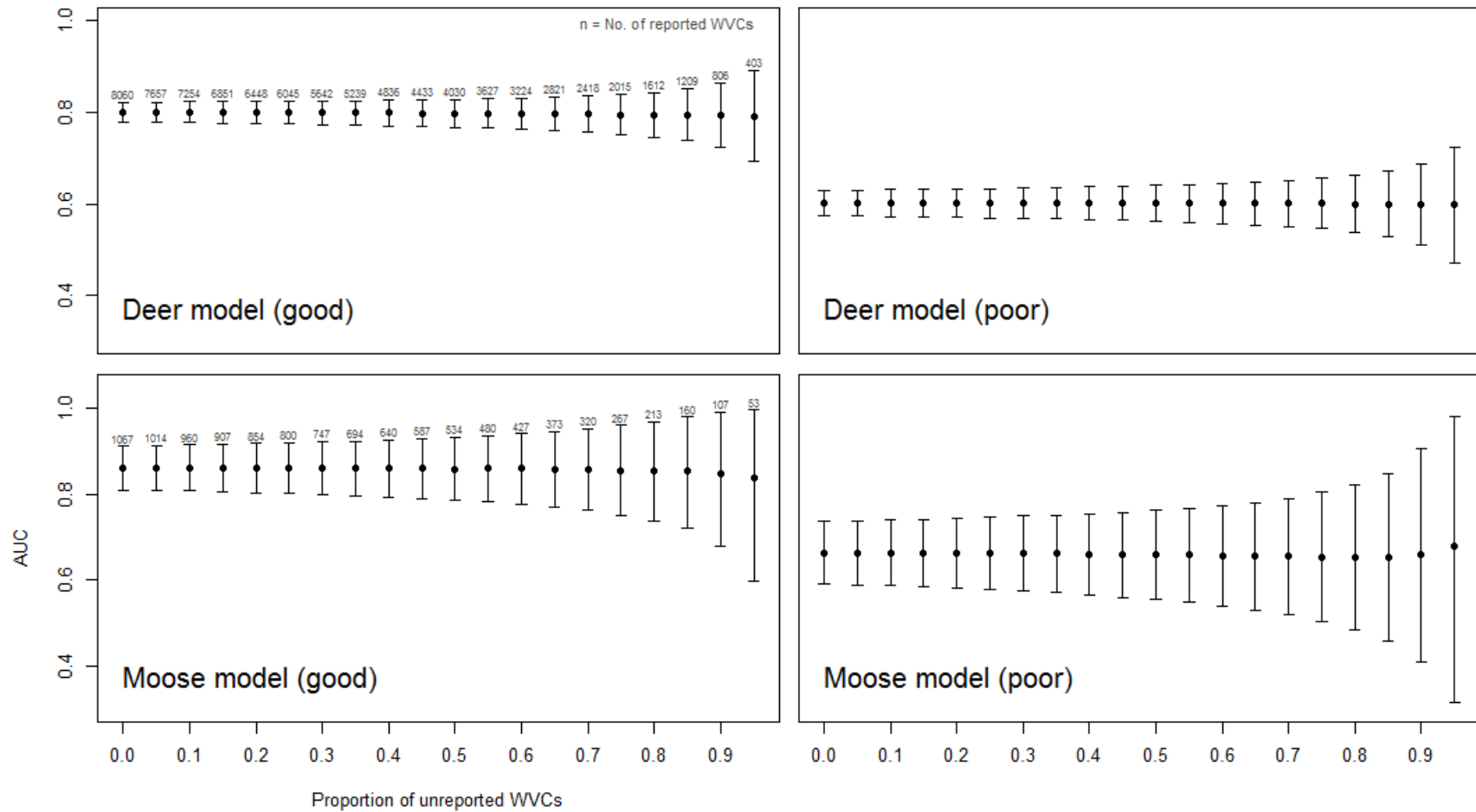


Table A.1 Description and sources of data used to examine the dynamic, space-time influences of deer-vehicle collisions (DVCs) throughout the Midwest, USA during 2000–2011.

