The Pennsylvania State University

The Graduate School

School of Forest Resources

METHODS OF ESTIMATING WHITE-TAILED DEER ABUNDANCE AT GETTYSBURG NATIONAL MILITARY PARK: TESTING ASSUMPTIONS OF DISTANCE SAMPLING

A Thesis in

Wildlife and Fisheries Science

by

David P. Stainbrook

Submitted in Partial Fulfillment
of the Requirements
for the degree of

Master of Science

December 2011

The thesis of David P. Stainbrook was reviewed and approved* by the following:

Duane R. Diefenbach
Adjunct Professor of Wildlife Ecology
Leader, PA Cooperative Fish and Wildlife Research Unit
Thesis Adviser

Tyler Wagner
Adjunct Assistant Professor of Fisheries Ecology
Assistant Unit Leader, PA Cooperative Fish and Wildlife Research Unit

Walter M. Tzilkowski
Associate Professor of Wildlife Science

Michael G. Messina
Director of the School of Forest Resources

*Signatures are on file in the Graduate School.

ABSTRACT

The mission at Gettysburg National Military Park and Eisenhower National Historic Site (GNMP-ENHS) is to preserve the historic character of the parks to enable current and future generations to understand and interpret the events that took place at each park. Management objectives include maintaining the landscape as it existed during the historic 1863 Civil War battle (e.g., dense understory in woodlots) in GNMP, and as it existed during the Eisenhowers' occupancy (e.g., patchwork of cropfields). Browsing by white-tailed deer (*Odocoileus virginianus*) diminished regeneration of native trees in woodlots and prevented crops from reaching maturity. Thus, to increase regeneration in woodlots and reduce crop damage, the National Park Service (NPS) began culling deer in 1995 to reach a density of 10 deer/km^2 of forest. However, park managers required an accurate population estimate to determine if this management goal has been met. Therefore, I captured and fitted adult and juvenile male and female deer with Global Positioning System (GPS) collars and performed surveys at dusk and at night, from April 2009 to November 2010, to estimate abundance using mark-resight methods. I found that the average detection probability ($\hat{p}$) during the April 2010 dusk count was 0.25, compared to 0.54 from research conducted over 20 years prior. Previous research used only marked female deer, and a number of factors that influence detectability of deer likely changed over time. Park managers can use my updated $\hat{p}$ as their sighting index to estimate deer abundance during future deer counts, which may provide more accurate estimates of abundance. However, factors that influence detection probability can change over time; thus, accuracy of future estimates of abundance could change.

Additionally, I conducted distance sampling surveys from roads during mark-resight surveys at night to test assumptions when roads are used as transects with distance sampling. A critical requirement of distance sampling is that transects are placed

randomly on the landscape to obtain a representative sampling of the study area and to meet the assumption that the distribution of deer is uniform with respect to perpendicular distances to transects. Roads have been used as transects for distance sampling and provide logistical advantages, but roads may be correlated with habitat characteristics and the distribution of the animals. Distance sampling can be a useful estimator for monitoring abundance; however, if roads are used as transects, the magnitude and direction of the bias are unknown unless information on the distribution of the animal is available. Therefore, I used GPS locations from GPS-collared deer to model the distribution of deer relative to roads using a resource selection function (RSF). During the hours when I conducted spotlight surveys, the distribution of deer was not uniform with respect to the location of roads in both forested and non-forested areas. Deer avoided areas close to roads, were more likely to be found near the park boundary, and selected for forested areas and open areas near forest edges. The estimator of detection probability was positively biased when deer avoided roads; thus, estimates of density in the sampled area were negatively biased. Although I failed to reject the null hypothesis that using roads as transects provided a representative sampling of the study area, extrapolating biased density estimates from the sample to the study area likely resulted in biased estimates of abundance in the study area. Further, estimates of abundance from distance sampling, using roads as transects, were lower than estimates of abundance from mark-resight estimators. Additionally, I demonstrated how the RSF can be used to account for non-random placement of transects to obtain more accurate estimates of abundance.

Both mark-resight and distance sampling estimators provided density estimates approximately 4 times greater than the park's goal of 10 deer/km^2 of forested land. I estimated abundance for April 2010 as 368 (95% CI = 322 – 421) and density ($\hat{D}$) as 43 deer/km^2 of forest during dusk mark-resight surveys; 403 (95% CI = 297 – 546; $\hat{D}$ = 48 deer/km^2 of forest) during spotlight mark-resight surveys; and 381 (95% CI = 238 – 607; $\hat{D}$ = 45 deer/km^2 of forest) during distance sampling surveys. I estimated abundance for November 2010 as 425 (95% CI = 196 – 921; $\hat{D}$ = 50 deer/km^2 of forest) during dusk mark-resight surveys; 598 (95% CI = 420 – 852; $\hat{D}$ = 71 deer/km^2 of forest) during spotlight mark-resight surveys; and 366 (95% CI = 255 – 525; $\hat{D}$ = 43 deer/km^2 of forest) during distance sampling surveys. However, density within the entire study area was not homogenous. I observed more deer on private lands in the study area and fewer deer on NPS owned property. Park staff have observed increased tree regeneration and reduced crop damage since culling was initiated, even though current deer densities are approximately 4 times greater than the goal specified in the GNMP-ENHS deer management plan. Consequently, the NPS may want to consider re-evaluating deer density goals.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

This project was made possible through the funding and cooperation between the National Park Service and the Pennsylvania Cooperative Fish and Wildlife Research Unit, of the U.S. Geologic Survey, located at the Pennsylvania State University. I would first like to thank my advisor Duane Diefenbach. Duane's hands-off advising style facilitated my professional development, independence, and responsibility; but he was always available to give me direction and guidance when necessary. I would also like to thank my committee members Tyler Wagner and Walter Tzilkowski for their assistance and feedback throughout the process of my research. Tyler also provided invaluable support regarding writing R code and interpreting outputs.

I am thankful for my hardworking technicians who assisted with deer capture and abundance surveys: Robbie Withington, Amanda Sommerer, Chris Nevius, Keith Baumgardner, Brandon Black, Jeremy Redding, and volunteer Phil Bietsch. I thank Zachary Bolitho and Sara Koenig, biologists at Gettysburg National Military Park, for their input and assistance with the research. I would like to thank all of the volunteers that assisted with the dusk surveys, especially Beth Brantley, instructor of forest technology at Penn State Mont Alto, for providing student volunteers. I would also like to thank all of the private landowners who allowed us to trap deer on their properties.

I thank Kay Christine and B.J. Scovern, administrative support assistants for the Pennsylvania Cooperative Fish and Wildlife Research Unit for their valuable support and assistance throughout my time at Penn State. I give special thanks to my fellow graduate students: Andrew Norton, Brooks Fost, Jason Hill, Cairsty Grassie, and Franny Buderman for their assistance, social breaks from work, and intellectual discussions that furthered my understanding of statistics. Finally, I would like to thank my friends and family, especially Elyssa Prince, for their constant support and encouragement.

CHAPTER 1:

HISTORY AND STUDY RATIONALE

The abundance of white-tailed deer (*Odocoileus virginianus*) in national parks has been a controversial issue for decades (Leopold 1963, Warren 1991). In the 1980s, resource managers at Gettysburg National Military Park and Eisenhower National Historic Site (GNMP-ENHS), Pennsylvania, were concerned that deer were adversely affecting park resources and leading to increased deer-vehicle collisions in and around the park (Frost et al. 1997). Tzilkowski and Storm (1993) estimated deer abundance on the park to be 1,018 deer in 1992, a density of 136 deer/km^2 of forested land, over 10 times the density recommended by the Pennsylvania Game Commission for Adams County at that time. Research at GNMP-ENHS concluded that deer reduced crop yields (Frost et al. 1997, Vecellio et al. 1994) and prevented forest regeneration (Storm et al. 1989).

The mission at GNMP-ENHS is to preserve the historic character of the parks to enable current visitors and future generations to interpret the significant historical events that took place at each park (U.S. Department of the Interior 1995). Management objectives for GNMP include maintaining the landscape as it existed prior to and during the historic 1863 Civil War battle, which included a dense understory of native vegetation in woodlots (U.S. Department of the Interior 1995, Frost et al. 1997, Newinski et al. 2006). Management objectives for ENHS include preserving the agricultural setting, so the former president's agrarian use of the farm and management strategies are properly understood by visitors, which included growing crops of corn, milo, soybeans, small grains, and hay (U.S. Department of the Interior 1995). Browsing by white-tailed deer (*Odocoileus virginianus*) diminished regeneration of native trees in woodlots and prevented crops from reaching maturity. The preferred management action identified in the Environmental Impact Statement was a combination of culling deer in the park and increased hunting efforts on private lands surrounding the park to reach a density of 10 deer/km^2 of forested land (U.S. Department of the Interior 1995), diminish deer browsing, and allow regeneration in woodlots (Frost et al. 1997).

An important component of the Resource Management Plan for GNMP-ENHS is a reliable estimate of the number of deer in the park. Estimates of abundance are the basis for population models to predict the response of populations to specific management actions and are important for making decisions regarding how many deer to kill to meet management objectives (Frost et al. 1997, Storm et al. 1992). Estimates of deer density are important in national parks (especially in the National Capital Region Network) that have implemented or are considering deer population reduction, as evidenced by the number of parks that have implemented deer monitoring programs (Bates 2006).

Buckland et al. (2000) classified the three major areas of wildlife population estimation as mark-recapture (including mark-resight and mark-recovery), harvest models (including catch-effort and removal methods), and distance sampling. All of these can be used to estimate abundance of deer, but each has a different set of assumptions and requirements, which may be difficult to meet to obtain unbiased and precise population estimates.

The mark-resight method, based on the Lincoln-Petersen estimator (Seber 1982) where resightings are used in place of recaptures, has been widely used to estimate abundance of ungulates (Rice and Harder 1977, Bartmann et al. 1987, McCullough and Hirth 1988, Storm et al. 1989, Neal et al. 1993, Focardi et al. 2002). Standard assumptions include: (1) geographically and demographically closed population; (2) marks are not lost; (3) marked individuals are correctly identified; (4) marked population has the same probability of sighting as unmarked population; (5) homogeneous sighting probabilities for all animals within a sampling occasion; and (6) individuals are not counted more than once within a sampling occasion (Neal et al. 1993, White and Shenk 2001). Neal et al. (1993) and McCullough and Hirth (1988) found the estimator was particularly sensitive to violations of assumption 5, when sighting probabilities vary among individuals (e.g., individual heterogeneity), resulting in bias. Bowden's estimator relaxes this assumption, which is often difficult to meet, but marked animals must be individually identifiable (Bowden and Kufeld 1995). The mark-resight method is typically used to assess population density on small areas because it is expensive and time consuming to capture a sufficient sample of marked deer in the population to obtain accurate estimates of abundance across large areas (Bartmann et al. 1987, Neal et al. 1993, Focardi et al. 2002, McClintock et al. 2006).

Harvest model methods do not require marked animals and could be far less expensive than the mark-resight method when used to estimate population size, but have limited applicability because removals (e.g., harvests or captures) are necessary to meet assumptions. Catch-effort methods assume catch per unit effort (e.g., animals killed per hour of culling) is proportional to population abundance and that the population is closed (Seber 1982, Buckland et al. 2000). However, assumptions may be difficult to meet (Quinn and Deriso 1999), and removal of a large proportion of the population is often required to obtain reliable estimates (Lewis and Farrar 1968, Lancia et al. 1996). Also, abundance can be estimated using change-in-ratio methods based on the change in the proportion of each class of animal after a differential harvest (e.g., cull only antlerless deer) and the total number harvested (Buckland et al. 2000). Assumptions include a geographically and demographically closed population, observed proportions of each class are unbiased and representative of the population (e.g., equal probability of harvest), and the total number of each class removed is known (Seber 1982, Conner et al. 1986). However, assumptions may be difficult to meet (e.g., not all harvested animals are retrieved) and a large proportion of the population must be harvested to observe a noticeable change in the proportion of each class and obtain reliable estimates (Conner et al. 1986, Buckland et al. 2000).

Distance sampling methods also do not require marked animals and could be far less expensive than mark-resight methods (Focardi et al. 2002) and more applicable to a wider range of species than harvest models. However, assumptions may be difficult to meet to obtain unbiased population estimates of highly mobile animals such as deer (Buckland et al. 2001, Koenen et al. 2002, Fewster et al. 2008). Assumptions include: (1) surveys are conducted from randomly-placed points or transects; (2) all objects on or near a point or transect are detected with certainty; (3) objects are detected at their initial location and any movement prior to detection is independent of observers; and (4) measurements are accurate (Buckland et al. 2001). Common methods of ground navigation of random transects or points include walking, horseback, and all-terrain vehicles; but these may result in deer moving in response to observers before detection, which results in negatively biased estimates of density (e.g., see Koenen et al. 2002). Aerial surveys can avoid the problem of deer movement in response to the observer, but are expensive, animals may move in response to a low-flying plane or helicopter, and it is difficult to ensure that all deer on the transect are detected, especially in forested landscapes (Naugle et al. 1996, Haroldson et al. 2003, Thomas et al. 2010). Surveying from roads using distance sampling is a convenient and commonly used method (Heydon et al. 2000, Koenen et al. 2002, Ruette et al. 2003, Ward et al. 2004, Bates 2006), which can reduce movement in response to observers (e.g., use of thermal imagers; Gill et al. 1997). However, roads are not random; thus, sampling from them violates the critical assumption of randomly placed transects and can result in biased estimates of density, which are unrepresentative of the population (Buckland et al. 2001). Assumptions can be met easily when applying distance sampling methods to count dung (Buckland et al. 2001, Marques et al. 2001). However, accuracy of density estimates rely on estimates of both defecation rates and dung decay rates, which often are estimated using penned deer, and can vary spatially, seasonally, and by differences in feeding behavior related to sex and age (Van Etten and Bennet 1965, Mitchell et al. 1985).

I used both mark-resight and distance sampling methods to estimate abundance of deer in the study area including GNMP-ENHS and used location information from a known, marked population to test assumptions behind these estimators. I investigated individual heterogeneity and temporal variation in sighting probabilities of deer during mark-resight surveys and evaluated the efficacy of applying mark-resight sighting probabilities to future surveys. In addition, I tested whether using roads as transects with distance sampling provided unbiased estimates of deer abundance and evaluated correction methods.

CHAPTER 2:

ESTIMATING ABUNDANCE USING MARK-RESIGHT METHODS

INTRODUCTION

It is well known that road-based and aerial surveys underestimate population size of ungulates when the proportion of missed observations, or visibility bias, is not accounted for in the survey design and estimation (Caughley 1974, Caughley 1977). Steinhorst and Samuel (1989) developed a sightability model to adjust for visibility bias using marked animals to estimate sighting probabilities for each animal group during surveys, which could be applied to future surveys to estimate abundance (Samuel et al. 1987, Cogan and Diefenbach 1998). Mark-resight data of marked animals also is used to estimate sighting probabilities, but at the population level, and applied to future surveys to estimate abundance (Seber 1982). Applying sighting probabilities to future surveys is a cheaper alternative than retaining a sample of marked deer in the population. However, if the assumptions of equal detectability (e.g., all individuals have the same probability of detection during a survey) or constant detectability over time (e.g., sighting probabilities do not change over time) fail, the estimator will be biased (Otis et al. 1978, Seber 1982, Pollock and Kendall 1987, Neal et al. 1993, Anderson 2001). Sighting probabilities can be influenced by variability among individuals (e.g., behavioral or physical differences such as group size, age, and sex; Downing et al. 1977, Samuel et al. 1992) and variability related to sampling and temporal changes (e.g., differences in observers, season, vegetation cover, and weather; Samuel et al. 1987, Cogan and Diefenbach 1998, Anderson 2001, McClintock et al. 2006).

Previous research conducted at GNMP-ENHS estimated sighting probability, termed the average detection probability, $\hat{\bar{p}}$, as the average proportion of 30 to 54 marked female deer that were resighted during April ($\hat{p}$ = 0.54) and November ($\hat{p}$ = 0.43) mark-resight surveys performed at dusk from 1987 to 1991 (Storm et al. 1992). Since 1993, the park has used $\hat{p}$ =0.54 from the Storm et al. (1992) study to adjust annual April dusk counts of unmarked deer to estimate abundance.

However, the mark-resight method used at GNMP-ENHS was the Lincoln-Petersen estimator, which assumes no heterogeneity in re-sighting probabilities (Seber 1982). Nevertheless, the estimator will be unbiased if the marked sample is representative of the population, such that the ratio of marked deer observed to marked deer available is the same as the ratio of unmarked deer seen to the total number of unmarked deer on the study area (Otis et al. 1978, Seber 1982, White et al. 1982). Because Storm et al. (1992) used only marked female deer to calculate $\hat{p}$ for all deer during mark-resight surveys, and male deer tend to exhibit lower sighting probabilities than females (McCullough 1982, Sage et al. 1983, McCullough and Hirth 1988), I believe the estimate of $\hat{p}$ was likely unrepresentative, leading to negatively biased estimates of abundance. This estimator only provides unbiased estimates if the assumption that male and female deer have the same probability of detection during surveys is true, which has not been investigated at GNMP-ENHS.

Furthermore, for the sighting probability to yield unbiased estimates of abundance in future surveys, the assumption that sighting probability has not changed over time must be met. This may not be valid at GNMP-ENHS because the detection probability of deer may have changed in response to a number of factors, including the reduction of deer density and habitat changes over the past 20 years.

Resource managers at GNMP-ENHS began culling antlerless deer in 1995 and continued culling every fall and winter from roads and fields on National Park Service

(NPS) owned property. Given the relatively small home ranges of antlerless deer, removals likely caused a reduced density of antlerless deer near park roads and fields. Additionally, the culling operations may have caused surviving antlerless and antlered deer to avoid the areas near park roads, such that fewer deer would be seen from those roads during April dusk counts. Further, culling operations led to reduced browsing by deer on park woodlands, which led to increased seedling tree density (Niewinski et al. 2006), which would decrease visibility in woodlots. Therefore, I hypothesized that culling operations resulted in a decrease in $\hat{\bar{p}}$ over time.

Additionally, several woodlots had trees removed to restore the park to its visual condition during the Civil War. Management included thinning or canopy opening of historic woodlots to stimulate tree regeneration (Niewinski et al. 2006) and removing entire sections of woodlots that were historically fields. I hypothesized that the increase in understory would reduce visibility and result in a decrease in $\hat{\bar{p}}$, but removal of entire sections of woodlots would result in an increase in $\hat{\bar{p}}$, such that the overall change in $\hat{\bar{p}}$ solely based on the effects of woodlot management would be difficult to predict.

Overall, I hypothesized that dusk mark-resight surveys would yield decreased average detection probability values from those found in Storm et al. (1992) because my marked population included all age and sex classes and because past management of the deer herd and habitat resulted in a denser understory of vegetation. My first research objective was to estimate a new average detection probability by monitoring the proportion of marked male and female deer seen during dusk mark-resight surveys at GNMP-ENHS, for managers to estimate abundance during future surveys. Next, I investigated individual heterogeneity and temporal variation in sighting probabilities of deer. Finally, I estimated abundance during dusk surveys using the bias adjusted Lincoln-Petersen (L-P) estimator and during spotlight surveys using Bowden's estimator.

STUDY AREA

The study area (Fig. 1) encompassed 2,913 ha (7,197 acres or 11.25 mi^2) of land, which included 1,790 ha of NPS owned land (61% of the study area) and 1,122 ha of private land surrounding Gettysburg, Pennsylvania. The study area was divided into 11 compartments (mean size of 264 ha each) to ease deer counting surveys (Storm et al. 1992). I used the same study area and compartment boundaries from Storm et al. (1992) for my analysis because the park manages its deer population based on this area. However, the study area size differs slightly from Storm et al. (1992) because they used a dot grid over aerial photographs to calculate areas, whereas I used a GIS (ArcView 9.3, Environmental Systems Research Institute. Redlands, California, USA).

According to Storm et al. (1989), approximately 48% (1,389 ha) of the study area was agricultural land; 26% (749 ha) was forested; 12% (355 ha) was forbs/shrubland; 8% (216 ha) was commercial, which also included a golf course and cemeteries; 5% (141 ha) was residential; and 1% (12 ha) consisted of lakes, ponds, and streams. As of 2009, approximately 50% was agricultural/grassland, 27% was forested/woodland, 17% was built-up (commercial, residential, and transportation), 3% was recently cleared land, 2% was shrubland, and 1% consisted of lakes, ponds, and streams (GNMP-ENHS, unpublished data). For my study, however, I combined forested/woodland and shrubland as forested because both are important as deer hiding cover, which accounted for 29% (848 ha) of the study area, of which 17% (502 ha) was owned by the NPS (Appendix B). The primary differences in vegetation between 1989 and 2009 included forbs/shrubland that aged into forest, recently cleared shrublands and forests, and an increase in built-up land.

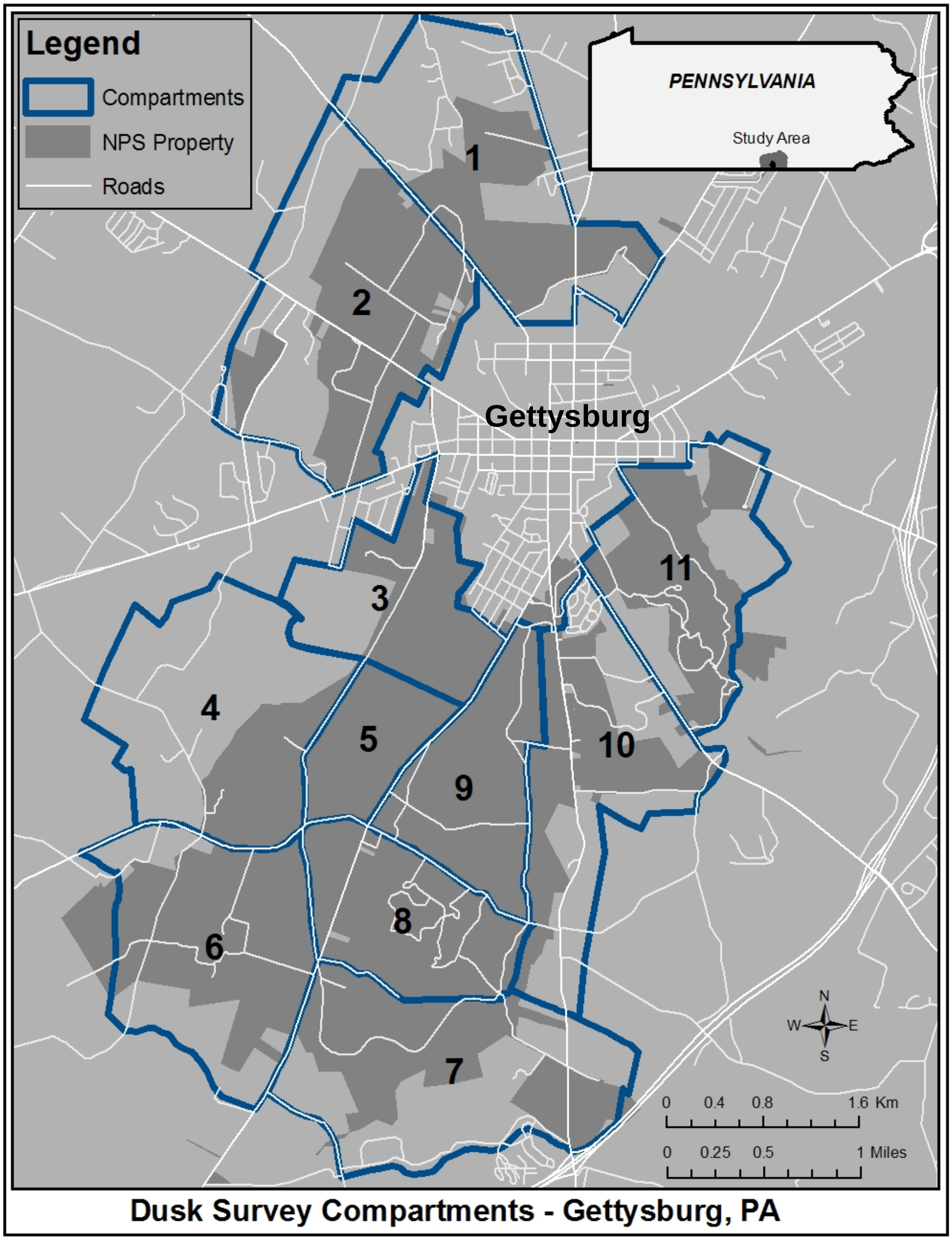


Figure 1. The 11 compartments in which white-tailed deer were counted during dusk mark-resight surveys from April 2009 to November 2010 in the 2,913 ha study area in Adams County, Pennsylvania. The areas in dark gray are National Park Service (NPS) owned property and areas in light gray are privately owned property.