Location	Data ^a	Description	Data source
ILLINOIS	DVCs	Vehicle crashes involving deer with >\$500 in property damage (2000-2009), >\$1,500 for insured drivers and >\$500 for uninsured drivers (2009-2012), or bodily injury	Illinois Department of Transportation
	Deer harvest	Estimated number of antlerless and antlered deer harvested by firearm and archery	Illinois Department of Natural Resources
	Traffic	Estimated annual vehicle miles traveled on roads by all vehicles	Illinois Department of Transportation
	Registered vehicles	Total number of registration counts	Illinois Secretary of State
	Roads	2011 shapefile of roads with federal functional classification representing public roads	Illinois Department of Transportation
IOWA	DVCs	Vehicle crashes involving deer with >\$1,000 in property damage (2000- June 2010), >\$1,500 (July, 2010-2012), or bodily injury	Iowa Department of Natural Resources
	Deer harvest	Estimated number of antlerless and antlered deer harvested by firearm and archery	Iowa Department of Natural Resources
	Traffic	Estimated annual vehicle miles traveled on roads by all vehicles	Iowa Department of Transportation
	Registered vehicles	Total number of registration counts	Iowa Department of Transportation
	Roads	2006 shapefile of roads with federal functional classification representing public roads	Iowa Department of Natural Resources Geographic Information Systems Library

Continued on next page

Table A.1 (cont'd)

Location	Data ^a	Description	Data source
MICHIGAN	DVCs	Vehicle crashes involving deer with >\$400 in property damage (2000-2003), >\$1,000 (2003-2012), or bodily injury	Michigan State Police
	Deer harvest	Estimated number of antlerless and antlered deer harvested by firearm and archery	Michigan Department of Natural Resources
	Traffic	The estimated total number of miles traveled annually by motor vehicles on Michigan trafficways	Michigan Department of Transportation
	Registered vehicles	Total number of registration counts excluding trailers and trailer coaches	Michigan State Police, Office of Highway Safety Planning
	Roads	2012 shapefile of roads with national functional classification representing public roads	Michigan Center for Geographic Information
WISCONSIN	DVCs	Vehicle crashes involving deer with >\$1,000 in property damage (2000-2012) or bodily injury	Wisconsin Department of Transportation
	Deer harvest	Estimated number of antlerless and antlered deer harvested by firearm and archery	Wisconsin Department of Natural Resources
	Traffic	The estimated total number of miles traveled annually by motor vehicles on Wisconsin trafficways	Wisconsin Department of Transportation
	Registered vehicles	Total numbers of current and non-expiring registrations	Wisconsin Department of Transportation, Bureau of Vehicle Services
	Roads	2013 shapefile of roads with roadway categories representing public roads.	Wisconsin Department of Transportation

Continued on next page

Table A.1 (cont'd)

Location	Data ^a	Description	Data source
REGION	County boundaries	2010 seamless national file with no overlaps or gaps between parts, designed to stand alone as an independent data set, or can be combined to cover the entire nation	U.S. Census Bureau (TIGER/Line Shapefile)
	Land-cover and land-use maps	30 m resolution land use and land cover maps for the conterminous United States generated from remote sensing with 79% (2001) and 78% (2006) overall accuracy	U.S. Geologic Survey (2001 and 2006 National Land Cover Database)

^aThe response variable was the number of deer vehicle collisions (DVCs) per county/yr. Explanatory variables were the numbers of antlered deer harvested per county/yr. as an index of deer population, annual vehicle miles traveled (TRAFFIC) per county/yr. as an index of traffic volume, number of registered vehicle per county/yr. as an index of the number of motorists, maps of all public roads for each county to calculate the densities of roads per county (km/km^2), maps of the boundaries of counties, and maps of the land cover and land use for each county.