METHODS

Deer Capture

I captured deer throughout the study area during January-April of 2009 and January-April of 2010 using rocket nets (Beringer et al. 1996, Haulton et al. 2001) and modified Clover traps (Clover 1956, Beringer et al. 1996, Haulton et al. 2001) over bait (shelled corn and apples). I sedated deer captured with rocket nets via an intramuscular injection of xylazine hydrochloride (100 mg/mL) at approximately 1 mg/1.8 kg body mass (Rosenberry et al. 1999). Prior to release (10-30 min.), I administered tolazoline hydrochloride (100 mg/mL) intramuscularly to reverse the effects of the xylazine hydrochloride at approximately 1 mg/0.2 kg body mass (Rosenberry et al. 1999). Deer captured in Clover traps were manually restrained and blindfolded, but not sedated if they could be safely handled. Non-sedated deer were released within 5 minutes. The capture and handling of deer followed protocols approved by The Pennsylvania State University Institutional Animal Care and Use Committee (IACUC # 29677).

Deer <1 year of age at time of capture were considered juveniles, and all deer >1 year of age were considered adults. I based age-class determination on body size and eruption of adult lower incisors and additionally, for bucks, evidence of antler growth in the previous year. I fitted all captured deer with ear tags (Original Tags™, Temple Tag Co., Temple, TX, USA) imprinted with a unique ID number and toll-free telephone number. Additionally, I fitted all captured deer with Global Positioning System (GPS) collars (PRO Light-3, VECTRONIC Aerospace GmbH, Berlin, Germany), which were imprinted with a toll-free telephone number for reporting, “Penn State University,” and “REWARD.”

I attached a narrow strip of red reflective tape to the front of the GPS-collar battery pack to aid in detection of GPS-collared deer. The reflective strip was positioned such that it could only be seen when deer were looking at a spotlight, yet it did not increase the probability of detection and it enabled me to easily determine whether a deer was GPS-collared or not after detection. I used 3 different sizes of collars (38 cm, 46 cm, and 56 cm) weighing 838, 863, and 891 g, respectively, with the harnesses capable of additional adjustment of size to best fit each deer. On mature bucks, I used the largest collars with a 4 cm layer of degradable foam to allow for growth and swelling of the neck during the rut. On juvenile males, I used the medium-sized collars with an 8 cm layer of degradable foam to keep the collar from being cast, yet allowing for growth and swelling of the neck during the rut. On all female deer, I used the smallest collars. I set the GPS schedule to take a GPS fix every 5 minutes during surveys; collars transmitted GPS locations daily to a computer via a cellular-phone network.

Dusk Surveys

I performed mark-resight surveys in the study area at dusk (60 minutes prior to sunset until about 30 minutes after sunset) during the same months that Storm et al. (1992) performed surveys: November 2009, April 2010, and November 2010. I used the protocol established by Storm et al. (1989, 1992), which required each compartment to be surveyed by at least 1 person, who continuously traveled on all survey roads in the assigned compartment. Observers used binoculars to classify deer as marked or unmarked, juvenile or adult, male or female, and noted the habitat (woods or field) in which they observed the deer (Storm et al. 1989, 1992). Observers surveyed the entire study area (Fig. 1) in 1 evening, meeting the assumptions of most mark-resight estimators regarding deer movements (i.e., a given deer could not be counted in multiple compartments during a single survey). I surveyed 3 consecutive days to increase precision in estimates. I postponed surveys if suitable weather conditions were not met (wind <32 km/hr, no rain or only drizzle, visibility >1.6 km).

In April 2009, I performed a survey using the method currently used by GNMP-ENHS for its April dusk deer counts, which, due to staff limitations, was adjusted from the method of Storm et al. (1992) so that only 2 people were required to perform a survey. Their current survey method for dusk deer counts included only 1 survey crew consisting of a driver and 1 to 2 observers seated in a pick-up truck using binoculars to count deer. The crew covered 2 to 4 of the 11 compartments of the study area (Fig. 1) each evening (did not continuously survey each compartment), such that a complete survey took approximately 3 or 4 days. They repeated surveys until the entire study area was surveyed 3 times.

Spotlight Surveys

I conducted additional mark-resight surveys after dusk in the study area prior to culling efforts (August), during culling efforts (November), after the culling operation (January), and following the deer capture season (April). I surveyed in April and November because these are the same months that Storm et al. (1992) performed surveys and to allow comparison with dusk count estimates. I chose to perform additional surveys in August and January to investigate the robustness of abundance estimates with respect to changes in visibility on the landscape. In August, visibility is low because of dense foliage in forests and corn and tall grasses in fields. In January, visibility is high in both forests and fields because there are no leaves on trees and all row crops have been harvested. I made minor adjustments to the day of the month for surveys during the 2010 season to avoid dates with high visitor traffic in and around the park.

I started spotlight surveys 30 minutes after sunset and continued for approximately 3-5 hours each night. It took approximately 2 to 3 nights to survey the study area once; I surveyed the study area 3 times to increase precision in estimates. I performed surveys using a driver and 2 observers standing in the bed of a pick-up truck, using handheld spotlights (240 SL BLITZ, TUFFLIGHTS.COM), each responsible for illuminating the area on their side of the road to detect deer.

I selected survey transects used in the spotlight mark-resight surveys from existing roads (Fig. 8) within the study area. For safety reasons, I chose survey routes located on roads with less vehicular traffic, such as NPS roads. I included additional roads and viewpoints to fields within the study area, but not viewable from transects to improve coverage of the study area. I traversed transects at 10-25 km/hr and varied initial starting points and routes to minimize temporal influences in deer detection that may have existed because of deer activity patterns. I divided transects into 3 groups for surveying each night based on geographic location (e.g. – North, Southwest, and Southeast). I did not survey on a particular night if adverse environmental conditions existed (wind $\geq$32 km/hr, rain, visibility $\leq$1.6 km).

Once an observer spotted a deer or group of deer, the driver stopped, the observer shined a spotlight on the deer or group of deer and looked for the presence of GPS collars with assistance from the other observer using binoculars. The observer also aged each deer as a fawn or adult and noted whether it was antlered or antlerless. If it was not possible to determine if a deer was GPS-collared, it was recorded as unidentified. I used a GIS to identify all observed unidentified deer as marked or unmarked and to individually identify all marked deer seen by comparing GPS locations taken during the time of the observation to the location of the observation.

Lincoln-Petersen Estimator: Dusk Surveys

I used the Lincoln-Petersen estimator (L-P) to estimate abundance (*N*) for each evening of each dusk count survey where

$$\hat{N} = \frac{(n_1+1)(n_2+1)}{(m_2+1)} - 1,$$

n_1 was the number of marked deer available (in the study area) each evening, n_2 was the total number of deer that were identified as marked or unmarked, and m_2 was the number of marked or GPS-collared deer seen (Seber 1982). I determined the number of GPS-collared deer in the study area (n_1) each evening for each survey by monitoring GPS locations. I calculated variance as

$$\hat{\text{var}}(\hat{N}) = \frac{(n_1+1)(n_2+1)(n_1-m_2)(n_2-m_2)}{(m_2+1)^2(m_2+2)},$$

standard error as $\hat{\text{SE}}(\hat{N}) = \sqrt{\hat{\text{var}}(\hat{N})}$**,** and the coefficient of variation (CV) as $\hat{\text{CV}} = \sqrt{\hat{\text{var}}(\hat{N})} / \hat{N} \cdot 100$ (Seber 1982). According to Chao (1989), I calculated log-normal 95% lower and upper confidence intervals, respectively, of $m + (\hat{N} - m)/C$, $m + (\hat{N} - m) \cdot C$, where $m = n_1 + n_2 - m_2$ and

$$C = \exp\left\{ z_{\alpha/2} \sqrt{\ln\left(1 + \frac{\hat{\text{var}}(\hat{N})}{(\hat{N}-m)^2}\right)} \right\}.$$

The L-P estimator assumes a closed population; thus, to meet this assumption, it is critical to survey the entire study area in one evening. Therefore, the current survey method used by park managers violated this assumption because it took multiple days to survey the entire study area, which could result in immigration or emigration and multiple observations of the same individual. Additionally, by only making one trip through each compartment, I expected that the sighting probability, and hence the number of deer seen, would be lower than when each compartment was surveyed continuously each evening.

Using the survey methods outlined by Storm et al. (1992), I surveyed the entire study area each evening; therefore, each evening provided an independent estimate of *N*. I calculated an arithmetic average of the L-P estimates for the 3 evenings, a standard error as the standard deviation of the 3 estimates, and log-normal 95% lower and upper confidence intervals, respectively, of $\hat{N}/C$, $N \times C$, according to Chao (1989), where

$$C = \exp\left\{1.96\sqrt{\ln\left(1+\frac{\hat{\mathrm{var}}(\hat{N})}{(\hat{N})^2}\right)}\right\}.$$

In addition, I used the joint hypergeometric maximum likelihood estimator (JHE) in program NOREMARK (Bartmann et al. 1987, White and Garrott 1990, Neal 1990, Neal et al. 1993, White 1996) to combine data from the 3 evenings for each survey into a single population estimate. The JHE is similar to the L-P estimator but uses data from multiple surveys to estimate *N* by maximizing the likelihood

$$£(N \mid M_i, n_i, m_i) = \prod_{i=1}^{k} \frac{\binom{M_i}{m_i}\binom{N-M_i}{n_i-m_i}}{\binom{N}{n_i}},$$

where M_i is the number of marked deer available each evening, n_i is the number of identified deer seen (marked plus unmarked) each evening, and m_i is the number of marked deer seen each evening. I used profile likelihood confidence intervals.

I calculated a new average detection probability for all deer for each dusk mark-resight survey as $\hat{\bar{p}}_j = \sum m_{2ij} \big/ \sum n_{1ij}$, where m_{2ij} equals the number of marked deer seen during the i^{th} day of the j^{th} survey and n_{1ij} equals the number of marked deer available during the i^{th} day of the j^{th} survey.

Bowden Estimator: Spotlight Surveys

The assumption of a closed population for the L-P estimator could not be met during spotlight surveys because it took 2 to 3 days to completely cover the study area. Therefore, I used Bowden's estimator to estimate abundance ($\hat{N}$) from each spotlight survey (Bowden and Kufeld 1995). The assumptions for Bowden's estimator are less restrictive than the assumptions for the L-P estimator and other mark-resight estimators (e.g., Bowden's estimator does not require the assumptions of homogeneous detection probabilities and independent sighting trials; Diefenbach 2009). Additionally, one does not have to survey the entire study area and the probability of sighting a deer can vary among individual deer over time (Bowden and Kufeld 1995, Diefenbach 2009). Bowden's estimator also assumes a geographically and demographically closed population, although temporary emigration is allowed (Bowden and Kufeld 1995). Other assumptions include individually identifiable marks (i.e., GPS-collars), marks are placed on a random sample of individuals, marks are not lost during the sighting attempts, and marked and unmarked deer are equally likely to be observed (Bowden and Kufeld 1995). Bowden's estimator estimates the number of deer that have used the study area at any point during the entire study (e.g., fate of marked individuals is typically unknown), whereas the L-P estimator estimates abundance of deer in the study area each evening (Bowden and Kufeld 1995). However, I estimated the number of deer that used the study area at any point during each survey period (e.g., 1 week) by monitoring GPS locations of the marked population.

For each mark-resight survey using Bowden's estimator, I let y_i be the number of sightings of marked deer *i*, $\bar{y} = \sum_{i=1}^{n}(y_i / n)$ be the mean number of times the *n* marked and available deer were sighted, $s_y^2 = \sum_{i=1}^{n}\left[(y_i - \bar{y})^2/(n-1)\right]$ be the sample variance of the mean sightings per marked deer, and Y be the total number of sightings of marked and unmarked deer (Bowden and Kufeld 1995). I determined the number of available marked deer (n) by monitoring GPS locations of all GPS-collared deer. If a marked deer was in the study area during the time that I performed a survey, then I considered it available for that survey.

An intuitive estimator of abundance is $\tilde{N} = Y/\bar{y},$ but tends to be positively biased with small sample sizes; therefore, I used an approximately unbiased estimator for N, developed by Bowden and Kufeld (1995), where

$$\hat{N} = \left(\tilde{N} + \frac{s_y^2}{\bar{y}^2}\right)\bigg/\left(1 + \frac{s_y^2}{n\bar{y}^2}\right),$$

with estimated variance

$$\hat{\mathrm{var}}(\hat{N}) = \hat{N}^2\left(\frac{1}{n} - \frac{1}{\hat{N}}\right)\left\{\left(\frac{s_y^2}{\bar{y}^2}\right)\bigg/\left(1 + \frac{s_y^2}{n\bar{y}^2}\right)^2\right\},$$

coefficient of variation (CV) $\mathrm{C\hat{V}} = \sqrt{\hat{V}(\hat{N})}\big/\hat{N}\cdot 100,$ and logarithm transformed 95% lower and upper confidence intervals, respectively, of $\hat{N}/C$ and $\hat{N}\cdot C,$ where

$$C = \exp\left(t_{1-\alpha/2,n-1}\sqrt{\left[\frac{1}{n} - \frac{1}{\hat{N}}\right]\frac{s_y^2}{\bar{y}^2}}\right).$$

RESULTS

Capture and Sample Sizes

From January 2009 to April 2009, I captured and GPS-collared 38 deer of all age and sex classes in the study area, of which 16 to 28 were available (present in the study area) during the dusk surveys (Table 1) and 18 to 31 were available during spotlight surveys (Table 2) prior to the second capture season. The total number available per survey decreased over time because of mortality, cast collars, dispersal, and temporary emigration from the study area, all of which were more prevalent in males (Table 1 and Table 2). At the start of the January 2010 trapping season, 18 GPS-collared deer from the 2009 trapping season were present on the study area, which I treated as adults during 2010. From January 2010 to April 2010, I captured 20 additional deer. After this second capture season, 14 to 23 GPS-collared deer were available during dusk surveys (Table 1) and 20 to 30 were available during spotlight surveys (Table 2).

Table 1. Number of marked white-tailed deer present in the study area, by age and sex class, during each complete round of dusk mark-resight surveys, Gettysburg, Pennsylvania, 2009-2010.

		Juveniles		Adults		
Survey	Round	Male	Female	Male	Female	Total
Apr 2009	1	6	3	5	11	25
	2	6	3	8	11	28
	3	6	3	6	11	26
Nov 2009	1	0	3	3	9	15
	2	1	3	4	9	17
	3	0	3	3	10	16
Apr 2010	1	3	1	6	13	23
	2	3	1	5	14	23
	3	3	2	6	12	23
Nov 2010	1	0	1	3	12	16
	2	0	1	5	11	17
	3	0	1	2	11	14

Table 2. Number of marked white-tailed deer, by age and sex class, that used the study area at any point during each spotlight mark-resight survey, Gettysburg, Pennsylvania, 2009-2010.

Survey	Juveniles Male	Juveniles Female	Adults Male	Adults Female	Total
Apr 9-16, 2009	7	3	9	12	31
Aug 3-9, 2009	7	3	8	12	30
Nov 20-25, 2009	1	3	4	10	18
Jan 5-8, 2010	1	2	5	11	19
Apr 1-4, 2010	4	2	9	15	30
Aug 25-30, 2010	3	1	8	12	24
Nov 15-19, 2010	0	2	6	12	20

Dusk Surveys

During the April 2009 survey, I used the current survey method used by park managers, where observers did not continuously survey each compartment each evening and it took multiple days to survey the study area completely. Thus, assumptions were violated and few marked and unmarked deer were observed relative to the number available (Table 3). No marked deer were observed for 2 of the 3 evenings; therefore, I could only estimate abundance for the first evening and could not estimate abundance using the JHE or calculate an arithmetic average of L-P estimates (Table 3, Fig. 2).

During the November 2009 survey, no marked deer were observed on the third evening (Table 3). However, using data from the first 2 evenings, I calculated a JHE abundance estimate of $\hat{N}$ = 574 (95% CI = 275 – 1,748) and an arithmetic average of L-P estimates of $\hat{N}$ = 410 (95% CI = 259 – 650, CV = 24), and $\hat{\bar{p}}$ = 0.13 (Table 3, Fig. 2). During the April 2010 survey, marked deer were seen during all 3 evenings, which allowed me to calculate a JHE abundance estimate of $\hat{N}$ = 414 (95% CI = 292 – 649), an arithmetic average of the L-P abundance estimates of $\hat{N}$ = 368 (95% CI = 322 – 421, CV=7), and $\hat{\bar{p}}$ = 0.25 (Table 3, Fig. 2). During the November 2010 survey, marked deer were seen during all 3 evenings, which allowed me to calculate a JHE abundance estimate of $\hat{N}$ = 465 (95% CI = 300 – 835), an arithmetic average of the L-P abundance estimates of $\hat{N}$ = 425 (95% CI = 196 – 921, CV = 41), and $\hat{\bar{p}}$ = 0.23 (Table 3, Fig. 2).

Table 3. Abundance estimates ($\hat{N}$) of white-tailed deer for the study area using the Lincoln-Petersen estimator and detection probabilities ($\hat{p}$) for each complete round of dusk mark-resight surveys, Gettysburg, Pennsylvania, 2009-2010.

Survey	Round	Day of Month	n_1 [b]	n_2 [c]	m_2 [d]	$\hat{N}$	SE	CV	95% CI	$\hat{p}$
Apr 2009[a]	1	7-9	25	48	2	421	193.5	46	198 - 1,034	0.08
	2	10-13	28	51	0					
	3	14-16	26	55	0					
Nov 2009	1	20	15	89	2	479	212.7	44	237 - 1,158	0.13
	2	21	17	56	2	341	151.9	45	168 - 826	0.12
	3	22	16	84	0					
Apr 2010	1	1	23	84	5	339	107.3	32	204 - 654	0.22
	2	2	23	94	5	379	120.4	32	227 - 732	0.22
	3	3	23	128	7	386	102.0	26	254 - 679	0.30
Nov 2010	1	15	16	92	5	263	77.5	29	168 - 496	0.31
	2	17	17	111	4	402	136.7	34	236 - 816	0.24
	3	18	14	121	2	609	269.4	44	302 - 1,471	0.14

[a] The April 2009 survey was performed such that a complete round took 3-4 days.
[b] n_1 = marked deer available.
[c] n_2 = total identified deer.
[d] m_2 = marked deer seen.

Spotlight Surveys

Estimates of abundance using the Bowden's estimator, ranged from 375 (95% CI = 279 – 502) to 691 (95% CI = 375 – 1,271; Table 4 and Fig. 2). The April 2009 and 2010 surveys, which occurred after each trapping season, had the largest *n* and yielded the most precise estimates across years with an average CV of 15% (Table 4). The August 2009 and November 2009 surveys, which coincided with high visitor use of the park (i.e., Labor Day in August and Remembrance Day in November), yielded the worst precision. Ignoring those two surveys, detection probability tended to be highest in August and April and slightly lower in November and January (Table 4).

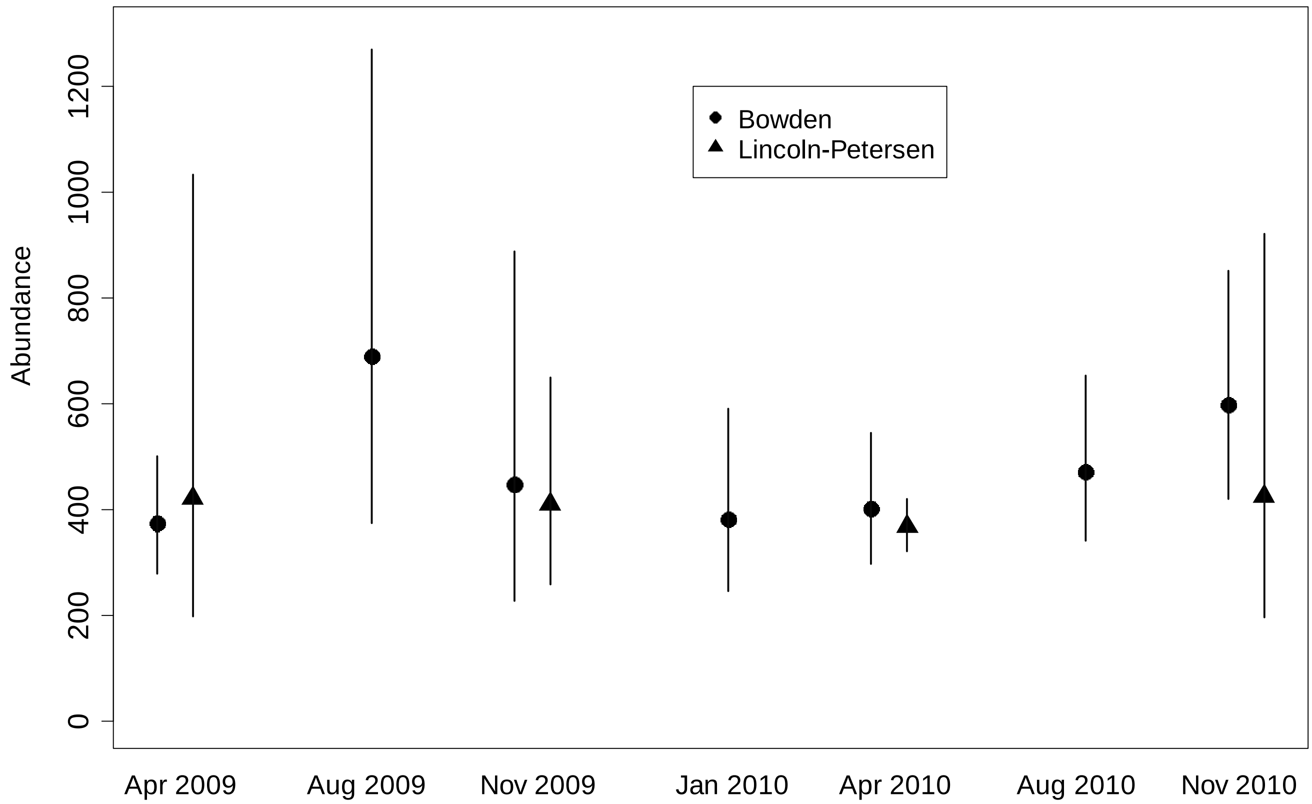


Figure 2. Abundance estimates ($\hat{N}$) of white-tailed deer and associated 95% confidence interval bars from mark-resight surveys using the Bowden estimator for spotlight surveys and an arithmetic average of the Lincoln-Petersen estimates for dusk surveys from April 2009 to November 2010 in the study area, Gettysburg, Pennsylvania.