LITERATURE CITED

LITERATURE CITED

- Alig, R. J., J. D. Kline, and M. Lichtenstein. 2004. Urbanization on the US landscape: looking ahead in the 21st century. *Landscape and Urban Planning* 69:219–234.
- Allen, A. W., J. W. Terrell, and P. A. Jordan. 1987. Habitat suitability index models: moose, Lake Superior region.
- Allen, A. W., J. W. Terrell, and P. A. Jordan. 1988. An overview of a habitat suitability index model for moose: Lake Superior region. *Alces* 24:118–125.
- Anderson, J., E. Hardy, J. Roach, and R. Witmer. 1976. A land use and land cover classification scheme for use with remote sensing data. United States Government Printing Office, Washington D.C., USA.
- Bailey, R. G. 1980. Description of the ecoregions of the United States. US Dept. of Agriculture, Forest Service, Washington D.C., USA.
- Bailey, R. G. 1983. Delineation of ecosystem regions. *Environmental Management* 7:365–373.
- Bailey, R. G. 1995. Description of the ecoregions of the United States. 2nd edition. Forest Service Miscellaneous Publication 1391 (rev.), US Department of Agriculture, Forest Service, Washington D.C., USA.
- Bailey, T. C., and A. C. Gatrell. 1995. Interactive spatial data analysis. Volume 413. Longman Scientific & Technical, Essex, UK.
- Banerjee, S., B. P. Carlin, and A. E. Gelfand. 2004. Hierarchical modeling and analysis for spatial data. Volume 101. Chapman & Hall/CRC, Boca Raton, FL, USA.
- Baum-Snow, N. 2007. Did highways cause suburbanization? *The Quarterly Journal of Economics* 122:775–805.
- Beyer, H. L. 2012. Isopleth. <http://www.spatial ecology.com/gme/isopleth.htm> Accessed February 01, 2014.
- Brockie, R. 2007. Notes on New Zealand mammals 4. Animal road kill “blackspots”. *New Zealand Journal of Zoology* 34:311–136.
- Brown, T. L., D. J. Decker, S. J. Riley, J. W. Enck, T. B. Lauber, P. D. Curtis, and G. F. Mattfeld. 2000. The future of hunting as a mechanism to control white-tailed deer populations. *Wildlife Society Bulletin* 28:797–807.

- Campa, H., S. J. Riley, S. R. Winterstein, T. L. Hiller, S. A. Lischka, and J. P. Burroughs. 2011. Changing landscapes for white-tailed deer management in the 21st century: Parcelization of land ownership and evolving stakeholder values in Michigan. *Wildlife Society Bulletin* 35:168–176.
- Clevenger, A. P., B. Chruszcz, and K. E. Gunson. 2001. Highway mitigation fencing reduces wildlife-vehicle collisions. *Wildlife Society Bulletin* 29:646–653.
- Clevenger, A. P., A. Hardy, and K. Gunson. 2006. *Analyses of Wildlife-Vehicle Collision Data: Applications for Guiding Decision-Making for Wildlife Crossing Mitigation and Motorist Safety*. Utah State University.
- Conner, C. A., B. A. Douthitt, and T. J. Ryan. 2005. Descriptive ecology of a turtle assemblage in an urban landscape. *The American midland naturalist* 153:428–435.
- Conover, M. R. 2010. *Resolving human-wildlife conflicts: the science of wildlife damage management*. CRC press, Boca Raton, Florida, USA.
- Conover, M. R., W. C. Pitt, K. K. Kessler, T. J. DuBow, and W. A. Sanborn. 1995. Review of human injuries, illnesses, and economic losses caused by wildlife in the United States. *Wildlife Society Bulletin* 23:407–414.
- Coulson, T. 1999. The science of overabundance: deer ecology and population management. *Biodiversity and Conservation* 8:1719–1721.
- Danks, Z. D., and W. F. Porter. 2010. Temporal, spatial, and landscape habitat characteristics of moose-vehicle collisions in western Maine. *Journal of Wildlife Management* 74:1229–1241.
- Decker, D. J., K. M. Loconti-Lee, and N. A. Connelly. 1990. Deer-related vehicular accidents in Tompkins County, New York: incidents, costs, and implications for deer management. *Transactions Northeast Section Wildlife Society* 47:21–26.
- DeNicola, A. J., K. C. VerCauteren, P. D. Curtis, and S. E. Hygnstrom. 2000. *Managing white-tailed deer in suburban environments*. Cornell Cooperative Extension.
- DeNicola, A. J., and S. C. Williams. 2006. Sharpshooting suburban white-tailed deer reduces deer-vehicle collisions. *Management*.
- Dilts, T. 2010. Topography tools for ArcGIS (9.3, 9.2, 9.1/9.0). <http://arcscrips.esri.com/details.asp?dbid¼15996> Accessed 01 Nov 2010.
- Drzyzga, S. A., and D. G. Brown. 1999. Land parcelization and forest cover fragmentation in three forested countries in Northern Lower Michigan. *Proceedings of the society of American Foresters 1998 National Convention* SAF-99-01:129–135.