Table 4. Abundance estimates ($\hat{N}$) of white-tailed deer and measures of precision using Bowden estimator for each spotlight mark-resight survey, Gettysburg, Pennsylvania, 2009-2010.

Survey	n [a]	Y [b]	$\bar{y}$ [c]	$\hat{N}$	SE	CV	95% CI
Apr 9-16, 2009	31	444	1.16	375	52.6	14	279 - 502
Aug 3-9, 2009	30	375	0.50	691	185.6	27	375 - 1,271
Nov 20-25, 2009	18	329	0.67	449	128.0	29	227 - 889
Jan 5-8, 2010	19	356	0.89	382	74.1	19	247 - 592
Apr 1-4, 2010	30	439	1.07	403	58.7	15	297 - 546
Aug 25-30, 2010	24	604	1.25	472	72.3	15	341 - 653
Nov 15-19, 2010	20	492	0.80	598	98.2	16	420 - 852

[a] n = total no. of individual marked deer that used the study area during the survey period.
[b] Y = total no. of identified deer seen during the survey period.
[c] $\bar{y}$ = mean no. of times an available marked deer was seen during the survey period.

Heterogeneity in Detection Rates

The few marked deer observed during dusk surveys precluded investigation of heterogeneity in detection rates for each dusk survey with respect to sex and age. I observed sufficient marked deer during spotlight surveys to investigate heterogeneity in detection rates, but estimates may only be relevant to detection at night and not at dusk. I found comparable rates of detection in adult males and adult females during the April surveys, higher rates in adult males than adult females during the August surveys and the January survey, and lower rates in adult males than adult females during the November surveys (Table 5). Detectability was low for both adult females and juvenile females during the August 2009 survey and for adult males during the November 2009 survey (Table 5).

Table 5. Mean no. of times an available marked white-tailed deer was seen ($\bar{y}$) during the entire survey period and mean per round ($\bar{\bar{y}}$), by age (A=Adult, J=Juvenile) and sex (M=Male, F=Female), for each spotlight mark-resight survey using Bowden's estimator, Gettysburg, Pennsylvania, 2009-2010. Estimates not calculated when < 3 marked deer were available.

	$\bar{y}$				$\bar{\bar{y}}$			
Survey	AM	AF	JM	JF	AM	AF	JM	JF
Apr 9-16, 2009	1.11	1.18	1.00	1.40	0.37	0.39	0.33	0.47
Aug 3-9, 2009	0.75	0.18	1.00	0.20	0.25	0.06	0.33	0.07
Nov 20-25, 2009	0.00	0.80		1.33	0.00	0.29		0.48
Jan 5-8, 2010	1.20	0.82			0.48	0.32		
Apr 1-4, 2010	1.00	0.93	1.75		0.33	0.31	0.58	
Aug 25-30, 2010	1.50	1.08	1.33		0.50	0.36	0.44	
Nov 15-19, 2010	0.67	0.92			0.22	0.31		

DISCUSSION

I estimated the average detection probability of both marked male and female deer during April and November dusk counts, which were, as I hypothesized, lower than the Storm et al. (1992) study for all 4 surveys (Table 3). Additionally, I estimated abundance during dusk counts using the L-P estimator and during spotlight surveys using Bowden's estimator. However, an important assumption for accuracy in average detection probability and L-P estimates is that rates of detection are constant across individuals within a given survey (Seber 1982). I found that detection rates varied by age and sex for each survey, with the exception of the April surveys, where adult males and adult females were detected at similar rates (Table 5). Therefore, if the marked sample was truly representative of the unmarked population, bias related to violating the assumption of homogenous detection probability would be minimized during surveys conducted in early April.

Several studies suggested that sightability of bucks is generally lower than sightability of does throughout the year, but increases during the summer months (Downing et al. 1977, McCullough et al. 1982, Sage et al. 1983). My results are consistent with these studies for the August survey, but I only observed lower rates of detection for adult males than for adult females during the November survey (Table 5). However, results of detection probability for the study area at GNMP-ENHS may not be representative of typical land in other areas of Pennsylvania. The NPS culls only antlerless deer on NPS owned property in the study area, but otherwise it is closed to hunting. Additionally, not all privately owned property in the study area is open to hunting. Therefore, bucks may exhibit greater rates of detection relative to does because antlerless deer are more likely to be culled, such that highly detectable antlerless deer are removed from population, but bucks are not, unless they are harvested off the park during the hunting season. Storm et al. (1992) occurred prior to culling, so although I concluded rates of detection between adult males and adult females were similar in April, rates of detection between males and females may not have been similar during that study, where detection probability was based on marked female deer.

Although I observed heterogeneous detection probabilities, estimates of average detection probability would be representative of the population if my marked sample represented all age and sex classes of the population (Otis et al. 1978, Seber 1982, White et al. 1982). Additionally, abundance estimates from the L-P and Bowden's estimators were similar (Fig. 2), suggesting that individual heterogeneity had minimal effect on accuracy of abundance estimates within a given survey period because the L-P estimator assumes homogeneous detection probabilities and Bowden's estimator does not. However, variance associated with L-P estimates may not be accurate. Additionally, caution should be used when comparing L-P and Bowden's estimates of abundance (Fig. 2). The L-P estimator estimates the number of deer in the study area during a given evening, whereas Bowden's estimator estimates the number of deer that used the study area during the entire survey period (e.g., 1 week) and is greater when temporary emigration occurs (Bowden and Kufeld 1995).

Additionally, for the average detection probability to yield unbiased estimates of abundance in future surveys, the assumption that detection probability will not change over time must be met. Therefore, park managers can use my estimate of average detection probability from the April 2010 survey of 0.25 as their updated sighting probability for future April dusk surveys, but they would have to assume that detection probability stays constant over time. Even if the same observers conduct future surveys, factors related to the environment (e.g., habitat conditions) and deer behavior may change over time (Anderson 2001). Therefore, using the sighting probability that I calculated may provide future density estimates that are inaccurate, such that detection probability may need to be updated again. For instance, park managers may observe an increasing trend in abundance over time using a constant sighting probability when abundance is actually stable because the true detection probability increased.

The April 2010 survey provided the greatest precision for both the L-P and Bowden's estimators (CV = 7% and 15% respectively; Fig. 2). Visibility was excellent during the April and January surveys because there was no foliage on deciduous trees and shrubs, nor any crops or high vegetation in fields. However, vegetation decreased visibility for August surveys and standing corn decreased visibility for August and November surveys, leading to less robust estimates. Additionally, the deer population is at its peak in the summer because of fawn recruitment and subsequently decreases because of mortality (e.g., NPS culling begins before the November survey; Fig. 2).

However, because the true population size is unknown, I could not evaluate accuracy. Park managers can use my estimates of abundance to update management strategies to reach their long-term density goal of 10 deer/km^2 (25 deer/mi^2) of forested land or 85 deer on the study area (Frost et al. 1997). I estimated density as 43 deer/km^2 (112 deer/mi^2) of forested land using the arithmetic average of L-P abundance estimates from the April 2010 dusk survey (Table 3). This is a reduction of 68% from the estimated density of 136 deer/km^2 of forested land in 1992 (Tzilkowski and Storm 1993), but four times greater than the park's goal of 10 deer/km^2 of forested land.

CHAPTER 3:

TESTING ASSUMPTIONS OF DISTANCE SAMPLING USING ROADS

INTRODUCTION

Distance sampling using line transects is a generalization of the strip transect sampling method, in which all objects within sample strips are detected (Buckland et al. 2001). Distance sampling allows a proportion of objects to be missed away from the line or transect, thus allowing a wider strip to be sampled and increasing sample size and efficiency (Buckland et al. 2001). Distance sampling often provides a practical, cost-effective method of estimating density for a broad range of applications, from walking transects to detect inanimate objects or plants in a terrestrial setting to traversing transects in a ship to detect moving objects such as whales in a marine setting (Thomas et al. 2010).

The appeal of the distance sampling estimator over mark-recapture methods, is that the method does not require marked animals and could be far less expensive when used to estimate population size (Focardi et al. 2002). However, it may be difficult to meet all assumptions to obtain accurate population estimates. A critical requirement or assumption with line-transect distance sampling is that randomly placed transects are used (Buckland et al. 2001, Thomas et al. 2010). A systematic placement of parallel transects located at a random starting point generally meets this design requirement, because these transects are randomly located and they sample across the area of interest (Buckland et al. 2001).

Random transects are critical to assume: (1) that regardless of the distribution of objects on the landscape, the distribution of objects is uniform with respect to perpendicular distances to transects (e.g., the uniformity requirement or assumption;

Fewster et al. 2008, Marques et al. 2010); and (2) that data collected are a representative sample of the population (e.g., the density of deer in the sample is representative of the larger area of interest; Buckland et al. 2001). The first is critical to model an unbiased estimate of detection probability as the decrease in the number of observations as distance from the transect increases; and the second is critical to extrapolate estimates from the sampled area to the larger area of interest (Buckland et al. 2001).

Further, deer density is often correlated to habitat types (e.g., forest cover vs. open areas), which can also influence the proportion of objects detected. Therefore, if transects are randomly placed, then the proportion of each habitat type in the sampled area should represent the proportion of each habitat type in the larger area of interest. The preferred sampling method to use when multiple habitats exist is to stratify transects by habitat (e.g., randomly place transects in fields and then randomly place additional transects in forests) in proportion to their occurrence in the whole study area (Buckland et al. 2001).

Existing roads have been used as transects with distance sampling (e.g., Gill et al. 1997, Heydon et al. 2000, Koganezawa and Li 2002**,** Ruette et al. 2003, Ward et al. 2004, Bates 2006), but an important concern is that roads are not randomly placed on the landscape and therefore, may not provide a representative sample of the study area (Anderson 2001, Buckland et al. 2001). If unrepresentative, surveys conducted from roads may only provide an estimate of density of the population near roads, which may have limited value for making management decisions (Buckland et al. 2001). Furthermore, if the distribution of deer was correlated with the location of roads, perhaps because the location of roads was correlated to habitat types important to deer, then the estimator for detection probability may be biased, leading to a biased estimator of density. The direction of the bias would depend on whether deer were avoiding or selecting for areas near roads, and the magnitude of the bias would depend on the amount of non-uniformity of the distribution of deer relative to transects. To my knowledge, no studies have tested both the uniformity requirement and representativeness of a sample from roads, or appropriately corrected for all bias.

If the distribution of deer relative to perpendicular distance to transects is not uniform, than detection probability will be biased. To demonstrate potential bias in detection probability, I created 5 hypothetical examples of the distribution of deer relative to the distribution of transects (dashed line) and compared these distributions to assumed distributions from distance sampling observation data (dotted line at y-intercept of solid line of detection function; Figs. 3 – 7). When there are fewer observations of objects near transects, the detection function y-intercept is typically an average of data from the first few bins. Because the true distribution of deer and the assumed distribution from observation data can have different y-intercepts, a simplified way to visualize the direction of the bias in the detection probability from a distance sampling survey under each scenario is to subtract the area under the dotted line from the area under the dashed line. A positive value indicates that detection probability would be positively biased from observation data, resulting in a negatively biased estimator of abundance.

The first scenario (Fig. 3) is the assumed distribution with random transects, where deer are distributed uniformly relative to perpendicular distance from transects. The actual distribution of deer (dashed line) is equal to the assumed distribution (dotted line), such that no bias is expected in detection probability (Fig. 3). The second scenario (Fig. 4) is where deer avoid areas near transects, then are distributed uniformly relative to perpendicular distance from transects after some distance x. The detection probability is actually 1.0 near transects, but there are less deer there to see (Fig. 4). Thus, detection probability would be positively biased because the assumed proportion of missed observations beyond distance x is less than actual (Fig. 4). The third scenario (Fig. 5) is where deer are distributed non-uniformly relative to perpendicular distance from transects, such that avoidance of areas near transects increases with distance from the transect. From the observation data, it would appear that detection probability is very high, such that few deer were missed (Fig. 5). However, because there are more deer at greater distances to see, and therefore miss, the detection probability would be positively biased (Fig. 5). The fourth scenario (Fig. 6) is the opposite of the third, where deer are distributed non-uniformly relative to perpendicular distance from transects, such that deer are attracted to transects and attraction decreases with distance from the transect. The detection function drops off quickly, and the assumed proportion of deer that are missed is greater than actual because there are less deer to see at further distances; therefore, detection probability would be negatively biased (Fig. 6). The last scenario (Fig. 7) is where deer are distributed non-uniformly relative to perpendicular distance from transects, such that deer avoid areas near transects and areas far away from transects. This is similar to the second scenario, except the number of deer decreases at further distances (Fig. 7). The assumed proportion of deer missed at further distances is greater than actual, but the assumed proportion of deer missed at intermediate distances is less than actual, such that overall, detection probability would be positively biased (Fig. 7).

These hypothetical examples demonstrate that without knowledge of the true distribution or density gradient of the object of interest with respect to the distribution of transects, inspection of the detection histogram provides no insight into the true detection function, unless assumed uniform because random transects were used. For example, the detection histogram from the fourth scenario (solid bins in Fig. 6) shows no avoidance of areas near the transect. Without knowledge that the object of interest is more prevalent near transects, abundance would be overestimated.

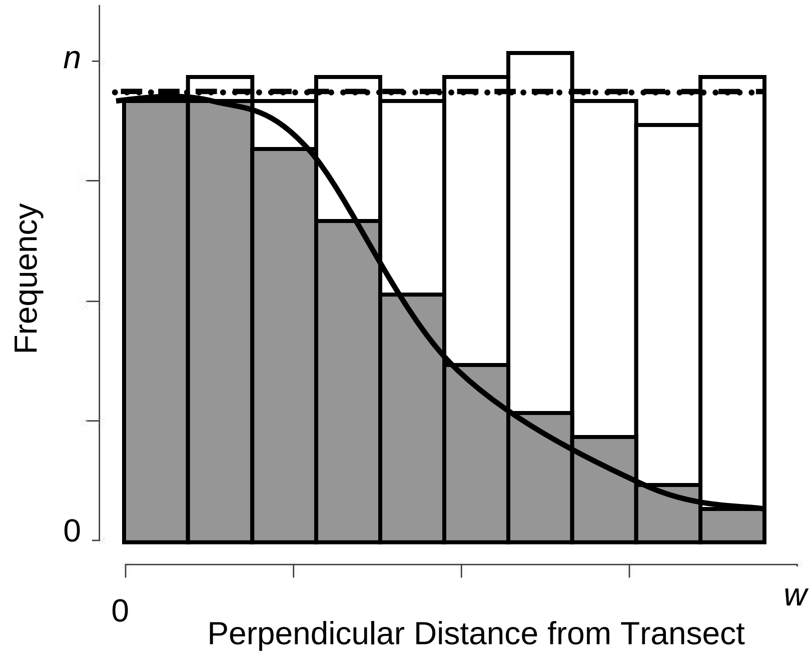


Figure 3. Hypothetical example where deer are distributed uniformly relative to perpendicular distance from transects. The solid portions of bins indicate observed deer from distance sampling survey and the open portion of bins indicate missed deer. The solid line is the fitted detection function, the dashed line is the true distribution of deer relative to the distribution of transects, and the dotted line is the assumed distribution of deer based on observations. This is the assumed distribution with random transects (i.e., dashed and dotted lines match), such that no bias is expected.

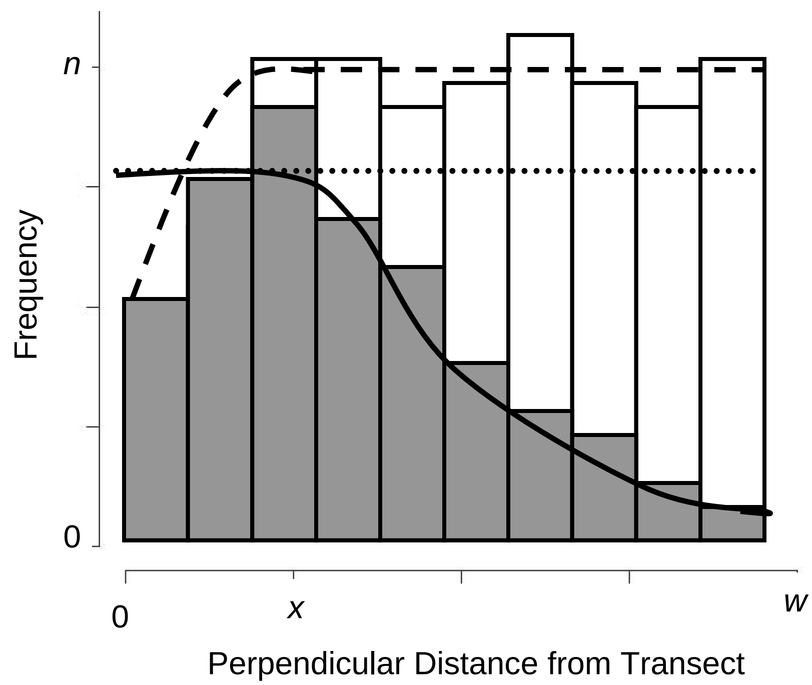


Figure 4. Hypothetical example where deer are avoiding areas near transects, but then distributed uniformly relative to perpendicular distance from transects after some distance *x*. The solid portions of bins indicate observed deer from distance sampling survey and the open portion of bins indicate missed deer. The solid line is the fitted detection function, the dashed line is the true distribution of deer relative to the distribution of transects, and the dotted line is the assumed distribution of deer based on observations. The detection probability would be positively biased with this distribution.

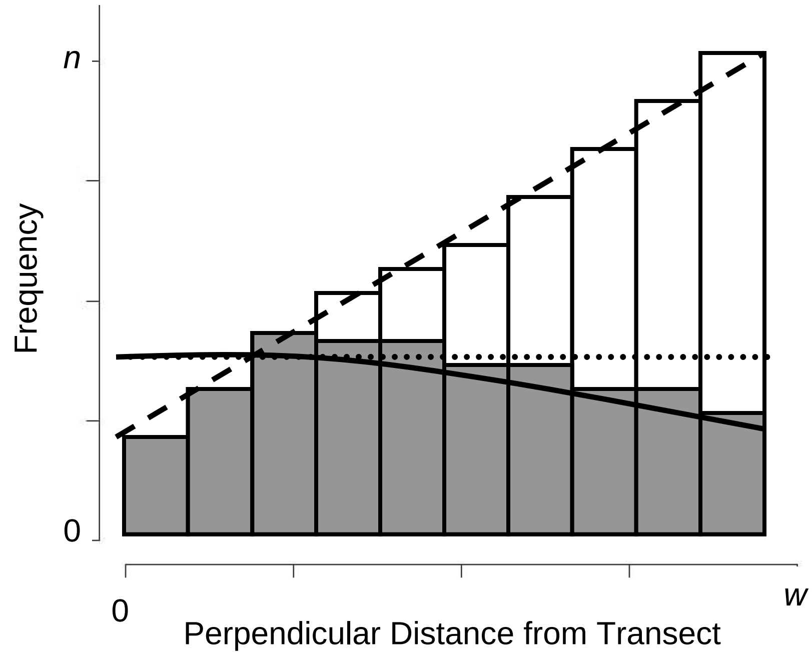


Figure 5. Hypothetical example where deer are distributed non-uniformly relative to perpendicular distance from transects, such that they are avoiding areas near transects. The solid portions of bins indicate observed deer from distance sampling survey and the open portion of bins indicate missed deer. The solid line is the fitted detection function, the dashed line is the true distribution of deer relative to the distribution of transects, and the dotted line is the assumed distribution of deer based on observations. The detection probability would be positively biased with this distribution.

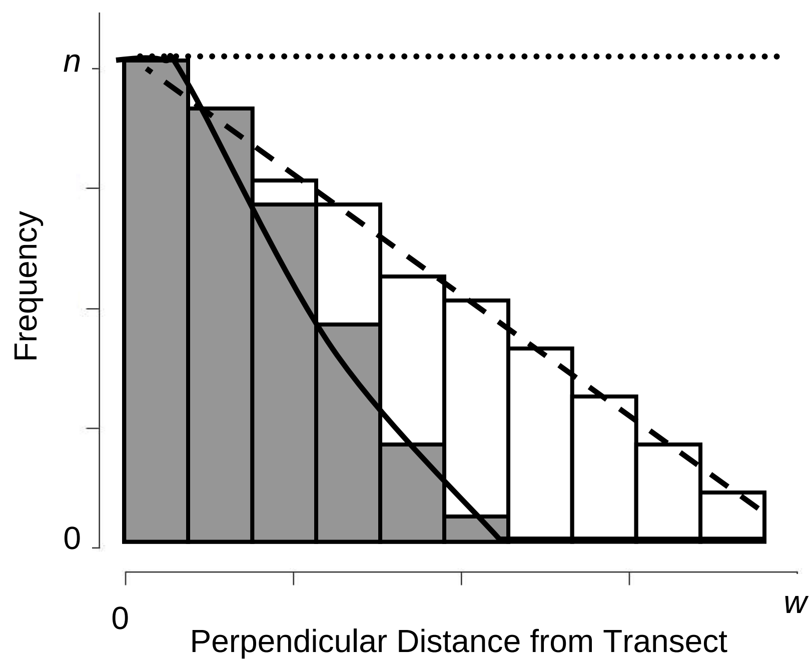


Figure 6. Hypothetical example where deer are distributed non-uniformly relative to perpendicular distance from transects, such that they are selecting for areas near transects. The solid portions of bins indicate observed deer from distance sampling survey and the open portion of bins indicate missed deer. The solid line is the fitted detection function, the dashed line is the true distribution of deer relative to the distribution of transects, and the dotted line is the assumed distribution of deer based on observations. The detection probability would be negatively biased with this distribution.

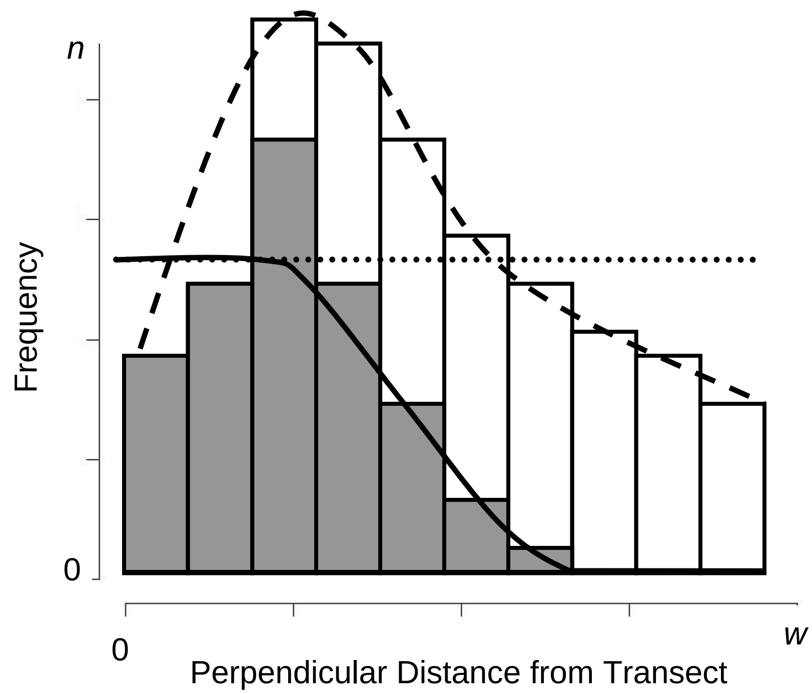


Figure 7. Hypothetical example where deer are distributed non-uniformly relative to perpendicular distance from transects, where they are avoiding areas near transects and far from transects. The solid portions of bins indicate observed deer from distance sampling survey and the open portion of bins indicate missed deer. The solid line is the fitted detection function, the dashed line is the true distribution of deer relative to the distribution of transects, and the dotted line is the assumed distribution of deer based on observations. The detection probability would be positively biased with this distribution.

To investigate whether using roads as transects violated critical assumptions, I collected GPS locations from marked deer on the study area during distance sampling surveys. I modeled these GPS locations with respect to the roads I used as transects to test whether the distribution of deer relative to the distribution of transects was uniform. Additionally, I used these GPS locations with landscape covariates to model relative habitat use of deer on the study area, which I used to test whether the roads I chose as transects provided a representative sample of the study area. Then, I developed methods to adjust for potential bias in detection probability when deer were distributed non-uniformly relative to transects and used the model of relative habitat use to adjust for a possible non-representative sample. My final objective was to compare bias adjusted estimates of abundance for the study area to estimates of abundance when violations of assumptions were ignored.

STUDY AREA

The study area, outlined by Storm et al. (1992), encompassed 2,913 ha (7,197 acres or 11.25 mi^2) of both public and private land surrounding Gettysburg, Pennsylvania (refer to Ch. 2 for more detail). I selected survey transects for distance sampling from existing roads within the study area (Fig. 8). I identified 26 survey routes or transects of similar length (range = 0.43–3.46 km, mean length = 1.83 km) rather than a few long routes to better estimate the variance related to encounter rate (Buckland et al. 2001). I chose as many roads as possible in the study area for more complete coverage, but selected roads with less vehicular traffic, such as NPS roads (e.g., closed to public travel after hours) rather than highways, for safety reasons. Transects included only segments of roads where spotlights could be used (e.g., sections near buildings, livestock, etc. were excluded).

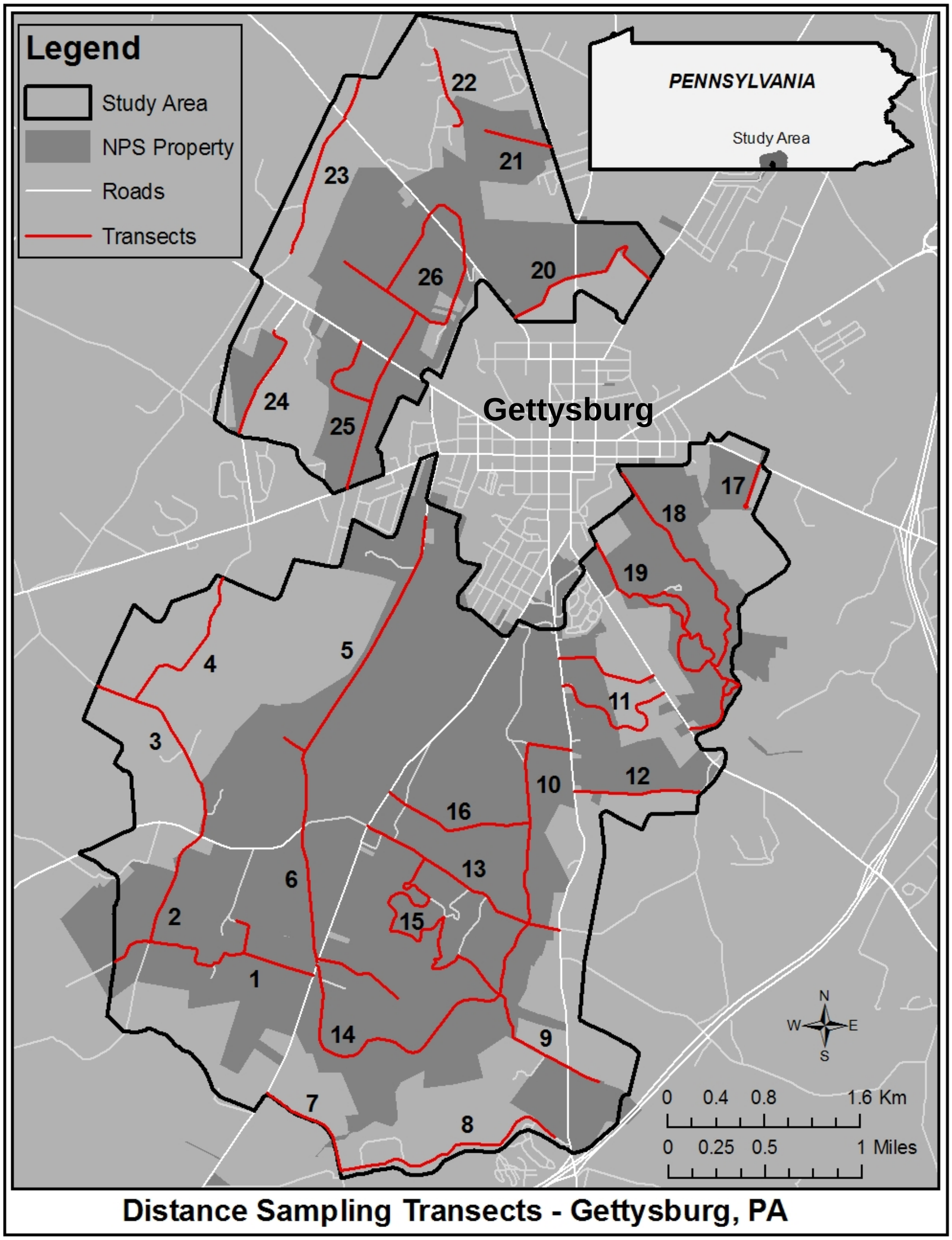


Figure 8. The 26 transects used during distance sampling and spotlight surveys for white-tailed deer, performed in the study area surrounding Gettysburg, Pennsylvania from April 2009 to November 2010. Roads not used as transects are shown in white. The areas in dark gray are National Park Service (NPS) owned property and areas in light gray are privately owned property.

METHODS

Global Positioning System Collars

I fitted all captured deer (refer to Ch. 2 for capture methods) with GPS-collars. These collars allowed me to record 1 location every 5 minutes (approximately 48 locations per deer per night) during the time that I conducted surveys.

Distance Sampling Surveys

I conducted distance sampling surveys during spotlighting mark-resight surveys at GNMP-ENHS prior to culling efforts (August), in the middle of culling efforts (November), after the culling operation (January), and following the capture season (April). I started surveys no earlier than 30 minutes after sunset, and surveys lasted approximately 3-5 hours. Two observers illuminated their respective sides of the transect with handheld spotlights while standing in the bed of a pick-up truck. When deer were detected, I recorded group size, distance and direction to the group, x-y coordinates of the observer, and whether the deer was located in a field or the forest. I defined groups based on behavioral cues and proximity to one another. Each deer in a group was no more than one-half the distance from the closest deer in its group than to the next closest deer of a neighboring group. I obtained distance and direction using a handheld PC/GPS unit (HP–iPAQ rx5900, Hewlett-Packard Company, Palo Alto, CA, USA), running Cybertracker data collection software (http://www.cybertracker.co.za), linked to a laser rangefinder (LTI-TruPulse 360, Laser Technology, Inc., Centennial, CO, USA) via Bluetooth. I calculated perpendicular distances as the shortest distance between the transect and the location of each observation, using a GIS.

I traversed transects at 10-25 km/hr and varied initial starting points to minimize temporal influences in deer detection that may have existed because of deer activity patterns. I divided transects into 3 groups and surveyed each night based on geographic location (e.g. – North, Southwest, and Southeast), such that it took approximately 2 to 3 nights to survey all transects once and approximately 6 to 9 nights to survey all transects 3 times. I did not survey on a particular night if adverse environmental conditions existed (wind ≥32 km/hr, rain, visibility ≤1.6 km).

Distance Sampling Analysis

I used program DISTANCE (Thomas et al. 2010) to estimate density of deer groups and employed a size-bias regression method to model group size as a function of distance from the transect. If this regression was not significant ($\alpha = 0.05$), I used mean group size. Because the detection function is likely different for open areas than for wooded areas, I used the habitat type for each observation (field or forest) as a covariate, using multiple covariate distance sampling (MCDS). To account for differences in observer detection rates (see Diefenbach et al. 2003), I tested additional models including both habitat type and observer as covariates. I used both half-normal and hazard-rate key functions to model the detection function. I constrained models to use no adjustment terms to ensure the detection function was monotonically non-increasing (Marques et al. 2007). I used right truncation distances of 250 meters and 80 meters, because few observations occurred beyond these distances in fields and forests, respectively. I used Goodness of Fit tests and Akaike's Information Criterion (AIC; Burnham and Anderson 1998) as aids in model selection for the detection function curve.

Resource Selection Model

I collected GPS locations of deer during the time I performed distance sampling surveys to model habitat use by creating a Resource Selection Function (RSF; Manly et al. 2002) for each survey using a zero-inflated negative binomial model (ZINB; count model was negative binomial with a log link and zero-inflated model was binomial with a logit link) for each deer. In a GIS, I randomly placed 3,000 sampling units (100 m diameter circles) across the study area. The response variable or measure of use (y) was the number of GPS locations in each sampling unit, and I used habitat covariates (X_k) as predictor variables in models to estimate each parameter (β_k), such that the $RSF = e^y$, *where* $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots \beta_k X_k$ (Millspaugh et al. 2006, Sawyer et al. 2006). I measured habitat covariates as the percentage of forested land within each sampling unit and distance (m from center of each sampling unit) to the nearest road, forest-field edge, and NPS property boundary. To reduce edge effects, I allowed center points of circles to extend to the study area boundary, and quantified covariates outside of the study area. I fitted ZINB models with the pscl library in program R (Bates et al. 2008 and R Developmental Core Team, 2009; Appendix K). I used the zero-inflated negative binomial model because count data are oftentimes overdispersed (negative binomial model) and ≥95% of sampling units contained no locations for each deer (zero-inflated model). In addition, I used offset terms equal to the natural logarithm of the total number of GPS locations per deer to scale the RSF equally among deer. To aid in model convergence I normalized each covariate $\left[(X_k - \bar{X}_k)/SD(X_k)\right]$.

I developed 8 models selected *a priori* that always included an intercept and various combinations of habitat covariates that I believed were important for selection or avoidance of deer in the study area (Table 6). I predicted a positive relationship between deer selection and increasing distance from the nearest road, such that deer would avoid areas near roads. I also modeled a quadratic relationship to allow for nonlinearity, where I predicted deer would avoid areas very close to roads the most, but avoidance would become less critical as distance from the road increased. Because forested areas are relatively sparse in the study area (<30% is forested) and open areas tended to be large, I predicted a positive relationship between deer selection and percent forest. I also modeled a quadratic relationship to allow for nonlinearity and incorporate a possible threshold level of fragmentation. However, I believed there would be seasonal variation (e.g., row crops such as corn in the summer and early fall may decrease selection of heavily forested areas). I predicted a negative relationship between deer selection and distance to forest-field edges, such that during surveys, deer would select for open areas near forest edges. In addition, I modeled a quadratic relationship to allow for nonlinearity. Finally, I predicted a negative relationship between deer selection and distance to NPS boundary, such that deer would select for areas close to the park boundary and near private property rather than in the interior. The NPS shoots deer primarily in the central regions of the park, and there are more anthropogenic disturbances because of park visitors. Furthermore, I included an interaction of distance to the nearest road and percent forest in model 7 (Table 6) because open areas and fields tend to be near roads (e.g., roads provide access to fields). Also, I included an interaction between percent forest and distance to the nearest forest-field edge in model 8 (Table 6) to examine whether deer were selecting for forest-field edges near heavily forested areas or in more open and fragmented areas.

I used an information-theoretic approach (Burnham and Anderson 2002) to compare models identified *a priori* (Table 6). I calculated an AIC value for each model for each deer, and then selected the model with the lowest sum of AIC values across deer as the best model. Using the best model, I calculated parameter estimates and weighted each i^{th} parameter estimate for each j^{th} deer ($\hat{\beta}_{ij}$) by the inverse of its estimate of standard error, x_{ij}, such that $x_{ij} = \hat{\beta}_{ij} * \left\{1 / \hat{S}E(\hat{\beta}_{ij})\right\}$. The weighted average for each i^{th} parameter for all deer is then

$$\bar{x}_i = \frac{\sum\left(\hat{\beta}_{ij} * \frac{1}{\hat{S}E(\hat{\beta}_{ij})}\right)}{\sum\frac{1}{\hat{S}E(\hat{\beta}_{ij})}}.$$

I calculated a weighted population variance estimate for each i^{th} parameter using each j^{th} deer as,

$$S_i^{\,2} = \frac{V_{1i}}{V_{1i}^{\,2} - V_{2i}} \sum_{j=1}^{n}\sum_{i=1}^{n'} w_{ij}(x_{ij} - \bar{x}_i)^2,$$

where $w_{ij} = 1/\sigma_{ij}^{\,2}$, $V_{1i} = \sum_{j=1}^{n}\sum_{i=1}^{n'} w_{ij}$, $V_{2i} = \sum_{i=j}^{n}\sum_{i=1}^{n'} w_{ij}^{\,2}$, n was the total number of deer, n' was the total number of parameters, and $\sigma_{ij}^{\,2}$ was the variance for each i^{th} parameter for each j^{th} deer.

Table 6. Zero-inflated negative binomial models selected *a priori* for analyzing resource selection for each white-tailed deer collared with a global positioning system device during each distance sampling survey, Gettysburg, Pennsylvania, 2009-2010.

Model	k[a]	Covariates[b]
1	5	I + PF + FFE
2	7	I + PF + PF^2 + FFE + FFE^2
3	7	I + PF + FFE + NPS + RD
4	8	I + PF + FFE + FFE^2 + NPS + RD
5	9	I + PF + PF^2 + FFE + FFE^2 + NPS + RD
6	11	I + PF + PF^2 + FFE + FFE^2 + NPS + NPS^2 + RD + RD^2
7	9	I + PF + FFE + FFE^2 + NPS + RD + PF×RD
8	10	I + PF + FFE + FFE^2 + NPS + RD + PF×FFE + PF×FFE^2

[a] k = no. of model parameters.

[b] I = Intercept, PF = Percent Forest, FFE = distance to nearest forest-field edge, NPS = distance to nearest National Park Service owned land boundary, RD = distance to nearest road, and a multiplication sign indicates interaction terms.

Resource Selection Map

I used a GIS to create a 5 m × 5 m grid of cells across the study area. Next, I quantified the same distance covariates used in the resource selection model to the center of each grid cell and used a 100 m diameter circle around each grid cell center point to quantify percent forest. Then, I calculated an RSF value for each grid cell as $RSF = e^y$, where $y= \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \ldots\beta_kX_k$, based on the weighted parameter estimates from the selected ZINB model. Greater RSF values represented greater use by deer.

Test Assumption 1: Are Deer Uniformly Distributed with Respect to Roads?

To test whether the distribution of deer relative to the distribution of transects was uniform, I modeled the distribution of GPS locations from GPS-collared deer with respect to the perpendicular distance to each transect for locations in open areas and locations in forests using program DISTANCE. First, I measured the distance from every GPS location to the nearest section of each transect using a GIS. Then, using DISTANCE, I fitted a uniform curve with no adjustment terms and right-truncated at 80 m and 250 m. In addition, I fitted a combination of key functions (uniform, half-normal, hazard-rate) and series expansions (cosine, simple polynomial, and hermite polynomial), and used automated selection of up to 3 adjustment terms based on AIC. Finally, I selected no constraints on the shape functions, which eliminated the restriction that the curve be monotonically non-increasing. If the uniform curve with no adjustment terms had the lowest AIC, then I concluded that the distribution of deer was uniform (not correlated) with respect to the distribution of transects.

Test Assumption 2: Do Roads Provide a Representative Sample of the Study Area?

In a GIS, I calculated the percentage of forested land near transects and compared that to the percentage of forested land within the entire study area. If transects were representative of the study area, the proportion of forested land within the sampled area would be similar to the proportion of forested land in the study area.

Furthermore, I used information from the RSF map to investigate whether resource selection near transects was representative of resource selection across the study area. I used the RSF to estimate the proportion of the study area population near transects during each survey, $\hat{p}_{RSF}$, where $\hat{p}_{RSF} = \frac{\sum RSF(survey\ zone)}{\sum RSF(study\ area)}$. I calculated $\hat{p}_{RSF}$ for two different truncation distances (w = 80 m and w = 250 m), termed survey zones (Appendix C). If transects were representative of the study area, I would expect $\hat{p}_{RSF}$ to equal the proportion of the study area surveyed using both 80 m and 250 m transects. I calculated standard errors and 95% confidence intervals for $\hat{p}_{RSF}$ using parametric bootstrapping, incorporating the standard errors for each parameter in the RSF model. If the 95% confidence interval for $\hat{p}_{RSF}$ overlapped the proportion of the study area surveyed, then I failed to reject the null hypothesis that transects were representative.

Correction Factor

To reduce the bias in detection probability, I estimated abundance close to transects, where detection probability is close to 1.0 (i.e., I saw most of the animals) for both fields and forests. The trade-off is that as width decreases, sample size decreases, and thus variance increases. I tested two truncation distances (w=80 m and w=250 m). However, only the 80 m survey zone would be considered narrow enough to possibly correct for the bias in detection probability (e.g., detection probability close to 1), but wide enough to include enough observations for reasonable estimates.

If all assumptions were met, abundance in each survey zone would simply be the estimate of density in each surveyed area multiplied by the area of each zone. However, it may not be reasonable to use the estimate of density from the MCDS distance sampling analysis for my calculations because, although I did use habitat type as a covariate, I sampled both field and forest on the same transects. Further, DISTANCE assumes the sampled area is representative of the study area. Therefore, it calculates the global probability of detection based on the number of observations in each habitat, and does not incorporate the proportion of each habitat. If the proportion of forests and fields sampled from roads was not representative of the study area then the MCDS estimator may be biased.

Fortunately, density is unnecessary when used to calculate abundance within the surveyed area. This is because

$$\hat{N}_{(survey\ zone)} = \hat{D} \cdot Area_{(survey\ zone)} = \frac{n \cdot E(S)}{2wL\hat{P}_a} \cdot Area_{(survey\ zone)} = \frac{n \cdot E(S)}{\hat{P}_a} \ ,$$

where the area of each survey zone ($Area_{(survey\ zone)}$) = 2*wL*, such that *w* is the width or right-truncation distance and *L* is the total length of transects. Thus, area cancels out of the equation, yielding the abundance within the surveyed area as the number of groups detected (*n*) within *w* multiplied by the mean cluster size $E(S)$ and divided by the probability of detection ($\hat{P}_a$). However, because the detection probability from the MCDS analysis may be biased if the habitat is not representative, I needed to incorporate the relative proportions of forests and fields in the sampled area into the estimate of detection probability.

I calculated average detection probabilities for fields ($\overline{p}_{Field}$) and forests ($\overline{p}_{Forest}$) separately and applied those across the area of each survey zone based on the proportion of each habitat. I calculated $\overline{p}_{Field}$ and $\overline{p}_{Forest}$ by first modeling detection functions for fields and forests separately using conventional distance sampling methods (CDS) for detections in each survey zone (80 m and 250 m) for each survey month. Because of the lack of observations in forests, I pooled observations across years by survey month to increase sample size and to aid in model convergence. Models included different combinations of key functions (uniform, half-normal, and hazard-rate) and series expansions for up to 3 adjustment terms (cosine, simple polynomial, and hermite polynomial) chosen with sequential automated selection based on AIC. I constrained models to be strictly monotonically non-increasing. I used Goodness of fit tests and AIC (Burnham and Anderson 1998) as aids in model selection for the best fit of the detection function curves for fields and forest. Additionally, I required best models to have a wide shoulder (high detection probability near transects).

Next, I created a 5 m × 5 m grid over each survey zone and calculated a detection probability value, $\hat{p}_i$, for each i^{th} grid cell based on its distance from the nearest transect and whether it was predominated by a field or forest (i.e., y value from the detection curve for fields or forests), such that

$$\overline{p}_{Field} = \frac{\sum \hat{p}_{i(Field)}}{n_{(\#grid\ cells\ in\ Fields)}}$$

and

$$\overline{p}_{Forest} = \frac{\sum \hat{p}_{i(Forest)}}{n_{(\#grid\ cells\ in\ Forests)}}.$$

I calculated standard errors and 95% confidence intervals for $\overline{p}_{Field}$ and $\overline{p}_{Forest}$ using parametric bootstrapping based on 1,000 samples.

I calculated abundance in each survey zone as,

$$\hat{N}_{(survey\ zone)} = \frac{\overline{n}_{Field}}{\overline{p}_{Field}} \cdot E(S)_{Field} + \frac{\overline{n}_{Forest}}{\overline{p}_{Forest}} \cdot E(S)_{Forest},$$

where $\overline{n}_{Field}$ was the number of observed groups of deer in fields divided by the number of complete rounds, $\overline{p}_{Field}$ was the average detection probability in fields, $E(S)_{Field}$ was the mean group size in fields, $\overline{n}_{Forest}$ was the number of observed groups of deer in forests divided by the number of complete rounds, $\overline{p}_{Forest}$ was the average detection probability in forests, and $E(S)_{Forest}$ was the mean group size in forests. I calculated each variable separately for each survey zone. In addition, I calculated standard errors and 95% confidence intervals for $E(S)_{Field}$ and $E(S)_{Forest}$ using non-parametric bootstrapping, where I sampled with replacement from the observed group sizes, such that I created 1,000 estimates of mean cluster size for fields and forests for each survey.

I calculated abundance for the entire study area for each survey as

$$\hat{N} = \frac{\hat{N}_{(survey\ zone)}}{\hat{p}_{RSF}},$$

using abundance estimates for each survey zone divided by the proportion of the study area population present in each survey zone. I calculated standard errors and 95% confidence intervals for $\hat{N}$ by using the 1,000 bootstrapped estimates of $\overline{p}_{Field}$, $\overline{p}_{Forest}$, $E(S)_{Field}$, $E(S)_{Forest}$, and $\hat{p}_{RSF}$ in the equation

$$\hat{N} = \frac{\dfrac{\overline{n}_{Field}}{\overline{p}_{Field}} \cdot E(S)_{Field} + \dfrac{\overline{n}_{Forest}}{\overline{p}_{Forest}} \cdot E(S)_{Forest}}{\hat{p}_{RSF}}.$$

RESULTS

Multiple Covariate Distance Sampling

Models including both observer and habitat as covariates performed worse (based on AIC) than models including only habitat as a covariate for all surveys, for both 250 m and 80 m truncation distances. The best models for a right-truncation distance of 250 m included a half-normal curve for all surveys except for the August 2009 survey, where the hazard-rate curve provided a better fit (Table 7). Estimates of abundance ranged from 253 (95% CI = 167 – 385) to 444 (95% CI = 304 – 650; Table 7). The best models for a right-truncation distance of 80 m included a half-normal curve for the April 2009 and 2010, January 2010, and November 2010 surveys and a hazard-rate curve for the August 2009 and 2010 and the November 2009 surveys (Table 8). Estimates of abundance from best models ranged from 218 (95% CI = 142 – 336) to 381 (95% CI = 238 – 607; Table 8). Model convergence issues occurred using the 80 m truncation distance, likely because of fewer observations of deer.