- Dussault, C., R. Courtois, and J. P. Ouellet. 2006. A habitat suitability index model to assess moose habitat selection at multiple spatial scales. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere* 36:1097–1107.
- Farrell, M. C., and P. A. Tappe. 2007. County-level factors contributing to deer-vehicle collisions in Arkansas. *Journal of Wildlife Management* 71:2727–2731.
- Fauchald, P., and T. Tveraa. 2003. Using first-passage time in the analysis of area-restricted search and habitat selection. *Ecology* 84:282–288.
- Fawley, B. J. 2012. Michigan deer harvest report: 2011 seasons. Michigan Department of Natural Resources. Report Wildlife Division Report 3548.
- Fawley, B. J., and B. A. Rudolph. 2014. 2012 deer hunting opinion survey. Michigan Department of Natural Resources. Report Wildlife Division Report 3580.
- Finder, R. A., J. L. Roseberry, and A. Woolf. 1999. Site and landscape conditions at white-tailed deer/vehicle collision locations in Illinois. *Landscape and Urban Planning* 44:77–85.
- Finley, A. O., S. Banerjee, and B. P. Carlin. 2007. spBayes: an R package for univariate and multivariate hierarchical point-referenced spatial models. *Journal of Statistical Software* 19:1–24.
- Finley, A. O., S. Banerjee, and A. E. Gelfand. 2012. Bayesian dynamic modeling for large space-time datasets using Gaussian predictive processes. *Journal of Geographical Systems* 14:29–47.
- Finley, A. O., H. Sang, S. Banerjee, and A. E. Gelfand. 2009. Improving the performance of predictive process modeling for large datasets. *Computational Statistics & Data Analysis* 53:2873–2884.
- Forman, R. T., D. Sperling, J. A. Bissonette, A. P. Clevenger, C. D. Cutshall, V. H. Dale, L. Fahrig, R. L. France, C. R. Goldman, and K. Heanue. 2003. *Road ecology: science and solutions*. Island Press, Washington DC, USA.
- Fry, J., G. Xian, S. Jin, J. Dewitz, C. Homer, L. Yang, C. Barnes, N. Herold, and J. Wickham. 2011. Completion of the 2006 National Land Cover Database for the Conterminous United States. *PE&RS* 77:858–864.
- Gehrt, S. D. 2004. Ecology and management of striped skunks, raccoons, and coyotes in urban landscapes. Pages 81–104 in N. Fascione, A. Delach, and M. E. Smith, editors. *People and predators: From conflict to coexistence*. Island Press, Washington DC, USA.
- Gelfand, A. E., S. Banerjee, and D. Gamerman. 2005. Spatial process modelling for univariate and multivariate dynamic spatial data. *Environmetrics* 16:465–479.