Table 7. Estimates of density ($\hat{D}$) and abundance ($\hat{N}$) of white-tailed deer with measures of precision from each distance sampling survey, using habitat type (field or forest) of each observation as a covariate and right truncating observations beyond 250 m, Gettysburg, Pennsylvania, 2009-2010.

Survey	Model[a]	k[b]	n[c]	AIC	Δ AIC	$\hat{D}$ (deer/km^2)	$E(S)$[d]	$\hat{P}$[e]	$\hat{N}$	95% CI	CV
Apr 2009	hn	2	106	1,113.82	0.00	11.1	3.41	0.49	324	232 – 453	0.17
	hz	3	106	1,117.02	3.20	9.8	3.41	0.55	285	205 – 396	0.17
Aug 2009	hz	3	143	1,515.74	0.00	8.7	2.00†	0.50	253	167 – 385	0.21
	hn	2	143	1,519.29	3.55	8.5	2.00†	0.51	248	164 – 375	0.21
Nov 2009	hn	2	143	1,456.27	0.00	10.6	1.71†	0.38	310	204 – 472	0.21
	hz	3	143	1,474.95	18.77	8.4	1.71†	0.48	244	162 – 369	0.21
Jan 2010	hn	2	99	1,066.71	0.00	9.6	3.35	0.60	280	189 – 414	0.20
	hz	3	99	1,071.13	4.55	8.1	3.35	0.72	235	161 – 343	0.19
Apr 2010	hn	2	90	924.28	0.00	15.3	3.52	0.38	444	304 – 650	0.19
	hz	3	90	925.86	1.58	13.1	3.52	0.44	381	262 – 553	0.19
Aug 2010	hn	2	167	1,758.84	0.00	10.7	2.33†	0.55	312	217 – 448	0.18
	hz	3	167	1,761.10	2.26	12.4	2.33†	0.48	361	250 – 521	0.19
Nov 2010	hn	2	184	1,893.38	0.00	13.1	1.90†	0.40	382	275 – 531	0.17
	hz	3	184	1,894.30	0.91	11.2	1.90†	0.47	326	236 – 452	0.16

[a] hn = half-normal and hz = hazard-rate.
[b] k = no. model parameters.
[c] n = no. of observed clusters.
[d] E(S) = expected cluster size (mean cluster size or †size-biased regressed cluster size).
[e] $\hat{P}$ = detection probability within 250 m of transects.

Table 8. Estimates of density ($\hat{D}$) and abundance ($\hat{N}$) of white-tailed deer with measures of precision from each distance sampling survey, using habitat type (field or forest) of each observation as a covariate and right truncating observations beyond 80 m, Gettysburg, Pennsylvania, 2009-2010.

Survey	Model[a]	k[b]	n[c]	AIC	Δ AIC	$\hat{D}$ (deer/km^2)	$E(S)$[d]	$\hat{P}$[e]	$\hat{N}$	95% CI	CV
Apr 2009	hn*	2	59	521.07	0.00	9.2	3.32	1.00	268	170 – 423	0.23
	hz*	3	59	522.11	1.05	9.3	3.32	0.99	270	171 – 426	0.23
Aug 2009	hz*	3	83	6.00	0.00	9.9	2.28	0.90	289	200 – 418	0.19
	hn*	2	83	725.05	725.05	10.4	2.28	0.86	303	209 – 437	0.19
Nov 2009	hz	3	98	861.70	0.00	10.9	2.08	0.95	318	215 – 469	0.19
	hn	2	98	861.87	0.17	11.3	2.08	0.92	328	222 – 485	0.19
Jan 2010	hn*	2	44	389.48	0.00	7.5	3.00	0.96	218	142 – 336	0.22
	hz*	3	44	391.62	2.14	7.2	3.00	1.00	210	137 – 323	0.21
Apr 2010	hn*	2	56	490.36	0.00	13.1	3.34	0.81	381	238 – 607	0.24
	hz	3	56	493.28	2.92	11.7	3.34	0.90	342	217 – 540	0.23
Aug 2010	hz	3	94	826.15	0.00	12.2	2.65	0.97	354	256 – 491	0.16
	hn*	2	94	826.55	0.40	12.5	2.65	0.95	363	261 – 504	0.16
Nov 2010	hn*	2	123	1,081.22	0.00	12.6	2.03	0.94	366	255 – 525	0.18
	hz	3	123	1,082.00	0.78	12.4	2.03	0.96	361	252 – 517	0.18

[a] hn = half-normal and hz = hazard-rate.
*denotes model convergence issues, such that results may not be reliable.
[b] k = no. model parameters.
[c] n = no. of observed clusters.
[d] E(S) = expected cluster size (mean cluster size or †size-biased regressed cluster size).
[e] $\hat{P}$ = detection probability within 80 m of transects.

Resource Selection Model

The number of GPS-collared deer used in the RSF ZINB model analysis ranged from 17 to 29 per survey, the total number of GPS locations taken per survey from those deer ranged from 4,542 to 15,472 and sample sizes were similar among deer (Table 9). Model 6, which included a quadratic term for the covariates percent forest, and distance to nearest road, forest-field edge, and NPS owned land boundary, was the parsimonious model for predicting resource selection for all deer during all 7 surveys (Table 10). Deer selected forested areas during all surveys, with the exception of the August surveys, which was the only time standing corn coincided with surveys (Appendix E). For all surveys, deer selected forest-field edges, avoided areas close to roads but selected areas at intermediate distances from roads (approximately 100 m to 150 m), and avoided areas in the park interior and near the park boundary (Appendix E). Additionally, for ease of interpretation, I plotted RSF parameter estimates (Appendix F) and developed fine-scale maps of relative use on the study based on RSF parameter estimates for each survey (Fig. 9, Appendix D).

Table 9. Sample sizes of white-tailed deer collared with global positioning system (GPS) devices and the no. of GPS locations used with the zero-inflated negative binomial model for each survey, Gettysburg, Pennsylvania, 2009-2010.

Survey	n_1 [a]	n_2 [b]	$\bar{n}_2$ [c]	min - max[d]
Apr 9-16, 2009	29	15,472	534	395-556
Aug 3-9, 2009	23	7,316	318	202-336
Nov 20-25, 2009	17	5,586	329	317-359[e]
Jan 5-8, 2010	17	3,688	217	200-232
Apr 1-4, 2010	26	7,263	279	274-300[f]
Aug 25-30, 2010	22	3,923	178	171-180
Nov 15-19, 2010	17	4,542	267	244-274

[a] n_1 = no. of collared deer used in analysis.
[b] n_2 = total no. of GPS locations recorded from n_1 deer.
[c] $\bar{n}_2$ = mean no. of GPS locations taken per deer.
[d] min - max = minimum to maximum no. of GPS locations taken per deer.
[e] One collar only obtained 60 locations, thus was not used as the minimum because it was not representative of the majority of collars.
[f] One collar only obtained 55 locations, thus was not used as the minimum because it was not representative of the majority of collars.

Table 10. Model selection results in the form of delta Akaike's Information Criterion (AIC) values for all 8 models, calculated using summed AIC values for each model across all white-tailed deer collared with global positioning system devices, for each survey, Gettysburg, Pennsylvania, 2009-2010.

Survey	1	2	3	4	5	6	7	8
Apr 2009	745.54	628.88	330.00	273.53	249.76	0	205.37	181.36
Aug 2009	991.07	918.72	414.76	349.60	317.45	0	391.06	267.95
Nov 2009	715.36	630.24	400.29	348.67	286.34	0	350.11	227.07
Jan 2010	479.12	440.87	163.59	170.51	145.90	0	113.57	35.14
Apr 2010	1,265.72	1,131.60	460.76	355.15	321.28	0	303.30	268.24
Aug 2010	1,078.18	949.27	361.62	292.93	305.47	0	273.95	199.18
Nov 2010	847.12	679.03	267.63	148.93	145.62	0	152.69	113.48

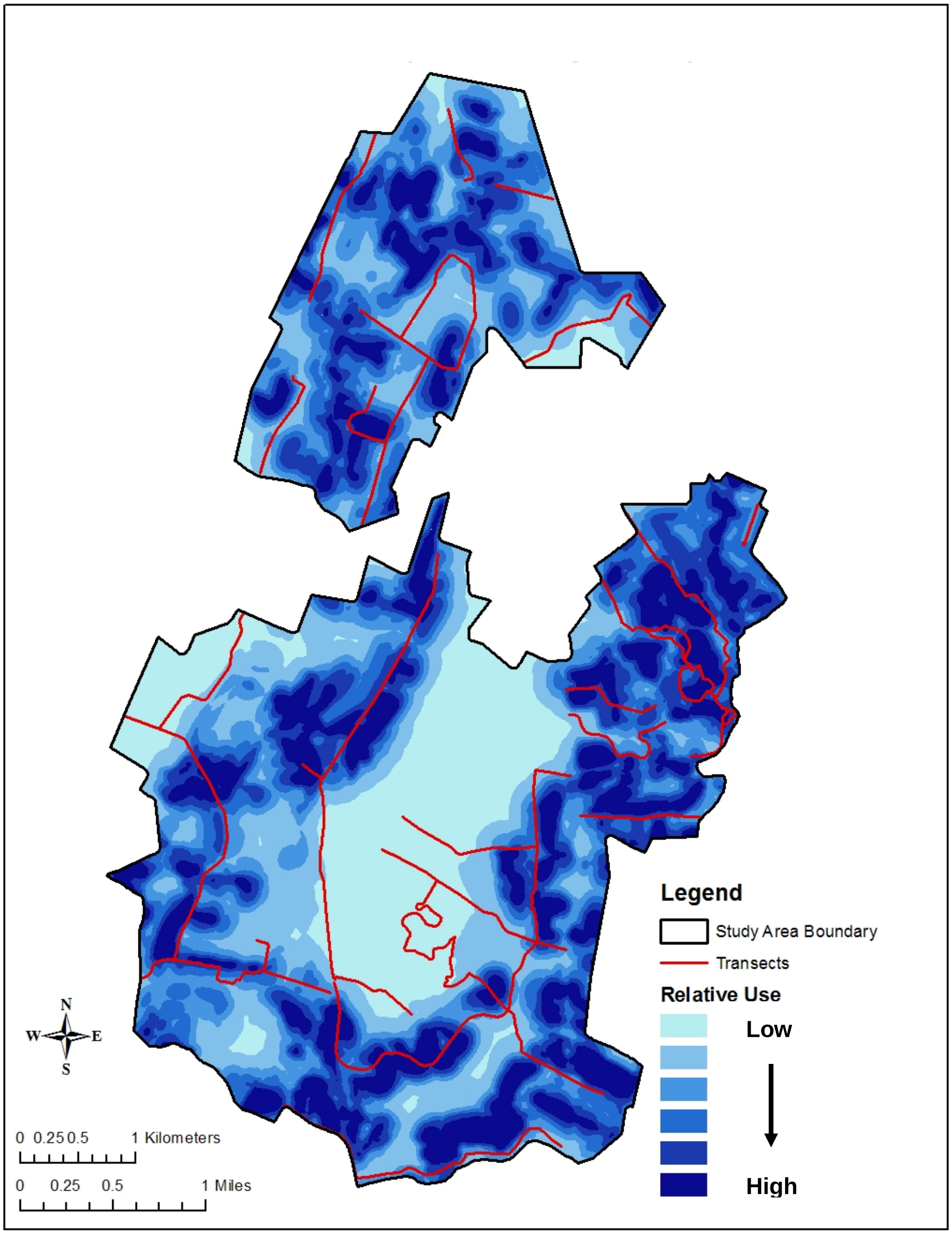


Figure 9. Map of relative use of white-tailed deer (5 x 5 m grid) in the study area during the April 2009 distance sampling survey, Gettysburg, Pennsylvania. (See Appendix D for maps from additional surveys)

Test Assumption 1: Are Deer Uniformly Distributed with Respect to Roads?

The number of GPS-collared deer used in the analysis ranged from 16 to 22 for open areas within 80 m of transects, 17 to 29 for open areas within 250 m of transects, 14 to 27 for forested areas within 80 m of transects, and 16 to 31 for forested areas within 250 m of transects (Table 11). For both open and forested areas, fewer deer used areas within 80 m of transects than within 250 m of transects (Table 11). Further, the uniform curve with no adjustment terms, performed worse (based on AIC) than models including adjustment terms in modeling the distribution of GPS locations for both locations in open areas and forested areas relative to transects during all surveys for both 80 m (Table 12) and 250 m (Table 13) right-truncation distances (Fig. 10, see Appendix A).

Table 11. Sample sizes used for the distribution of global positioning system (GPS) locations of GPS-collared white-tailed deer relative to perpendicular distance to each transect during each survey; where n_1= no. of GPS-collared deer and n_2 = no. of GPS locations, separated by forested and non-forested areas and for 80 m and 250 m from transects, Gettysburg, Pennsylvania, 2009-2010.

	Open Habitat				Forested Habitat			
	80 m		250 m		80 m		250 m	
Survey	n_1	n_2	n_1	n_2	n_1	n_2	n_1	n_2
Apr 9-16, 2009	22	625	29	3,445	27	3,762	31	9,104
Aug 3-9, 2009	20	1,380	25	4,991	22	1,168	25	2,736
Nov 20-25, 2009	16	432	19	2,393	15	632	17	2,406
Jan 5-8, 2010	16	384	18	1,548	14	727	18	1,571
Apr 1-4, 2010	26	756	28	3,413	23	1,027	28	3,171
Aug 25-30, 2010	19	1,491	25	7,691	20	671	24	3,486
Nov 15-19, 2010	16	560	17	1,496	14	987	16	2,120

Table 12. Model selection results from modeling a uniform curve with 0 parameters versus a best-fit curve with up to 3 parameters to binned data of distances from each global positioning system (GPS) location to each transect during each survey, Gettysburg, Pennsylvania, 2009-2010. The GPS locations were collected in open and forested habitats from GPS-collared white-tailed deer and right-truncated beyond 80 m from transects.

	Open Habitat				Forested Habitat			
Survey	Model[a]	k[b]	AIC[c]	ΔAIC[d]	Model[a]	k[b]	AIC[c]	ΔAIC[d]
Apr 2009	un + cos	3	5,363	0	hn + cos	2	32,390	0
	un	0	5,478	114	un	0	32,970	580
Aug 2009	un + cos	3	11,857	0	un + cos	3	10,015	0
	un	0	12,094	238	un	0	10,236	222
Nov 2009	un + cos	2	3,700	0	un + cos	3	5,361	0
	un	0	3,786	86	un	0	5,539	178
Jan 2010	hn + cos	3	3,321	0	un + cos	2	6,342	0
	un	0	3,365	45	un	0	6,372	30
Apr 2010	un + cos	3	6,313	0	un + cos	3	8,929	0
	un	0	6,626	312	un	0	9,001	72
Aug 2010	un + cos	3	12,496	0	un + cos	3	5,641	0
	un	0	13,067	571	un	0	5,881	240
Nov 2010	un + hm	2	4,898	0	un + cos	3	8,545	0
	un	0	4,908	9	un	0	8,650	105

[a] un = uniform, hn = half-normal, cos = cosine, and hm = hermite polynomial.
[b] k = no. of model parameters.
[c] Akaike's Information Criterion (AIC).
[d] Difference in AIC values between each set of models.

Table 13. Model selection results from modeling a uniform curve with 0 parameters versus a best-fit curve with up to 3 parameters to binned data of distances from each global positioning system (GPS) location to each transect during each survey, Gettysburg, Pennsylvania, 2009-2010. The GPS locations were collected in open and forested habitats from GPS-collared white-tailed deer and right-truncated beyond 250 m from transects.

	Open Habitat				Forested Habitat			
Survey	Model[a]	k[b]	AIC[c]	ΔAIC[d]	Model[a]	k[b]	AIC[c]	ΔAIC[d]
Apr 2009	un + cos	3	37,635	0	un + cos	2	99,252	0
	un	0	38,043	408	un	0	100,535	1,283
Aug 2009	hn + cos	3	54,638	0	hn + cos	3	29,982	0
	un	0	55,115	477	un	0	30,213	231
Nov 2009	un + cos	2	26,087	0	un + cos	3	26,439	0
	un	0	26,426	339	un	0	26,569	130
Jan 2010	un + cos	2	17,071	0	hn + cos	3	17,235	0
	un	0	17,094	23	un	0	17,348	114
Apr 2010	un + cos	3	37,068	0	un + cos	3	34,446	0
	un	0	37,689	622	un	0	35,017	571
Aug 2010	un + cos	3	83,872	0	un + cos	3	37,999	0
	un	0	84,931	1,059	un	0	38,496	497
Nov 2010	un + hm	3	16,102	0	un + hm	2	23,229	0
	un	0	16,222	120	un	0	23,411	182

[a] un = uniform, hn = half-normal, cos = cosine, and hm = hermite polynomial.
[b] k = no. of model parameters.
[c] Akaike's Information Criterion (AIC).
[d] Difference in AIC values between each set of models.

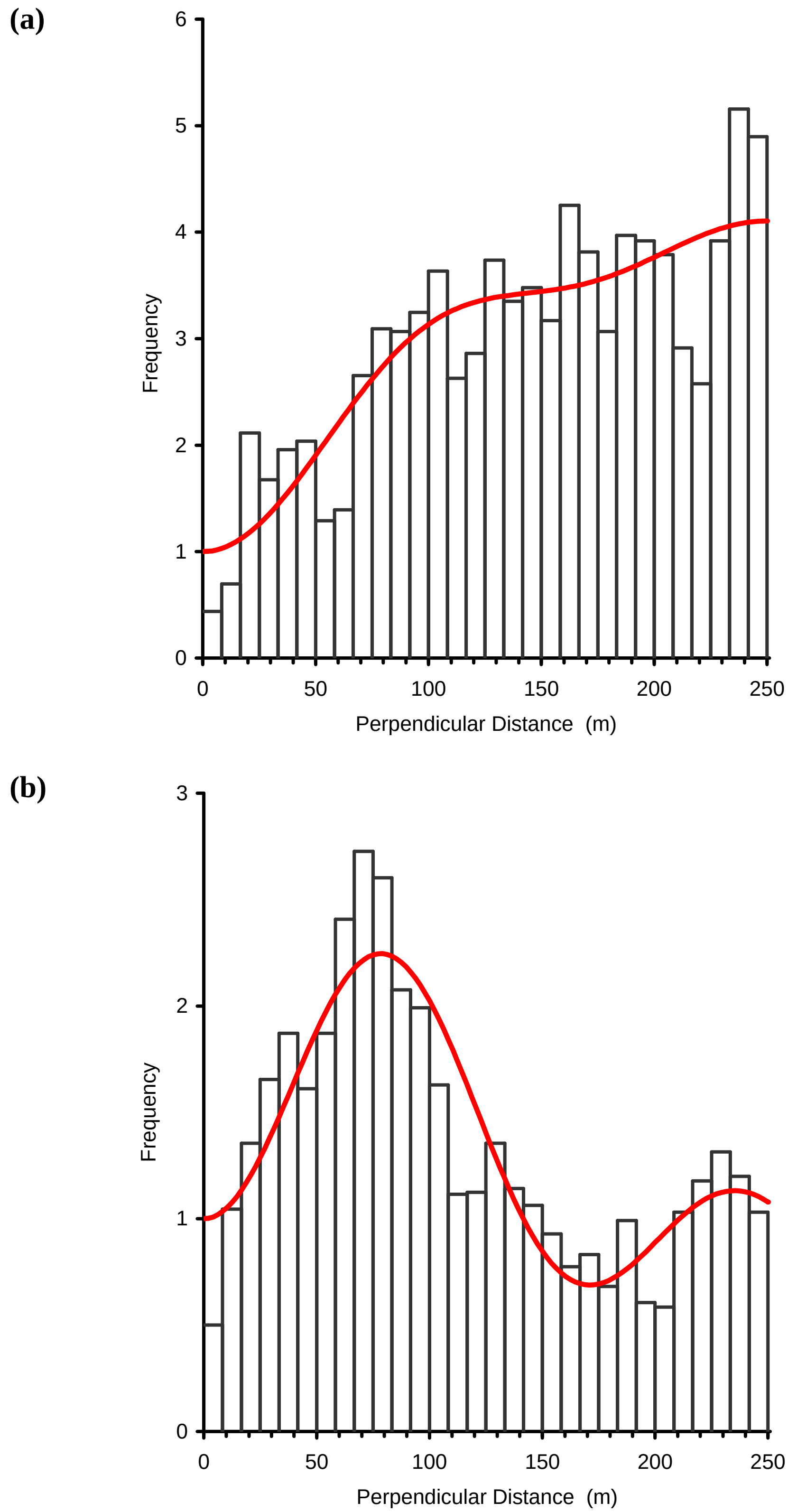


Figure 10. Binned data for the distribution of global positioning system (GPS) locations from GPS-collared white-tailed deer in (a) open areas and (b) forested areas relative to perpendicular distance from transects, with associated best-fit line, during the first distance sampling survey, April 9-16, 2009, performed in the study area at Gettysburg, Pennsylvania (See Appendix A for figures during additional surveys).

Test Assumption 2: Do Roads Provide a Representative Sample of the Study Area?

I sampled slightly more forested areas from the roads I chose as transects relative to the proportion of forested areas in the study area; 4% more forested areas within 250 m of transects and 6% more forested areas within 80 m of transects (Table 14).

Resource use within 250 m of transects ranged from 1% to 16% greater than expected (0.58 = the proportion of the total study area surveyed), ranging from 0.585 (95% CI = 0.099 – 0.626) to 0.672 (95% CI = 0.472 – 0.739), such that the average $\hat{p}_{RSF}$ was 0.647 (SD = 0.029) and the average percent difference was 12% (SD = 0.051; Table 15). Because bootstrapped 95% confidence intervals of $\hat{p}_{RSF}$ were wide, I failed to reject the null hypothesis that transects were representative for all surveys except the August 2009 survey (Table 15).

Resource use within 80 m of transects showed much variation across years and seasons, ranging from -28% to 27% more than expected (0.224 = the proportion of the total study area surveyed), such that $\hat{p}_{RSF}$ ranged from 0.1613 (95% CI = 0.014 – 0.291) to 0.2841 (95% CI = 0.113 – 0.355; Table 16). However, the average $\hat{p}_{RSF}$ across seasons and years was 0.220 (SD= 0.037), such that the average percent difference was only -2% (SD=0.16; Table 16). Bootstrapped 95% confidence intervals of $\hat{p}_{RSF}$ contained the value of 0.224 for every survey. Therefore, I failed to reject the null hypothesis that transects were representative for all surveys (Table 16).

Table 14. Land area and forested area (in km^2) quantified in 2008 for the study area and the 250 m and 80 m survey zones, the proportion of the study area that each survey zone constituted, and the percent forested land in the study area and in each survey zone, Gettysburg, Pennsylvania.

Zone	Area (km^2)	Prop. of Study Area	Forested Area (km^2)	% Forested
Study Area	29.13	1.000	8.48	29
250m Zone	16.89	0.580	5.61	33
80m Zone	6.53	0.224	2.28	35

Table 15. Estimates of the proportion of the study area population of white-tailed deer ($\hat{p}_{RSF}$) within the 250 m survey zone during each survey, Gettysburg, Pennsylvania, 2009-2010.

Survey	$\hat{p}_{RSF}$	SE($\hat{p}_{RSF}$)	95% CI	% Difference[a]
Apr 9-16, 2009	0.672	0.068	0.472 – 0.739	16
Aug 3-9, 2009	0.667	0.026	0.603 – 0.701	15[*]
Nov 20-25, 2009	0.652	0.107	0.247 – 0.727	12
Jan 5-8, 2010	0.647	0.053	0.494 – 0.684	12
Apr 1-4, 2010	0.645	0.121	0.204 – 0.719	11
Aug 25-30, 2010	0.585	0.133	0.099 – 0.626	1
Nov 15-19, 2010	0.663	0.053	0.494 – 0.684	14

[a] Difference between $\hat{p}_{RSF}$ and p = 0.580, the proportion of the total study area that the 250 m survey zone comprised, such that a positive value denotes selection toward areas less than 250 m from transects.
[*]Denotes significant difference.

Table 16. Estimates of the proportion of the study area population of white-tailed deer ($\hat{p}_{RSF}$) within the 80 m survey zone during each survey, Gettysburg, Pennsylvania, 2009-2010.

Survey	$\hat{p}_{RSF}$	SE($\hat{p}_{RSF}$)	95% CI	% Difference[a]
Apr 9-16, 2009	0.223	0.072	0.097 – 0.367	0
Aug 3-9, 2009	0.226	0.068	0.108 – 0.364	1
Nov 20-25, 2009	0.216	0.096	0.050 – 0.403	-4
Jan 5-8, 2010	0.232	0.065	0.113 – 0.355	4
Apr 1-4, 2010	0.200	0.081	0.043 – 0.353	-11
Aug 25-30, 2010	0.161	0.076	0.014 – 0.291	-28
Nov 15-19, 2010	0.284	0.065	0.113 – 0.355	27

[a] Difference between $\hat{p}_{RSF}$ and p = 0.224, the proportion of the total study area that the 80 m survey zone comprised, such that a positive value denotes selection toward areas less than 80 m from transects.

Correction Factor: Adjust for Non-Representative Sample of Habitat

Estimated detection probabilities using CDS for areas within 250 m of transects ranged from 0.56 to 0.80 for open areas and 0.24 to 0.35 for forested areas (Table 17). Contrary to what I had expected because I sampled more forested areas from transects, my correction to the detection probability for the proportion of open areas versus forested areas within 250 m of transects increased estimates of detection probability in fields and forests by an average of 5% and 8%, respectively (Table 17).

Estimated detection probabilities using CDS equaled 1.0 for open areas and ranged from 0.60 to 1.0 for forested areas within 80 m of transects (Table 18). My correction to the detection probability for the proportion of open areas versus forested areas within 80 m of transects had little effect on estimated detection probability because it only increased detection probability in forests by an average of 1% (Table 18).

Table 17. Detection probabilities of white-tailed deer for fields and forests in the 250 m survey zone for each survey month (pooled across years), Gettysburg, Pennsylvania, 2009-2010.

	CDS[a]				Corrected[b]			
Month	p_{Field}	$SE(p_{Field})$	p_{Forest}	$SE(p_{Forest})$	$\overline{p}_{Field}$	$SE(\overline{p}_{Field})$	$\overline{p}_{Forest}$	$SE(\overline{p}_{Forest})$
Apr	0.69	0.06	0.34	0.02	0.73	0.05	0.42	0.02
Aug	0.59	0.03	0.24	0.04	0.66	0.04	0.31	0.05
Nov	0.56	0.04	0.30	0.02	0.63	0.04	0.38	0.02
Jan	0.80	0.09	0.35	0.05	0.83	0.08	0.44	0.05

[a] Detection probabilities for field and forest and associated standard errors were calculated separately using conventional distance sampling (CDS) methods in program DISTANCE.

[b] Corrected average detection probabilities for field and forest were calculated by applying the detection function across the 5 m grid over the landscape based on whether each grid cell was forested or open and its distance to the nearest transect. Standard errors for corrected detection probabilities were calculated using non-parametric bootstrapping.

Table 18. Detection probabilities of white-tailed deer for fields and forests in the 80 m survey zone for each survey month (pooled across years), Gettysburg, Pennsylvania, 2009-2010.

	CDS[a]				Corrected[b]			
Month	p_{Field}	$SE(p_{Field})$	p_{Forest}	$SE(p_{Forest})$	$\overline{p}_{Field}$	$SE(\overline{p}_{Field})$	$\overline{p}_{Forest}$	$SE(\overline{p}_{Forest})$
Apr	1.00	0.00	0.82	0.10	1.00	0.00	0.82	0.08
Aug	1.00	0.00	0.60	0.09	1.00	0.00	0.61	0.08
Nov	1.00	0.00	0.89	0.06	1.00	0.00	0.90	0.17
Jan	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00

[a] Detection probabilities for field and forest and associated standard errors were calculated separately using conventional distance sampling (CDS) methods in program DISTANCE.

[b] Corrected average detection probabilities for field and forest were calculated by applying the detection function across the 5 m grid over the landscape based on whether each grid cell was forested or open and its distance to the nearest transect. Standard errors for corrected detection probabilities were calculated using non-parametric bootstrapping.

Correction Factor: Biases from Non-Random Transects

I used the RSF to adjust for bias related to a non-representative sample when using a 250 m survey zone; however, I could not correct for bias in detection probability related to the non-uniform distribution of deer relative to transects. Therefore, “corrected” estimates using the 250 m survey zone only accounted for some of the bias. These adjusted estimates of abundance ranged from 255 (95% CI = 207 – 477) to 461 (95% CI = 389 – 1,332; Table 19) and differed from MCDS estimates (not corrected) between -74% and 32%, with an average percent difference of -5% (SD=0.33; Table 21). Precision for the corrected estimates was greater than that of MCDS estimates for all surveys except the August 2010 survey (Fig. 11).

I minimized bias in the detection probability related to the non-uniform distribution of deer relative to transects by using a very narrow width (80 m survey zone), where overall detection probability was close to 1.0. Then, to obtain a more representative estimate of abundance in the study area, I adjusted the estimate of abundance in the 80 m survey zone by the proportion of deer present in the 80 m zone based on the RSF. Corrected estimates of abundance for the entire study area ranged from 222 (95% CI = 137 – 420) to 547 (95% CI = 294 – 2,943; Table 20). Corrected estimates differed from MCDS estimates between -19% and 35%, with an average percent difference of 5% (SD=0.19; Table 21, Fig. 12). However, precision for corrected estimates was worse than that of MCDS estimates for every survey (Fig. 12).

Table 19. Estimates of abundance of white-tailed deer for the study area ($\hat{N}_{Corrected}$; corrected for bias from non-random placement of transects, but not for a non-uniform distribution of deer relative to transects) with associated measures of precision and parameters for the 250 m survey zone, Gettysburg, Pennsylvania, 2009-2010.

Survey	$\bar{p}_{Field}$ [a]	$\bar{p}_{Forest}$ [a]	$E(S)_{Field}$ [b]	$E(S)_{Forest}$ [b]	$\bar{n}_{Field}$ [c]	$\bar{n}_{Forest}$ [c]	$\hat{N}_{Survey\ Zone}$ [d]	$\hat{p}_{RSF}$ [e]	$\hat{N}_{Corrected}$	$SE(\hat{N}_{Corrected})$	95% CI
Apr 2009	0.73	0.42	3.73	2.90	21	14	205	0.6723	305	68.95	253 - 404
Aug 2009	0.66	0.31	2.47	1.92	44	4	189	0.6670	283	31.01	239 - 346
Nov 2009	0.63	0.38	2.20	2.10	40	11	203	0.6516	311	1,389.57	255 - 496
Jan 2010	0.83	0.44	3.46	2.95	31	8	186	0.6472	287	81.63	229 - 395
Apr 2010	0.73	0.42	3.89	2.50	22	8	165	0.6446	255	831.16	207 - 477
Aug 2010	0.66	0.31	3.12	1.85	51	4	270	0.5849	461	3,915.83	389 - 1,332
Nov 2010	0.63	0.38	2.27	1.81	42	19	244	0.6632	367	82.22	323 - 465

[a] $\bar{p}$ = the average detection probability in fields and forests.

[b] $E(S)$ = mean cluster size in fields and forests.

[c] $\bar{n}$ = average number of observed clusters in fields and forests per complete round.

[d] $\hat{N}_{Survey\ Zone}$ = the estimate of abundance within the 250 m survey zone.

[e] $\hat{p}_{RSF}$ = the proportion of the study area population within the 250 m survey zone.

Table 20. Estimates of abundance of white-tailed deer for the study area ($\hat{N}_{Corrected}$; corrected for bias from non-random placement of transects and for a non-uniform distribution of deer relative to transects) with associated measures of precision and parameters for the 80 m survey zone, Gettysburg, Pennsylvania, 2009-2010.

Survey	$\overline{p}_{Field}$ [a]	$\overline{p}_{Forest}$ [a]	$E(S)_{Field}$ [b]	$E(S)_{Forest}$ [b]	$\overline{n}_{Field}$ [c]	$\overline{n}_{Forest}$ [c]	$\hat{N}_{Survey\ Zone}$ [d]	$\hat{p}_{RSF}$ [e]	$\hat{N}_{Corrected}$	$SE(\hat{N}_{Corrected})$	95% CI
Apr 2009	1.00	0.82	3.71	3.06	8	12	73	0.2233	327	186.33	193 - 730
Aug 2009	1.00	0.61	2.35	1.82	24	4	67	0.2260	298	115.58	185 - 566
Nov 2009	1.00	0.90	2.07	2.10	24	11	76	0.2157	353	4,470.72	188 - 1,232
Jan 2010	1.00	1.00	2.93	3.20	11	6	52	0.2320	222	102.76	137 - 420
Apr 2010	1.00	0.82	3.97	2.50	11	8	67	0.2004	333	1,981.78	178 - 1040
Aug 2010	1.00	0.61	2.78	1.85	27	4	88	0.1613	547	16,080.75	294 - 2,943
Nov 2010	1.00	0.90	2.22	1.80	23	18	87	0.2841	307	220.81	242 - 797

[a] $\overline{p}$ = the average detection probability in fields and forests.

[b] $E(S)$ = mean cluster size in fields and forests.

[c] $\overline{n}$ = average number of observed clusters in fields and forests per complete round.

[d] $\hat{N}_{Survey\ Zone}$ = the estimate of abundance within the 80 m survey zone.

[e] $\hat{p}_{RSF}$ = the proportion of the study area population within the 80 m survey zone.

Table 21. Percent difference between abundance estimates of white-tailed deer using multiple covariate distance sampling and bias-adjusted estimates of abundance using the correction factor for each survey Gettysburg, Pennsylvania, 2009-2010.

Survey	250 m Zone[a]	80 m Zone[a]
Apr 2009	-6%	18%
Aug 2009	11%	3%
Nov 2009	0%	10%
Jan 2010	3%	2%
Apr 2010	-74%	-15%
Aug 2010	32%	35%
Nov 2010	-4%	-19%

[a] A positive value denotes the correction factor increased the estimate of abundance from the distance sampling estimator when roads were used as transects.

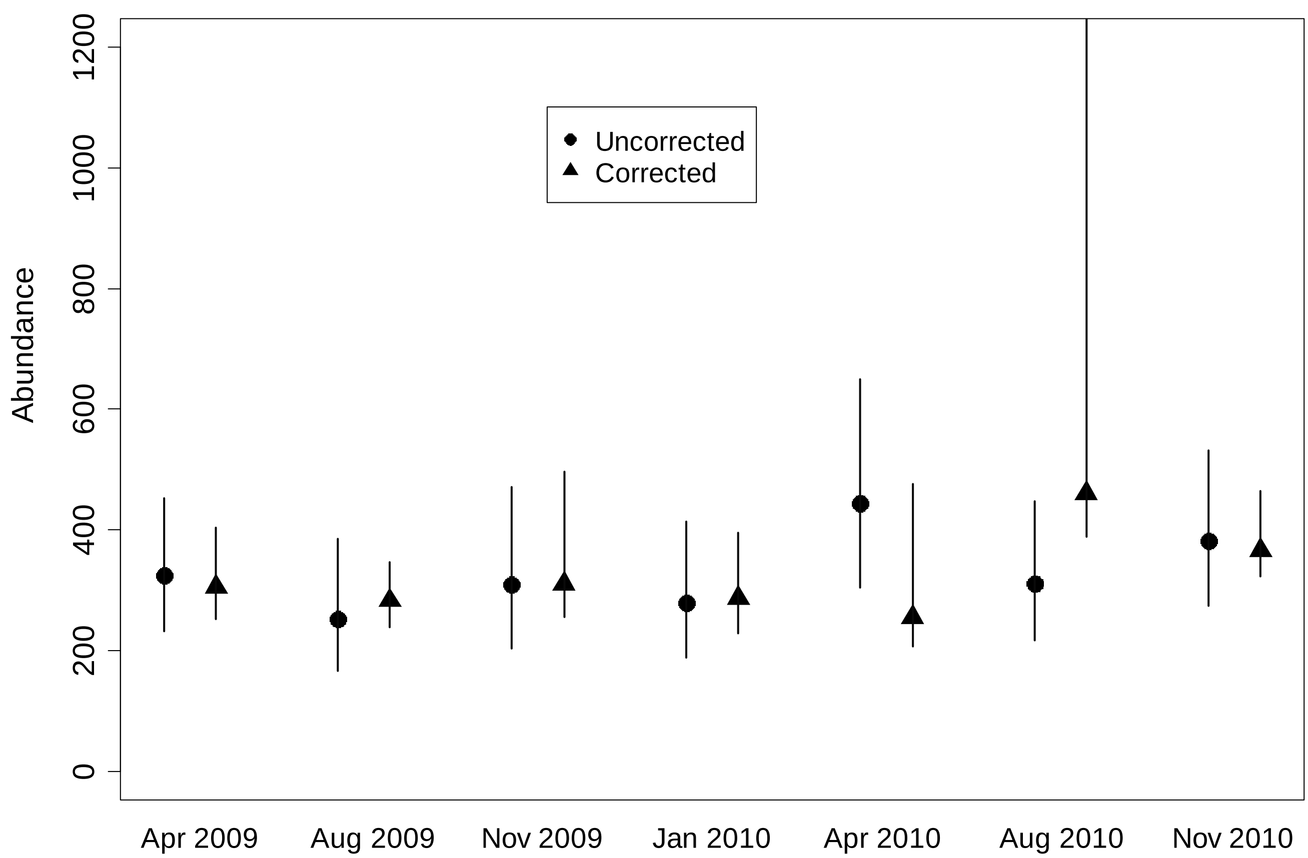


Figure 11. Abundance estimates ($\hat{N}$) of white-tailed deer in the study area and associated 95% confidence interval bars using multiple covariate distance sampling (MCDS: ignoring any violations of assumptions) and bias-adjusted estimates of abundance using the correction factor for each survey using the 250 m survey zone, Gettysburg, Pennsylvania, 2009-2010. The upper value of the 95% confidence interval for the August 2010 survey (1,332) extends beyond the y-axis limit shown.

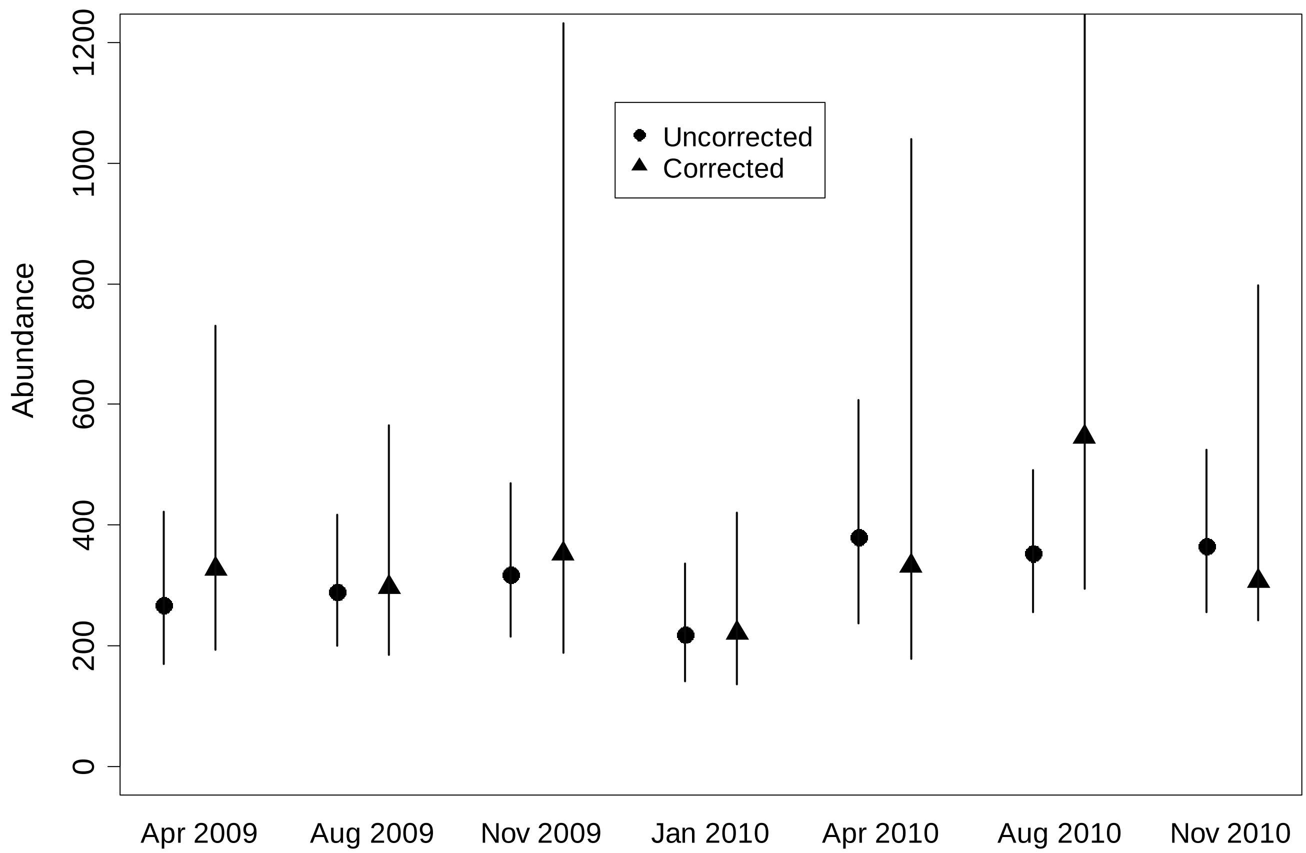


Figure 12. Abundance estimates ($\hat{N}$) of white-tailed deer in the study area and associated 95% confidence interval bars using multiple covariate distance sampling (MCDS: ignoring any violations of assumptions) and bias-adjusted estimates of abundance using the correction factor for each survey using the 80 m survey zone, Gettysburg, Pennsylvania, 2009-2010. The upper value of the 95% confidence interval for the August 2010 survey (2,943) extends beyond the y-axis limit shown.

DISCUSSION

Test Assumptions

My results documented bias in abundance estimates related to using existing, non-randomly placed roads as transects for distance sampling. I used GPS locations from GPS-collared adult and juvenile male and female deer in the study area to test two assumptions that are critical to meet for unbiased estimates, and are typically met using randomly placed transects. The first is that the distribution of deer is uniform with respect to perpendicular distance from transects, which allows the estimator for detection probability to be unbiased (Buckland et al. 2001). The second is that the sample from transects is representative of the entire study area, which allows the estimate of density in the sample to be extrapolated to abundance in the larger area of interest (Buckland et al. 2001). Previous studies that investigated roads as transects failed to investigate both critical assumptions; and conclusions tended to be speculative, based only on observation data, or relied on VHF telemetry locations, where the error associated with locations prevented fine scale measurement of avoidance of roads (e.g., Heydon et al. 2000, Ruette et al. 2003, Ward et al. 2004, Butler et al. 2005, Venturato et al. 2010, Erxleben et al. 2011).

I found that the distribution of GPS-collared deer was not uniform with respect to the roads I used as transects in both forested and non-forested areas (Table 12, Table 13), which violated the first assumption. GPS-collared deer tended to avoid areas within approximately 50 m of transects and the avoidance distance was larger in open areas than in forested areas (see Fig. 10, Appendix A). Additionally, I observed fewer unmarked deer near roads during surveys. Several studies using roads as transects with distance sampling also observed fewer detections near transects than expected for deer

(*Odocoileus hemionus*; e.g., Rost and Bailey 1979, Kie and Boroski 1995; *Cervus nippon*; e.g., Koganezawa and Li 2002, *Capreolus capreolus*; e.g., Ward et al. 2004), moose (*Alces alces*; e.g., Yost and Wright 2001), and foxes (*Vulpes vulpes*; e.g., Heydon et al. 2000, Ruette et al. 2003), but were unable to definitively test why. Fewer detections near the road could be caused by a number of reasons, including avoidance of the areas near roads (e.g., because of disturbance or correlation of habitat with roads; Fewser et al. 2008), movement away from roads in response to observers, or missed observations near roads (Buckland et al. 2001). Considering I also observed fewer GPS locations near transects during surveys, I believe the likely cause was avoidance of roads. I do not rule out movement of deer in response to our vehicle or the spotlight before detection. However, upon inspection of intensive GPS data, I rarely observed flight responses of GPS-collared individuals with respect to observers before detection, and observers were trained to always look ahead to ensure all observations on the transect were detected and that observations were recorded at their initial location. Thus, because deer avoided areas close to roads, using the distance sampling estimator with roads as transects would lead to positively biased estimates of detection probability (e.g., similar to Fig. 7), which would lead to negatively biased estimates of density.

However, I found that the roads I chose as transects provided a representative sample of forested areas on the study area. Additionally, I failed to reject the null hypothesis that habitat use (or the proportion of deer present) was in proportion to availability for both the 80 m and 250 m survey zones during all surveys, except the August 2009 survey. However, the statistical power to reject the null hypothesis was poor because 95% confidence intervals were wide (Table 15, Table 16), likely because of the RSF calculation and bootstrapping process; thus, the null hypothesis test may not have been a meaningful test. From GPS data, I observed a non-linear trend in the number of locations in relation to roads, so there is likely a similar trend in deer density in relation to roads. Therefore, finding that the sample within 80 m from roads was representative of deer density in the study area could be because mean density within that interval was representative. For instance, the interval from 0–40 m may under-represent density, whereas the interval from 40–80 m may over-represent density, such that the average is representative. However, the decrease in precision associated with smaller transect widths precluded testing the representativeness of smaller intervals.