- Gesch, D., M. Oimoen, S. Greenlee, C. Nelson, M. Steuck, and D. Tyler. 2002. The National Elevation Dataset. *Photogrammetric Engineering and Remote Sensing* 68:5–11.
- Gesch, D. B. 2007. The National Elevation Dataset. Pages 99–118 in D. Maune, editor. *Digital Elevation Model Technologies and Applications: The DEM Users Manual*. 2nd Edition. American Society for Photogrammetry and Remote Sensing, Bethesda, Maryland, USA.
- Gitzen, R. A., J. J. Millspaugh, and B. J. Kernohan. 2006. Bandwidth selection for fixed-kernel analysis of animal utilization distributions. *Journal of Wildlife Management* 70:1334–1344.
- Glista, D. J., T. L. DeVault, and J. A. DeWoody. 2009. A review of mitigation measures for reducing wildlife mortality on roadways. *Landscape and Urban Planning* 91:1–7.
- Gomes, L., C. Grilo, C. Silva, and A. Mira. 2009. Identification methods and deterministic factors of owl roadkill hotspot locations in Mediterranean landscapes. *Ecological Research* 24:355–370.
- Gould, N., and W. Andelt. 2011. Reproduction and denning by urban and rural San Clemente Island foxes (*Urocyon littoralis clementae*). *Canadian Journal of Zoology* 89:976–984.
- Grilo, C., J. A. Bissonette, and M. Santos-Reis. 2009. Spatial-temporal patterns in Mediterranean carnivore road casualties: consequences for mitigation. *Biological Conservation* 142:301–313.
- Grinder, M., and P. R. Krausman. 2001. Morbidity-mortality factors and survival of an urban coyote population in Arizona. *Journal of Wildlife Diseases* 37:312–317.
- Groves, R. M. 2004. Survey errors and survey costs. Volume 536. John Wiley & Sons, Inc., Hoboken, New Jersey.
- Harden, C. D., A. Woolf, and J. Roseberry. 2005. Influence of exurban development on hunting opportunity, hunter distribution, and harvest efficiency of white-tailed deer. *Wildlife Society Bulletin* 33:233–242.
- Harris, C. D. 1943. Suburbs. *American Journal of Sociology*:1–13.
- Hatchwell, B., D. Chamberlain, and C. Perrins. 1996. The reproductive success of Blackbirds *Turdus merula* in relation to habitat structure and choice of nest site. *Ibis* 138:256–262.
- Heffner, R. A., M. J. Butler, and C. K. Reilly. 1996. Pseudoreplication revisited. *Ecology* 77:2558–2562.
- Homer, C., J. Dewitz, J. Fry, M. Coan, N. Hossain, C. Larson, N. Herold, A. McKerrow, J. N. VanDriel, and J. Wickham. 2007. Completion of the 2001 National Land Cover Database

- for the Counterminous United States. *Photogrammetric Engineering and Remote Sensing* 73:337–341.
- Hosmer, D. W., S. Lemeshow, and R. X. Sturdivant. 2013. *Applied logistic regression*. John Wiley & Sons, Hoboken, NJ.
- Hothorn, T., R. Brandl, and J. Müller. 2012. Large-scale model-based assessment of deer-vehicle collision risk. *PloS one* 7:e29510.
- Hubbard, M. W., B. J. Danielson, and R. A. Schmitz. 2000. Factors influencing the location of deer-vehicle accidents in Iowa. *The Journal of Wildlife Management*:707–713.
- Huijser, M. P., J. W. Duffield, A. P. Clevenger, R. J. Ament, and P. T. McGowen. 2009. Cost-benefit analyses of mitigation measures aimed at reducing collisions with large ungulates in the United States and Canada: a decision support tool. *Ecology and Society* 14:15. Available from <http://www.ecologyandsociety.org/vol14/iss12/art15/>.
- Huijser, M. P., P. T. McGowen, J. Fuller, A. Hardy, and A. Kociolek. 2008. *Wildlife-Vehicle Collision Reduction Study: Report to Congress*. U.S. Department of Transportation, Federal Highway Administration.
- Hundertmark, K. 2007. Home range, dispersal and migration. Pages 303–335 in A. Franzmann, and C. Schwartz, editors. *Ecology and management of the North American moose*. University Press of Colorado, Boulder, Colorado, USA.
- Hurlbert, S. H. 1984. Pseudoreplication and the design of ecological field experiments. *Ecological Monographs* 54:187–211.
- Iverson, A. L., and L. R. Iverson. 1999. Spatial and temporal trends of deer harvest and deer-vehicle accidents in Iowa. *Ohio Journal of Science* 99:84–94.
- Jackson, H. B., and L. Fahrig. 2012. What size is a biologically relevant landscape? *Landscape Ecology* 27:929–941.
- Jordan, S., J. P. Ross, and K. G. Usowski. 1998. US suburbanization in the 1980s. *Regional Science and Urban Economics* 28:611–627.
- Kanda, L. L., T. K. Fuller, P. R. Sievert, and R. L. Kellogg. 2009. Seasonal source-sink dynamics at the edge of a species' range. *Ecology* 90:1574–1585.
- Kernohan, B. J., R. A. Gitzen, and J. J. Millspaugh. 2001. Analysis of animal space use and movements. Pages 125–166 in J. J. Millspaugh, and J. M. Marzluff, editors. *Radio tracking and animal populations*. Academic Press, San Diego, California, USA.
- Kéry, M. 2010. *Introduction to WinBUGS for ecologists: A Bayesian approach to regression, ANOVA and related analyses*. Academic Press.

- Kling, G. W., K. Hayhoe, L.B. Johnson, J.J. Magnuson, S. Polasky, S.K., B. J. S. Robinson, M.M. Wander, D.J. Wuebbles, D.R. Zak, R.L. Lindroth, and a. M. L. W. S.C. Moser. 2003. Confronting climate change in the Great Lakes region: impacts on our communities and ecosystems. Union of Concerned Scientists, Cambridge, Massachusetts, and Ecological Society of America, Washington, D.C., USA.
- Knapp, K. K., C. Kienert, and A. Witte. 2005. Statewide and Upper Midwest summary of deer-vehicle crash and related data from 1993 to 2003. Midwest Regional University Transportation Center, Deer-Vehicle Crash Information Clearinghouse, University of Wisconsin.
- Koitzsch, K. B. 2002. Application of a moose habitat suitability index model to Vermont Wildlife Management Units. *Alces* 38:89–107.
- Krebs, C. J. 1999. Sample Size Determination and Statistical Power. Pages 229–260 *in* Ecological methodology, Second Edition. Addison Wesley Longman, Inc., Menlo Park, California.
- Laver, P. N., and M. J. Kelly. 2008. A critical review of home range studies. *Journal of Wildlife Management* 72:290–298.
- Lavrakas, P. J. 2008. Sampling Bias. *Encyclopedia of Survey Research Methods*. Pages 785–786 *in* P. J. Lavrakas, editor. SAGE Publications, Inc., Thousand Oaks, CA.
- Leptich, D. J., and J. R. Gilbert. 1989. Summer home range and habitat use by moose in northern Maine. *Journal of Wildlife Management* 53:880–885.
- Levine, N. 2004. CrimeStat III--A spatial statistics program for the analysis of crime incident locations. Ned Levine & Associates: Houston, Texas, and the National Institute of Justice.
- Lovely, K. R., W. J. Mcshea, N. W. Lafon, and D. E. Carr. 2013. Land parcelization and deer population densities in a rural county of Virginia. *Wildlife Society Bulletin* 37:360–367.
- Maine Office of GIS. 2010. Geographic name information system (GNIS). State of Maine. <http://www.maine.gov/megis/catalog/> Accessed 02 August 2010.
- Malo, J. E., F. Suárez, and A. Díez. 2004. Can we mitigate animal–vehicle accidents using predictive models? *Journal of Applied Ecology* 41:701–710.
- Marcoux, A., and S. Riley. 2010. Driver knowledge, beliefs, and attitudes about deer–vehicle collisions in southern Michigan. *Human-Wildlife Interactions* 4:47–55.
- McGarigal, K., S. A. Cushman, M. C. Neel, and E. Ene. 2002. FRAGSTATS: spatial pattern analysis program for categorical maps. *in* University of Massachusetts, Amherst, USA.