In conclusion, the distribution of deer with respect to the roads used as transects was not uniform, such that estimates of density from the sample were negatively biased. Therefore, regardless of whether the sample from roads was representative of the larger study area, extrapolating density from the sample to the study area would provide negatively biased estimates of density for the study area, unless estimates of density in the sample can be corrected.

Correction Methods

Several studies attempted to correct estimates of density when few detections were observed near transects by employing left-truncation methods or a wide first interval (e.g., Heydon et al. 2000, Koganezawa and Li 2002, Ruette et al. 2003, Ward et al. 2004). There are two types of left-truncation; rescaling (displace the transect line to some distance x and censor truncated observations; Buckland et al. 2001) or full-left truncation (where the detection function is extrapolated back to distance 0 from some distance x; Alldredge and Gates 1985). If the distribution of animals with respect to perpendicular distance to transects is uniform, then in certain circumstances, both left-truncation methods can provide more representative estimates (e.g., when observations on the transect are missed directly under an airplane; Buckland et al. 2001). In the case of roads as transects, if animals are avoiding the areas close to the road, but then the distribution relative to transects becomes uniform after some distance *x* (e.g., similar to Fig. 4), then the rescaling method can provide more representative estimates (Buckland et al. 2001). However, detection probability at distance *x* must be 1.0 and the density of deer in the area less than distance *x* from roads needs to be incorporated into the estimate of overall abundance (Buckland et al. 2001). Additionally, the assumption that the sample from roads is representative of the population would need to be tested and a correction applied if found to be unrepresentative. Full-left truncation and increasing the width of the first interval would not be appropriate to use because there are fewer animals in the area less than distance *x*, which would result in an over-estimate of density near transects (Buckland et al. 2001). Neither method of left-truncation would be appropriate to use given the distributions I observed (Appendix A), because the distributions did not typically become uniform after road avoidance. Fitting a hazard-rate function to data can be useful to assign a detection probability of 1.0 to distances close to the transect. However, because the true distribution is non-uniform, detection probability would remain positively biased, such that when the detection function does begin to decline, actual detection probability may be much lower than estimated (see Fig. 7). These problems highlight the necessity for new methods to correct estimates when animals are distributed non-uniformly relative to transects.

I developed a correction factor to yield more representative estimates of abundance for the study area. First, I reduced potential for bias in the detection probability by using a narrow transect width (80 m) where detection probability was near

1.0 (Table 18). However, because deer avoided areas near roads (Appendix A), the estimate of density in the 80 m survey zone was likely unrepresentative of the study area. Therefore, I adjusted the estimate in the 80 m survey zone by the proportion of the study area population present in the 80 m survey zone, based on the RSF (Table 16, Table 20). However, I believe the estimated detection probabilities from CDS in forested areas were higher than actual (e.g., I doubt that we detected every deer in forests within 80 m of transects during the January 2010 survey; Table 18), such that corrected estimates of abundance were negatively biased. Using an even narrower distance would be more appropriate to reduce bias in detection probability; however, the narrower the width, the less precision for two reasons. First, the number of observations decreases, which increases the variance associated with the distance sampling estimator. Second, because $\hat{p}_{RSF}$ becomes smaller as w decreases, the bootstrapped estimates of $\hat{p}_{RSF}$ lead to less robust estimates of abundance because $\hat{p}_{RSF}$ is used in the denominator.

The method I used to correct for bias in the detection probability was not appropriate for the 250 m survey zone. Therefore, the estimator of detection probability was likely positively biased for both the MCDS and RSF corrected estimators, resulting in negatively biased estimates of abundance. Additionally, the correction to the detection probability for the proportion of forested and open areas inflated the detection probability even more. I believe this was because sections of forest obstructed observer detection in many secluded open areas within 250 m. However, grid cells for these secluded areas were assigned a detection probability based on the distance from transect, when in reality detection probability should have been 0. Given the proportion of forested areas within the sampled area was representative of the study area, using the original CDS detection probabilities would have been more appropriate. Nevertheless, my goal was to demonstrate a method to correct for a non-representative sample of habitat.

Original estimates of abundance using MCDS varied little from RSF corrected estimates of abundance for both the 250 m and 80 m survey zones (Figs. 10 and 11). I believe this was because my correction for detection probability was not sufficient, such that both MCDS and RSF corrected estimates of abundance were negatively biased. Further, because I observed avoidance of areas near roads, I expected the RSF correction factor to increase estimates of abundance; however, this was complicated by increased use of intermediate distances from roads (Table 21, Appendix A). Additionally, because I minimized bias in the detection probability for the 80 m survey zone, I expected the difference between MCDS and RSF corrected estimates of abundance for the 80 m survey zone to be greater than for the 250 m survey zone; however, this was rarely the case (Table 21). I believe the RSF correction did little because transects happened to be near areas of high use (Fig. 9 , Appendix D) and provided a representative sample of habitat important for deer (Tables 14-16). If the roads used as transects happened to be in areas of low resource use or if forested areas along transects were underrepresented, the correction factor would have had a greater effect on $\hat{N}$.

I modeled relative use on the study area using the RSF with GPS locations and landscape covariates important for selection or avoidance by deer. For results to be accurate, the assumption that the marked population was representative of the unmarked population, with respect to habitat use, must be met. The proportion of each age and sex class marked should be representative of the population and the marked population should be distributed on landscape representative to the density of deer on the landscape. I captured all age and sex classes of deer throughout the study area and relative to local densities. Although I did not know the true sex ratio or age structure, I believe my marked population was representative of the unmarked population; thus, I believe I satisfied these assumptions. However, the issue of the NPS culling deer on the park may have resulted in inaccuracy in the RSF model. For instance, we observed fewer deer on NPS owned property compared to private lands, likely because of the culling operations on NPS owned property. It was difficult to model differences in density of deer related to hunting and culling (e.g., suitable habitat unoccupied by deer) with the RSF, but I attempted this using distance to the nearest NPS boundary, with the assumption that the majority of the culling occurs in the central regions of the park (e.g., further distances). A more realistic covariate would have been to designate areas on the landscape as hunted (e.g., areas where the NPS culls deer or private areas where hunting occurs) or not hunted. Thus, I do not believe I adequately modeled the heterogeneity in the density of deer on the landscape with the RSF. For example, we observed few deer on areas of NPS owned property with favorable habitat for deer, where the RSF predicted use would be high. I do not believe this led to considerable bias, but if bias related to the RSF predicting higher use in the park than expected existed, it would have introduced negative bias in the corrected density estimator.

I used GPS locations during the time I conducted surveys (typical accuracy <20 m) to examine the distribution of deer relative to transects. However, using ground-based VHF telemetry data to test assumptions may not provide the accuracy or number of intensive locations required to accurately model the distribution of animals relative to transects during the time that surveys are conducted (e.g., Butler et al. 2005, Venturato et al. 2010, Erxleben et al. 2011). For instance, the lack of accuracy with VHF telemetry locations may prevent detection of avoidance of areas close to roads (e.g., <50 m).

Furthermore, it is paramount that locations are collected during the time that surveys are conducted, especially when temporal and seasonal variation in habitat use exists. For instance, modeling the distribution of VHF locations taken during the day over a period of 5 months may not be appropriate for making inferences to the distribution of animals during surveys conducted in a one-week period or at night. Hounsome et al. (2005) properly addressed these issues with VHF locations, which they used to predict the proportion of badgers (*Meles meles*) using open areas during distance sampling surveys, but they experienced issues related to a low sample size of locations during surveys.

In addition, measuring the distance from each animal's location to the nearest road may not represent the distribution of that animal relative to the distribution of roads or the roads used as transects. For instance, in areas with a high density of roads, there may be more areas close to roads and few areas more than 200 m from a road. Conversely, in areas of low road density, there may be more areas further from roads than near roads. Therefore, it is vital to measure the distance from each location to each transect used, not just to the nearest transect. Venturato et al. (2010) used distances from VHF locations to the nearest road, but compared these to random locations, such that a reasonable comparison could be made; however, the analysis lacked statistical power to detect a difference.

In conclusion, my correction method could have provided more representative estimates of abundance for the study area than MCDS estimates (i.e., ignoring violations of assumptions) when roads were used as transects to sample deer with distance sampling. However, the method did not account for all of the bias and resulted in decreased precision because of the narrow width (e.g., 80 m) required to minimize bias in detection probability.

CHAPTER 4:

CONCLUSIONS AND FUTURE RESEARCH

My results suggested the assumption of homogeneity in detection rates for the L-P estimator was violated. If the sample of marked individuals was truly representative of the unmarked population (i.e., adult and juvenile male and female deer in proportion to the true sex and age ratio of deer on the study area) then the L-P estimator would be accurate. However, the variance may not be representative, especially if individuals exhibited heterogeneity in detection rates. Bowden's estimator relaxes the assumption of homogeneous detection probability; thus, estimates were likely accurate. Although mark-resight methods can provide accurate and precise estimates of abundance, methods are difficult and expensive, especially when implemented at a large scale. Capturing deer is time consuming and must be conducted on an annual or biennial basis to retain a sufficient sample size of marked deer in the population. Additionally, to use detection probability from mark-resight surveys to estimate abundance in future surveys (e.g., as a sighting index), the assumption that detection probability does not change over time must be met, which is highly unlikely because many factors related to detection rates can change over time (Anderson 2001). Therefore, the updated sighting index I calculated may not provide an unbiased estimator of abundance for future surveys.

There is no perfect solution for meeting all assumptions of distance sampling when surveying for highly mobile animals such as deer. Even if random transects are used, it is difficult to detect all animals on the transect from aerial surveys (Fewster et al. 2008) and walking transects often results in avoidance of the observer (e.g., Koenen et al. 2002). Additionally, as discussed in Buckland et al. (2001) and Fewster et al. (2008), the use of non-random roads or tracks as transects for distance sampling can result in considerable bias because the location of roads or the disturbance from roads may affect the distribution of animals. My results concluded that the distribution of deer was correlated to the distribution of the roads I surveyed (Table 12, Table 13). Deer avoided areas near transects and often selected for intermediate distances (Fig. 10, Appendix A). Therefore, I expected detection probability was positively biased, leading to negatively biased estimates of abundance.

I compared distance sampling estimates of abundance to mark-resight estimates of abundance (Fig. 13) and found that distance sampling estimates were always lower. This supported the conclusion that distance sampling from roads provided positively biased detection probabilities. I attempted to correct for bias when roads were used as transects for distance sampling by restricting analysis to a narrow width and then adjusting estimates by the RSF (Fig. 12). However, this method may not have corrected for all of the bias (e.g., point estimates were also negatively biased compared to mark-resight; Fig. 13) and resulted in decreased precision. My results regarding bias in detection probability can be generalized to distance sampling using roads as transects for any animal not distributed independently of roads, especially highly mobile animals, and where roads do not provide a representative sample of the larger area of interest.

The argument can be made that an inaccurate estimator with good precision may be more useful for management and predicting trends than an accurate estimator with poor precision. Additionally, an argument can be made that the logistical advantages of using roads as transects outweigh the disadvantages (Heydon et al. 2000). Nevertheless, any study using roads or tracks as transects with distance sampling should carefully consider the effects on bias. For instance, if roads are used as transects and animals avoid

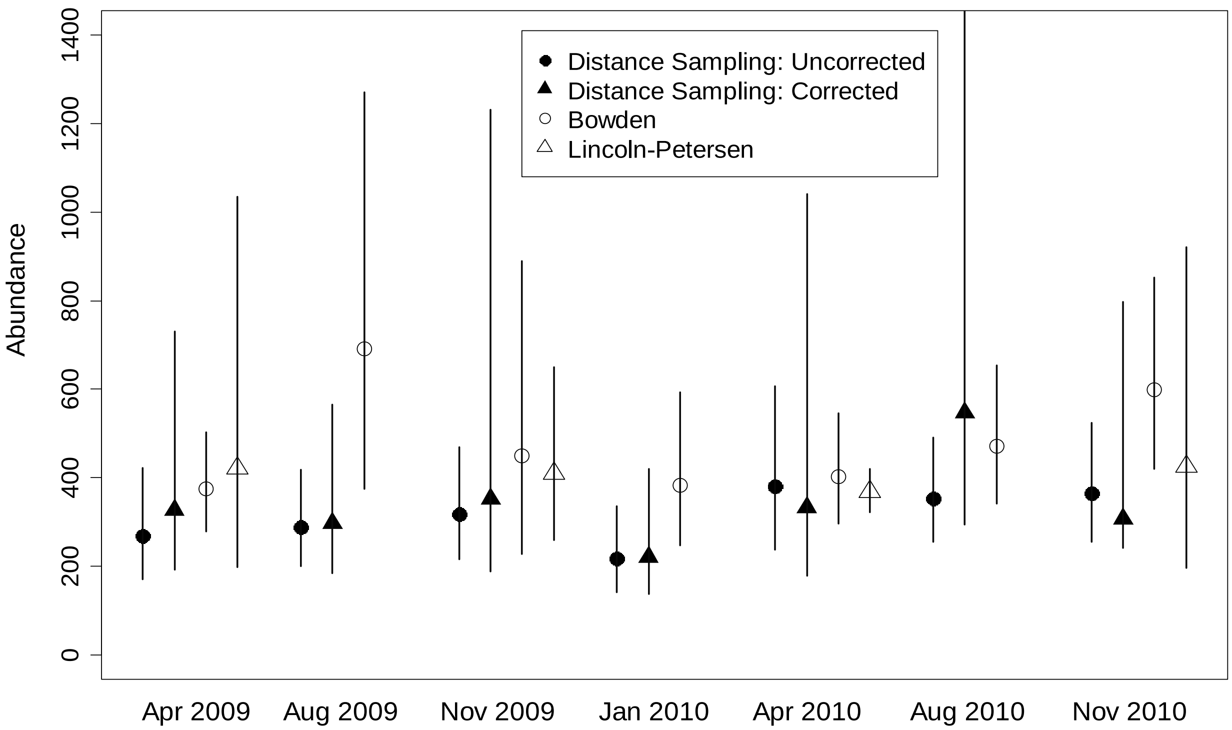


Figure 13. Abundance estimates ($\hat{N}$) of white-tailed deer in the study area and associated 95% confidence interval (CI) bars using distance sampling methods (solid symbols) using the 80 m survey zone and mark-resight methods (hollow symbols) for each survey, Gettysburg, Pennsylvania, 2009-2010. The upper value of the 95% CI for the August 2010 survey (2,943) extends beyond the y-axis limit shown.

the road, abundance estimates should be treated as indices of abundance rather than point estimates. Anderson (2001) highlights some of the problems with using indices of abundance. However, using distance sampling estimates as indices of abundance can reduce some sampling variability because the method can incorporate differences in observer detection rates (see Diefenbach et al. 2003) and model detection probability with additional covariates, such as habitat types (Marques et al. 2007).

If using roads as transects is the only feasible option, the following methods may decrease potential bias. As proposed by Buckland et al. (2001), selecting a large number of short sections and stratifying transects by habitat such that the proportion of total lengths in each habitat are equal to the proportion of their availability in the larger study area can reduce bias. Additionally, using roads or tracks with very low or no traffic volume may reduce bias. Rost and Bailey (1979) observed an increase in avoidance of areas near roads by deer depending on the level of traffic volume. Further, Gill et al. (1997) used a thermal imager along unimproved forest tracks, which were closed to public traffic, to survey deer using distance sampling methods. They did not evaluate the distribution of deer relative to those tracks, but did not observe fewer detections near tracks, nor did they conclude that deer were moving in response to their vehicle. Therefore, use of thermal imagers could reduce movement of animals in response to observers, ensure that all animals on or near the transect are detected, increase the number of observations, and reduce disturbance to the public. However, without investigating the true distribution of animals relative to transects, the magnitude and direction of the bias remain unknown.

I provided estimates of abundance of deer on the study area containing GNMP-ENHS using both mark-resight and distance sampling estimators (Appendix J), which negatively biased or not, were all above the deer density goal of 10 deer/km^2 (25 deer/mi^2) of forested land. However, there are several logistical constraints preventing park managers from reaching this goal. The NPS owns 61% of the land within the study area, but NPS owned forested land comprises only 17% of the study area. Park managers can only cull deer on park owned property and from roads where areas on both sides are owned by the park. Because of public roads and buildings in and around the park, managers can safely shoot deer only on a small proportion of the park, which tends to be areas in the park interior. The small area that managers can safely cull deer may be the reason it has proven difficult to reduce deer density on the entire study area to 10 deer/km^2 of forested land.

Furthermore, although I found that deer density in the study area was above the long-term goal, park managers have observed adequate regeneration of hardwood seedling trees on park woodlands (Niewinski et al. 2006; Randy Krichten, GNMP-ENHS vegetation manager, personal communication). Additionally, I observed few deer in the interior of the park during surveys. Most of the deer tended to be on private lands and on the edge of the park along private lands (i.e., areas where park managers cannot shoot deer). Given that an important objective at GNMP-ENHS is to preserve the historic character of the landscape, the requirement for a precise and often expensive measure of deer abundance for meeting deer culling goals may be unnecessary. Regardless of whether or not the deer density goal of 10 deer/km^2 of forested land in the study area is reached, the objective of preserving the historic character of the landscape still may be attained. The NPS may want to consider re-evaluating deer density goals for GNMP-ENHS if data to assess current vegetation conditions are available (e.g., measures of regeneration, species diversity, and deer browsing).

Future Methods

The method I used to reduce bias in detection probability was to restrict analysis to a width in which detection probability was likely near 1.0. However, this approach likely provided positively biased estimates of detection probability because it was unlikely that most of the deer within 80 m were seen in forested areas. Perhaps a better method to obtain unbiased estimates of detection probability would be to model the detection function with respect to the true distribution of animals. This method is similar to the correction method used by Anderson and Pospahala (1970), but I used an estimated distribution of animals relative to transects (Fig. 14) rather than assuming it was uniform. The method involved fitting a non-linear curve to detections of deer (Fig. 15), then rescaling so the y-intercept was equal to the non-linear curve fit to GPS location data (Fig. 14). Thus, detection probability was the ratio of the integral of the observation curve divided by the integral of the distribution curve (Fig. 16). An important assumption remains that all animals on the transect are observed and if the sample is not representative of the study area, appropriate correction methods (e.g., RSF) must be used. Additionally, a variance estimator for the resulting detection probability is needed.

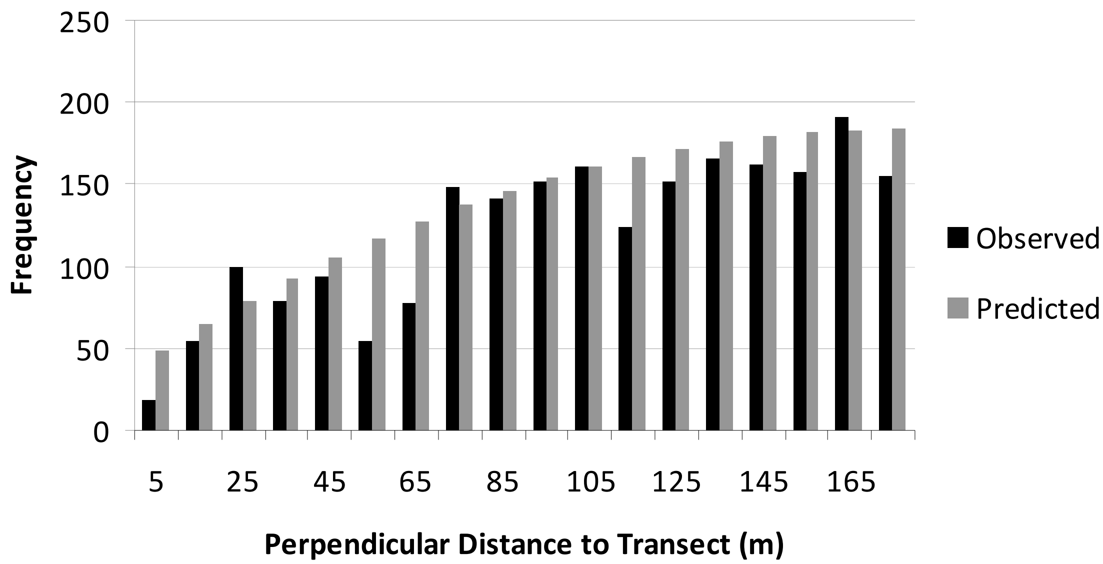


Figure 14. Binned data of the no. of global positioning system (GPS) locations with respect to perpendicular distance to each transect (black) and binned data of the predicted values from a best-fit model (gray), based on data collected from GPS-collared white-tailed deer during distance sampling surveys, Gettysburg, Pennsylvania, 2009-2010.

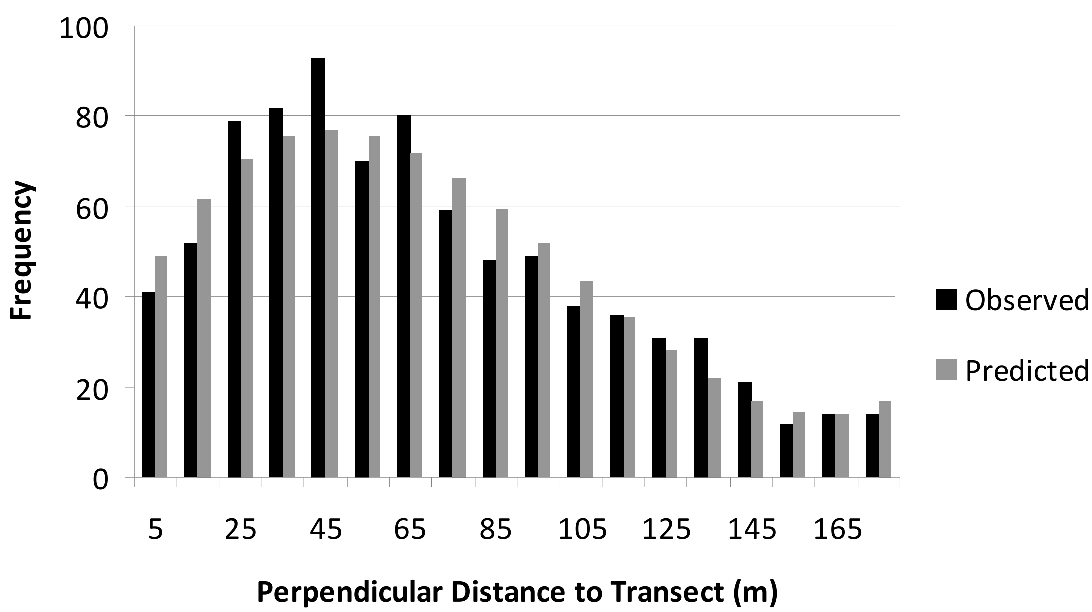


Figure 15. Binned data of the no. of observations from distance sampling surveys with respect to perpendicular distance to each transect (black) and binned data of the predicted values from a best-fit model (gray), based on groups of white-tailed deer observed during distance sampling surveys, Gettysburg, Pennsylvania, 2009-2010.