- Michigan Department of Transportation. 2013. Federal-aid Eligibility: Using NFC Maps.
http://www.michigan.gov/mdot/0,1607,7-151-9622_11033_11155-25881--,00.html
 Accessed 23 July 2013.
- Mitchell, B. R. 2005. Fragstatsbatch extension for ArcGIS 9. Arcscripts.
<http://arcscripsts.esri.com/details.asp?dbid=13995> Accessed 1 Nov 2010.
- Moore, C. M., and P. W. Collins. 1995. *Urocyon littoralis*. Mammalian Species:1–7.
- Morris, K. 1999. Moose Assessment. State of Maine, Department of Inland Fisheries and Wildlife.
- National Oceanic and Atmospheric Administration Coastal Services Center. 2012. NOAA Coastal Change Analysis Program (C-CAP) Regional Land Cover Database.
<http://www.csc.noaa.gov/digitalcoast/data/ccapregional/> Accessed 17 January 2010.
- Nelson, T. A., and B. Boots. 2008. Detecting spatial hot spots in landscape ecology. *Ecography* 31:556–566.
- Ng, J. W., C. Nielson, and C. C. St Clair. 2008. Landscape and traffic factors influencing deer-vehicle collisions in an urban enviroment. *Human-Wildlife Interactions* 2:34–47.
- Nielsen, C. K., R. G. Anderson, and M. D. Grund. 2003. Landscape influences on deer-vehicle accident areas in an urban environment. *The Journal of Wildlife Management*:46–51.
- Nystrom, S. K. 2007. Modelling Spatial Characteristics of Deer-vehicle Accident Sites in Onondaga County, New York. Thesis, State University of New York, Syracuse., Syracuse.
- Okabe, A., T. Satoh, and K. Sugihara. 2009. A kernel density estimation method for networks, its computational method and a GIS-based tool. *International Journal of Geographical Information Science* 23:7–32.
- Okabe, A., and K. Sugihara. 2012. Spatial analysis along networks. John Wiley & Sons, West Sussex, UK.
- Openshaw, S., and P. J. Taylor. 1981. The modifiable areal unit problem. Routledge and Kegan Paul, London, UK.
- Peek, J. M. 2007. Habitat relationships. Pages 351–376 in A. W. Franzmann, and C. C. Schwartz, editors. *Ecology and Management of the North American Moose*. University Press of Colorado, Boulder, USA.
- Quinn, A. C. D., D. M. Williams, and W. F. Porter. 2012. Landscape structure influences space use by white-tailed deer. *Journal of Mammalogy* 94:398–407.

- Ramp, D., J. Caldwell, K. A. Edwards, D. Warton, and D. B. Croft. 2005. Modelling of wildlife fatality hotspots along the snowy mountain highway in New South Wales, Australia. *Biological Conservation* 126:474–490.
- Reed, D. F., T. D. Beck, and T. N. Woodard. 1982. Methods of reducing deer-vehicle accidents: benefit-cost analysis. *Wildlife Society Bulletin*:349–354.
- Resnik, J. R. 2012. Home range, site fidelity, reproductive ecology, and den site characteristics of the San Clemente Island fox. Thesis, Colorado State University, Fort Collins, Colorado, USA.
- Ribeiro, P. J., and P. J. Diggle. 2001. geoR: A package for geostatistical analysis. *R news* 1:14–18.
- Robin, X., N. Turck, A. Hainard, N. Tiberti, F. Lisacek, J.-C. Sanchez, and M. Müller. 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 12:1–8.
- Romin, L. A., and J. A. Bissonette. 1996. Deer: vehicle collisions: status of state monitoring activities and mitigation efforts. *Wildlife Society Bulletin* 24:276–283.
- Roseberry, J. L., and A. Woolf. 1998. Habitat-population density relationships for white-tailed deer in Illinois. *Wildlife Society Bulletin* 26:252–258.
- Rosenblatt, D. L., E. J. Heske, S. L. Nelson, D. M. Barber, M. A. Miller, and B. MacALLISTER. 1999. Forest fragments in east-central Illinois: islands or habitat patches for mammals? *The American midland naturalist* 141:115–123.
- Seaman, D. E., J. J. Millspaugh, B. J. Kernohan, G. C. Brundige, K. J. Raedeke, and R. A. Gitzen. 1999. Effects of sample size on kernel home range estimates. *Journal of Wildlife Management*:739–747.
- Seaman, D. E., and R. A. Powell. 1996. An evaluation of the accuracy of kernel density estimators for home range analysis. *Ecology*:2075–2085.
- Snow, N. P., W. F. Andelt, and N. P. Gould. 2011. Characteristics of road kill locations of San Clemente Island foxes. *Wildlife Society Bulletin* 35:32–39.
- Snow, N. P., W. F. Andelt, T. R. Stanley, J. R. Resnik, and L. Munson. 2012. Effects of roads on survival of San Clemente Island foxes. *The Journal of Wildlife Management* 76:243–252.
- Snow, N. P., D. M. Williams, and W. F. Porter. 2014. A landscape-based approach for delineating hotspots of wildlife-vehicle collisions. *Landscape Ecology* 29:817–829.

- Spinks, P. Q., G. B. Pauly, J. J. Crayon, and H. Bradley Shaffer. 2003. Survival of the western pond turtle (*Emys marmorata*) in an urban California environment. *Biological Conservation* 113:257–267.
- Stock, J. H., and M. W. Watson. 2012. Disentangling the Channels of the 2007–2009 Recession. National Bureau of Economic Research.
- Stohlgren, T. J., Y. Otsuki, C. A. Villa, M. Lee, and J. Belnap. 2001. Patterns of plant invasions: a case example in native species hotspots and rare habitats. *Biological Invasions* 3:37–50.
- Sudharsan, K., S. Riley, B. Riley, and B. Maurer. 2005. Deer-vehicle crash patterns across ecoregions in Michigan, Pages 246–255 in Conference Deer-vehicle crash patterns across ecoregions in Michigan. D. L. Nolte, and K. A. Fagerstone.
- Sullivan, J. M. 2011. Trends and characteristics of animal-vehicle collisions in the United States. *Journal of Safety Research* 42:9–16.
- Sullivan, T. L., A. F. Williams, T. A. Messmer, L. A. Hellinga, and S. Y. Kyrychenko. 2004. Effectiveness of temporary warning signs in reducing deer-vehicle collisions during mule deer migrations. *Wildlife Society Bulletin* 32:907–915.
- Sward, W., and R. Cohen. 1980. Plant community analysis of San Clemente Island. Report to Naval Ocean System Center, San Diego, California, USA.
- Thompson, I. D., M. D. Flannigan, B. M. Wotton, and R. Suffling. 1998. The effects of climate change on landscape diversity: an example in Ontario forests. *Environmental Monitoring and Assessment* 49:213–233.
- Thorne, R. F. 1976. The vascular plant communities of California. Pages 1–31 in J. Latting, editor. *Plant communities of Southern California*. California Native Plant Society Special Publication no. 2, Riverside, California, USA.
- U.S. Department of the Interior | U.S. Geological Survey. 2012. GAP Land Cover Updated. <http://gapanalysis.usgs.gov/gaplandcover/gap-land-cover-updated/> Accessed 12 Nov 2013.
- United States Census Bureau. 2012. Population Estimates. United States Department of Commerce. <http://www.census.gov/popest/data/intercensal/county/county2010.html> Accessed 01 Jan 2014.
- University of Illinois Extension. 2013. Living with white-tailed deer in Illinois. <http://web.extension.illinois.edu/deer/strategy.cfm> Accessed 06 Nov 2013.
- Walter, W. D., K. C. VerCauteren, H. Campa III, W. R. Clark, J. W. Fischer, S. E. Hygnstrom, N. E. Mathews, C. K. Nielsen, E. M. Schaubert, and T. R. Van Deelen. 2009. Regional

- assessment on influence of landscape configuration and connectivity on range size of white-tailed deer. *Landscape Ecology* 24:1405–1420.
- Wickham, J., S. Stehman, J. Fry, J. Smith, and C. Homer. 2010. Thematic accuracy of the NLCD 2001 land cover for the conterminous United States. *Remote Sensing of Environment* 114:1286–1296.
- Wiens, J. A. 1976. Population responses to patchy environments. *Annual Review of Ecology and Systematics* 7:81–120.
- Wiens, J. A. 1989. Spatial scaling in ecology. *Functional Ecology* 3:385–397.
- Williams, D. M., A. C. D. Quinn, and W. F. Porter. 2012. Landscape effects on scales of movement by white-tailed deer in an agricultural–forest matrix. *Landscape Ecology* 27:45–57.
- Worton, B. J. 1989. Kernel methods for estimating the utilization distribution in home-range studies. *Ecology* 70:164–168.
- Xie, Z., and J. Yan. 2008. Kernel density estimation of traffic accidents in a network space. *Computers Environment and Urban Systems* 32:396–406.