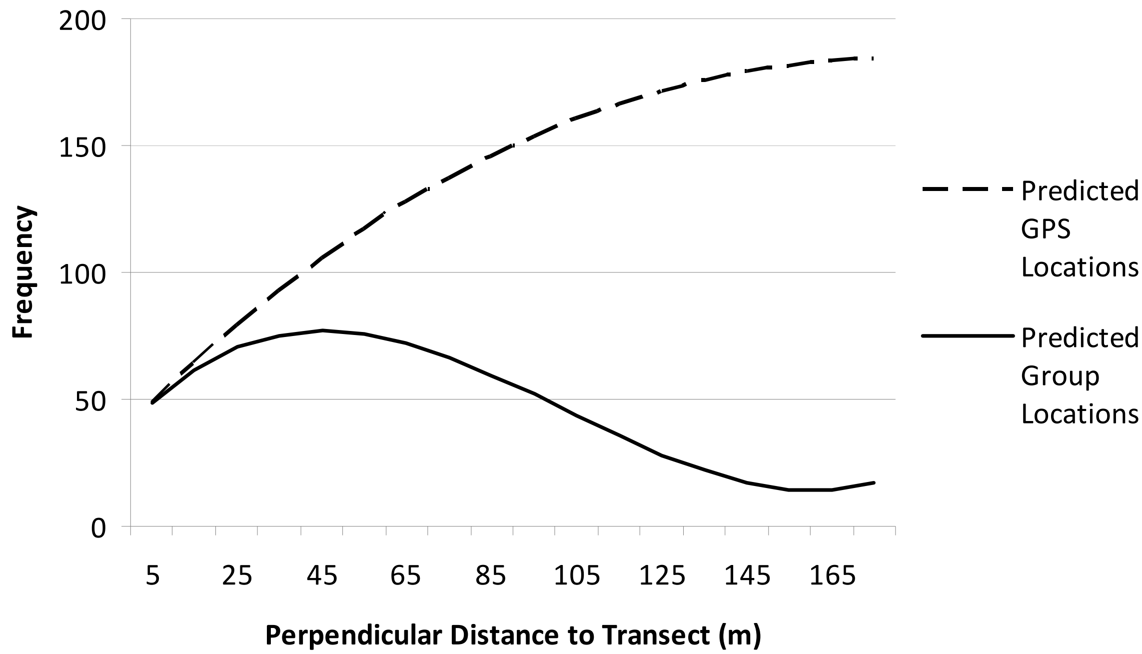


Figure 16. Best-fit curve of global positioning system (GPS) data (dashed line), which exhibits a non-uniform distribution relative to transects, and best-fit curve of distance sampling observations (solid line) rescaled so y-intercepts are the same. The detection probability for a given distance interval is the integral of the observation curve divided by the integral of GPS curve for that interval. The curve for predicted GPS locations is based on data collected from GPS-collared white-tailed deer and the curve for predicted group locations is based on observations of groups of deer during distance sampling surveys, Gettysburg, Pennsylvania, 2009-2010.

Marques (2007), proposed a likelihood-based approach to incorporate a density gradient from transects to provide unbiased estimates of abundance. The method measures the density gradient using secondary transects sampled perpendicular to the primary transects in which the density gradient exists. For my purposes, the GPS distribution curve could serve as the measurement of the density gradient, and replace the requirement for secondary transects. Additionally, Marques et al. (2010) developed a method using random point transects along linear features such as roads and included sighting angles to provide an unbiased estimate of detection probability. However, for both methods, if the sample were unrepresentative of the larger area of interest, then a correction would need to be applied.

LITERATURE CITED

Alldredge, J. R., and C. E. Gates. 1985. Line transect estimators for left-truncated distributions. Biometrics 41:273-280.

Anderson, D. R. 2001. The need to get the basics right in wildlife field studies. Wildlife Society Bulletin 29:1294-1297.

Anderson, D. R., and K. P. Burnham. 2002. Avoiding pitfalls when using information-theoretic methods. Journal of Wildlife Management 66:912-918.

Anderson, D. R., and R. S. Pospahala. 1970. Correction of bias in belt transects of immotile objects. Journal of Wildlife Management 34:141-146.

Bartmann, R. M., G. C. White, L. H. Carpenter, and R. A. Garrott. 1987. Aerial mark-recapture estimates of confined mule deer in pinyon-juniper woodland. Journal of Wildlife Management 51:41-46.

Bates, S. 2006. White-tailed Deer Density Monitoring Protocol Version 1.1. Inventory and Monitoring Program, National Capital Region Network, National Park Service. Washington, D.C., USA.

Beringer, J., L.P. Hansen, W. Wilding, J. Fischer, and S.L. Sheriff. 1996. Factors affecting capture myopathy in white-tailed deer. Journal of Wildlife Management 60:373-380.

Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas. 2001. Introduction to distance sampling: estimating abundance of biological populations. Oxford University Press, Oxford, United Kingdom.

Buckland S. T., I. B. J. Goudie, and D. L. Borchers. 2000. Wildlife population assessment: past developments and future directions. Biometrics 56:1-12.

Burnham K. P., and D. R. Anderson. 1998. Model selection and inference: a practical information-theoretic approach. Springer, New York, New York, USA.

Bowden, D. C., and R. C. Kufeld. 1995. Generalized mark-sight population size estimation applied to Colorado moose. Journal of Wildlife Management 59:840-851.

Caughley, G. 1974. Bias in aerial survey. Journal of Wildlife Management 38:921-933.

Caughley, G. 1977. Sampling in aerial survey. Journal of Wildlife Management 41:605-615.

Chao, A. 1989. Estimating population size for sparse data in capture-recapture experiments. Biometrics 45:427-438.

Clover, M. R. 1956. Single-gate deer trap. California Fish and Game 42:199-201.

Cogan, R. D., and D. R. Diefenbach. 1998. Effect of undercounting and model selection on a sightability-adjustment estimator for elk. Journal of Wildlife Management 62:269-279.

Conner, M. C., R. A. Lancia, and K. H. Pollock. 1986. Precision of the change-in-ratio technique for deer population management. Journal of Wildlife Management 50:125-129.

Diefenbach, D. R. 2009. Estimating avian population size using Bowden's estimator. The Auk 126:211-217.

Diefenbach, D. R., D. W. Brauning, and J. A. Mattice. 2003. Variability in grassland bird counts related to observer differences and species detection rates. The Auk 120:1168-1179.

Downing, R. L., E. D. Michael, and J. P. Robert. 1977. Accuracy of sex and age ratio counts of white-tailed deer. Journal of Wildlife Management 41:709-714.

Erxleben, D. R., M. J. Butler, W. B. Ballard, M. C. Wallace, M. J. Peterson, N. J. Silvy, W. P. Kuvlesky, D. G. Hewitt, S. J. DeMaso, J. B. Hardin, and M. K. Dominguez-Brazil. 2011. Wild turkey (*Meleagris gallopavo*) association to roads: implications for distance sampling. European Journal of Wildlife Resources 57:57-65.

Fewster, R. M., C. Southwell, D. L. Borchers, S. T. Buckland, and A. R. Pople. 2008. The influence of animal mobility on the assumption of uniform distances in aerial line-transect surveys. Wildlife Research 35:275-288.

Frost, H. C., G. L. Storm, M. J. Batcheller, and M. J. Lovallo. 1997. White-tailed deer management at Gettysburg National Military Park and Eisenhower National Historic site. Wildlife Society Bulletin 25:462-469.

Gill, R. M. A., M. L. Thomas, and D. Stocker. 1997. The use of portable thermal imaging for estimating deer population density in forest habitats. Journal of Applied Ecology 34:1273-1286.

Haroldson, B. S., E. P. Wiggers, J. Beringer, L. P. Hansen, and J. B. McAninch. 2003. Evaluation of aerial thermal imaging for detecting white-tailed deer in a deciduous forest environment. Wildlife Society Bulletin 24:37-43.

Haulton, S.M., W.F. Porter, and B.A. Rudolph. 2001. Evaluating 4 methods to capture white-tailed deer. Wildlife Society Bulletin 29:255-264.

Hounsome, T. D., R. P. Young, J. Davison, R. W. Yarnell, I. D. Trewby, B. T. Garnett, R. J. Delahay, and G. J. Wilson. 2005. An evaluation of distance sampling to estimate badger (*Meles meles*) abundance. Journal of Zoology 266:81-87.

Hudson, D. J. 1971. Interval estimation from the likelihood function. Journal of the Royal Statistical Society Series B 33:256-262.

Koenen, K. G., S. DeStefano, and P. R. Krausman. 2002. Using distance sampling to estimate seasonal densities of desert mule deer in a semidesert grassland. Wildlife Society Bulletin 30:53-63.

Koganezawa, M., and Y. Li. 2002. Sika deer response to spotlight counts: implications for distance sampling of population density. Mammal Study 27:95-99.

Kie, J. G., and B. B. Boroski. 1995. Using spotlight counts to estimate mule deer population size and trends. California Fish and Game 81:55-70.

Lancia, R. L., J. W. Bishir, M. C. Conner, and C. S. Rosenberry. 1996. Use of catch-effort to estimate population size. Wildlife Society Bulletin 24:731-737.

Leopold, A. S. 1963. Study of wildlife problems in national parks. Transactions of the North American Wildlife Conference 28:28-45.

Lewis, J. C., and J. W. Farrar. 1968. An attempt to use the Leslie method on deer. Journal of Wildlife Management 32:760-764.

Manly, B. F. J., L. L. McDonald, D. L. Thomas, T. L. McDonald, and W. P. Erickson. 2002. Resource selection by animals: statistical design and analysis for field studies. Second Edition. Kluwer Academic, Dordrecht, the Netherlands.

Marques, F. C., S. T. Buckland, D. Goffin, C. E. Dixon, D. L. Borchers, B. A. Mayle, and A. J. Peace. 2001. Estimating deer abundance from line transect surveys of dung: sika deer in southern Scotland. Journal of Applied Ecology 38:349-363.

Marques T. A. 2007. Incorporating measurement error and density gradients in distance sampling surveys. Dissertation, University of Saint Andrews, Scotland, United Kingdom.

Marques T. A., L. Thomas, S. G. Fancy, and S. T. Buckland. 2007. Improving estimates of bird density using multiple-covariate distance sampling. The Auk. 124:1229-1243.

McClintock, B. T., G. C. White, and K. P. Burnham. 2006. A robust design mark-resight abundance estimator allowing heterogeneity in resighting probabilities. Journal of Agricultural, Biological, and Environmental Statistics 11:231-248.

McClintock, B. T., G. C. White, M. F. Antolin, and D. W. Tripp. 2009. Estimating abundance using mark-resight when sampling is with replacement or the number of marked individuals is unknown. Biometrics 65:237-246.

McCullough, D. R. 1982. Evaluation of night spotlighting as a deer study technique. Journal of Wildlife Management 46:963-973.

McCullough, D. R., and D. H. Hirth. 1988. Evaluation of the Petersen-Lincoln estimator for a white-tailed deer population. Journal of Wildlife Management 52:534-544.

Millspaugh, J. J., R. M. Nielson, L. L. McDonald, J. M. Marzluff, R. A. Gitzen, C. D. Rittenhouse, M. W. Hubbard, and S. L. Sheriff. 2006. Analysis of resource selection using utilization distributions. Journal of Wildlife Management 70:384-395.

Mitchell, B. D., J. J. Rowe, P. R. Ratcliffe, and M. Hinge. 1985. Defecation frequency in roe deer (*Capreolus capreolus*) in relation to the accumulation rates of faecal deposits. Journal of Zoology 207:1-7.

Naugle, D. E., J. A. Jenks, and B. J. Kernohan. 1996. Use of thermal infrared sensing to estimate density of white-tailed deer. Wildlife Society Bulletin 24:37-43.

Neal, A. K. 1990. Evaluation of mark-resight population estimates using simulations and field data from mountain sheep. Thesis, Colorado State University, Fort Collins, USA.

Neal, A. K., G. C. White, R. B. Gill, D. F. Reed, and J. H. Olterman. 1993. Evaluation of mark-resight model assumptions for estimating mountain sheep numbers. Journal of Wildlife Management 57:436-450.

Niewinski, A. T., T. W. Bowersox, R. L. Laughlin. 2006. Vegetation status in selected woodlots at Gettysburg National Military Park pre and post white-tailed deer management. Technical Report NPS/NER/NRTR-2006/037. National Park Service. Philadelphia, Pennsylvania, USA.

Otis, D. L., K. P. Burnham, G. C. White, and D. R. Anderson. 1978. Statistical inference from capture data on closed animal populations. Wildlife Monographs 62:3-135.

Pollock, K. H., and W. L. Kendall. 1987. Visibility bias in aerial surveys: a review of estimation procedures. Journal of Wildlife Management 51:502-510.

Quinn, T. J. II, and R. B. Deriso. 1999. Quantitative Fish Dynamics. Oxford University Press, New York, New York, USA.

Rice, W. R., and J. D. Harder. 1977. Application of multiple aerial sampling to a mark-recapture census of white-tailed deer. Journal of Wildlife Management 41:197-206.

Rosenberry, C. S., R. A. Lancia, and M. C. Conner. 1999. Population effects of white-tailed deer dispersal. Wildlife Society Bulletin 27:858-864.

Rost, G. R., and J. A. Bailey. 1979. Distribution of mule deer and elk in relation to roads. Journal of Wildlife Management 43:634-641.

Ruette, S., P. Stahl, and M. Albaret. 2003. Applying distance-sampling methods to spotlight counts of red foxes. Journal of Applied Ecology 40:32-43.

Sage, R. W., W. C. Tierson, G. F. Mattfeld, and D. F. Behrend. 1983. White-tailed deer visibility and behavior along forest roads. Journal of Wildlife Management 47:940-953.

Samuel, M. D., E. O. Garton, M. W. Schlegel, and R. G. Carson. 1987. Visibility bias during aerial surveys of elk in northcentral Idaho. Journal of Wildlife Management 51:622-630.

Samuel, M. D., R. K. Steinhorst, E. O. Garton, and J. W. Unsworth. 1992. Estimation of wildlife population ratios incorporating survey design and visibility bias. Journal of Wildlife Management 56:718-725.

Sawyer, H., R. M. Nielson, F. Lindzey, and L. L. McDonald. 2006. Winter habitat selection of mule deer before and during development of a natural gas field. Journal of Wildlife Management 70:396-403.

Seber, G. A. F. 1982. The estimation of animal abundance and related parameters. Second edition. MacMillan, New York, New York, USA.

Steinhorst, R. K., and M. D. Samuel. 1989. Sightability adjustment methods for aerial surveys in wildlife populations. Biometrics 45:415-425.

Storm, G. L., D. F. Cottam, R. H. Yahner, and J. D. Nichols. 1992. A comparison of two techniques for estimating deer density. Wildlife Society Bulletin 20:197-203.

Storm, G. L., R. H. Yahner, D. F. Cottam, and G. M Vecellio. 1989. Population status, movements, habitat use, and impact of white-tailed deer at Gettysburg National Military Park and Eisenhower National Historic Site, Pennsylvania. Technical

Report NPS/MAR/NRTR-89/043. National Park Service. Philadelphia, Pennsylvania, USA.

Thomas, L., S. T. Buckland, E. A. Rexstad, J. L. Laake, S. Strindberg, S. L. Hedley, J. R. B. Bishop, T. A. Marques, and K. P. Burnham. 2010. Distance software: design and analysis of distance sampling surveys for estimating population size. Journal of Applied Ecology 47:5-14.

Tzilkowski, W. M., and G. L. Storm. 1993. Detecting change using repeated measures analysis: white-tailed deer abundance at Gettysburg National Military Park. Wildlife Society Bulletin 21:411-414.

U.S. Department of the Interior. 1995. Final environmental impact statement: white-tailed deer management plan; Gettysburg National Military Park and Eisenhower National Historic Site. U.S. Department of the Interior, National Park Service.

Van Etten, R. C., and C. L. Bennet. 1965. Some sources of error in using pellet-group counts for censusing deer. Journal of Wildlife Management 29:723-729.

Vecellio, G. M., R. H. Yahner, and G. L. Storm. 1994. Crop damage by deer at Gettysburg Park. Wildlife Society Bulletin 22:89-93.

Venturato, E., P. Cavallini, and F. Dessi-Fulgheri. 2010. Are pheasants attracted or repelled by roads? A test of a crucial assumption of transect censuses. European Journal of Wildlife Resources 56:233-237.

Venzon, D. J., and S. H. Moolgavkar. 1988. A method for computing profile-likelihood based confidence intervals. Applied Statistics 37:87-94.

Ward, A. I, P. C. L. White, and C. H. Critchley. 2004. Roe deer *Capreolus capreolus* behaviour affects density estimates from distance sampling surveys. Mammal Review 34:315-319.

Warren, R. J. 1991. Ecological justification for controlling deer populations in eastern national parks. Transactions of the North American Wildlife and Natural Resources Conference 56:56-66.

White, G. C., and R. A. Garrott. 1990. Analysis of wildlife radio-tracking data. Academic Press, New York, New York, USA.

White, G. C., and T. M. Shenk. 2001. Population estimation with radio-marked animals.. Pages 329-350. in J. J Millspaugh, and J. M. Marzluff, editors. Design and analysis of wildlife radiotelemetry studies. Academic Press, San Diego, California, USA.

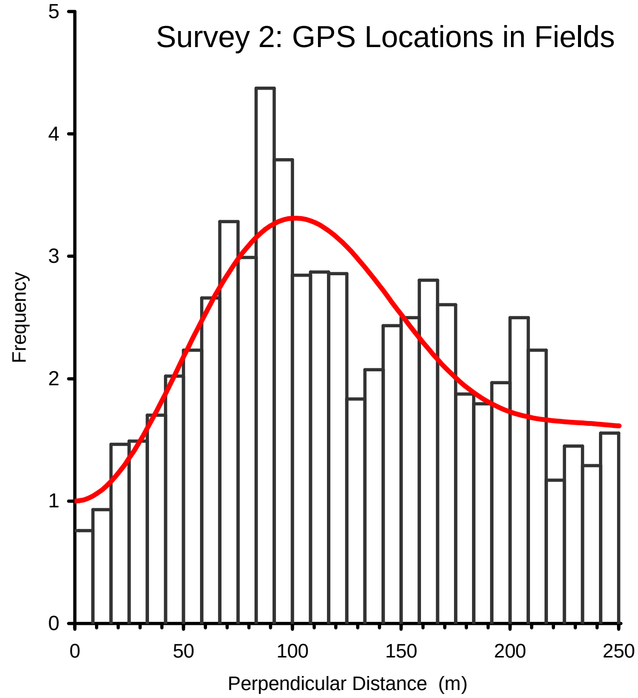
White G. C. 1996. NOREMARK: Population estimation from mark-resighting surveys. Wildlife Society Bulletin 24:50-52.

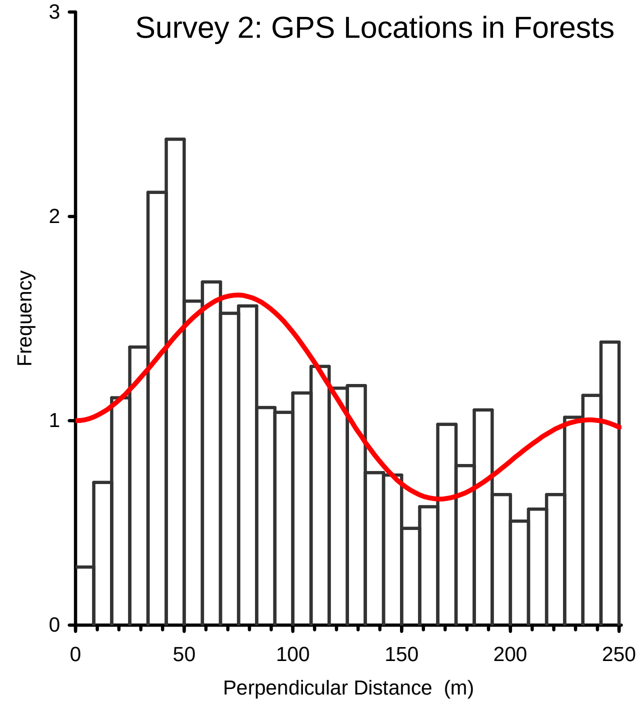
White, G. C. 1996. NOREMARK. [Online] Available at **www.cnr.colostate.edu/~gwhite/software.html**. Accessed 8 Aug 2009.

Yost, A. C., and R. G. Wright. 2001. Moose, caribou, and grizzly bear distribution in relation to road traffic in Denali National Park, Alaska. Arctic 54:41-48.

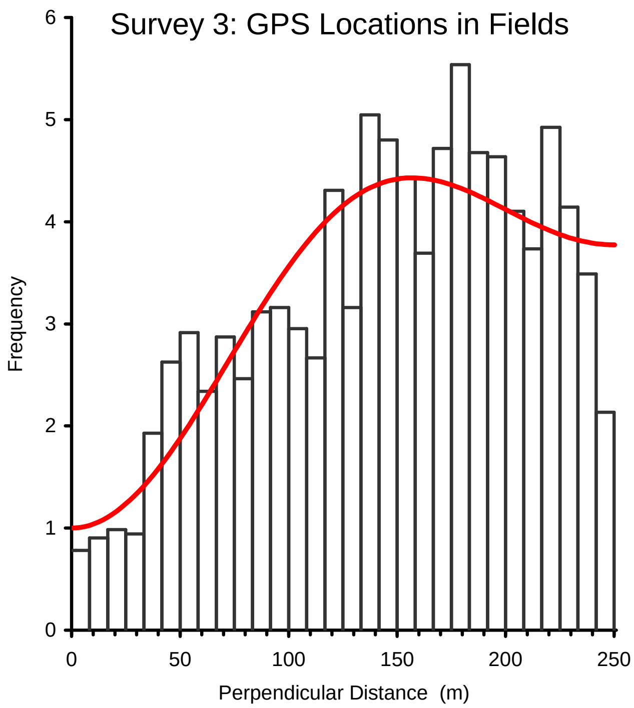
APPENDIX A: Distribution of Deer Relative to Transects

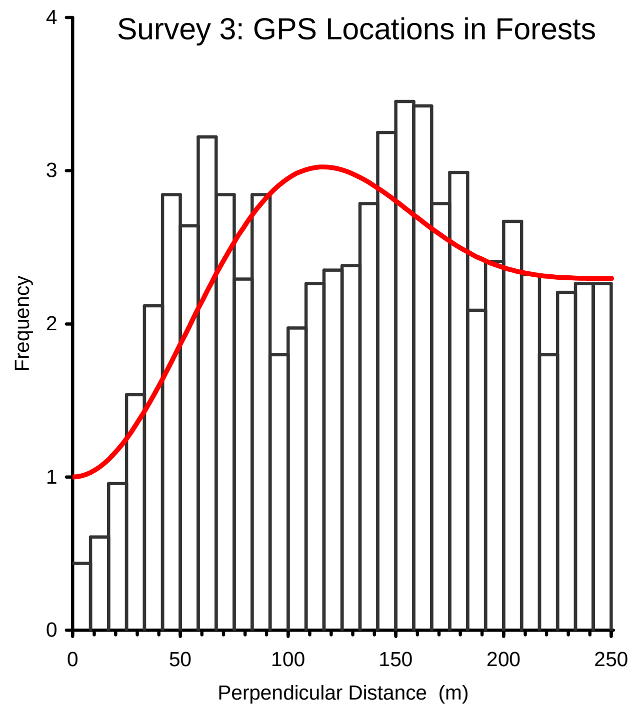
(a) Distribution of global positioning system (GPS) locations in open areas and forested areas relative to perpendicular distance from transects during the second distance sampling survey, August 3-9, 2009.



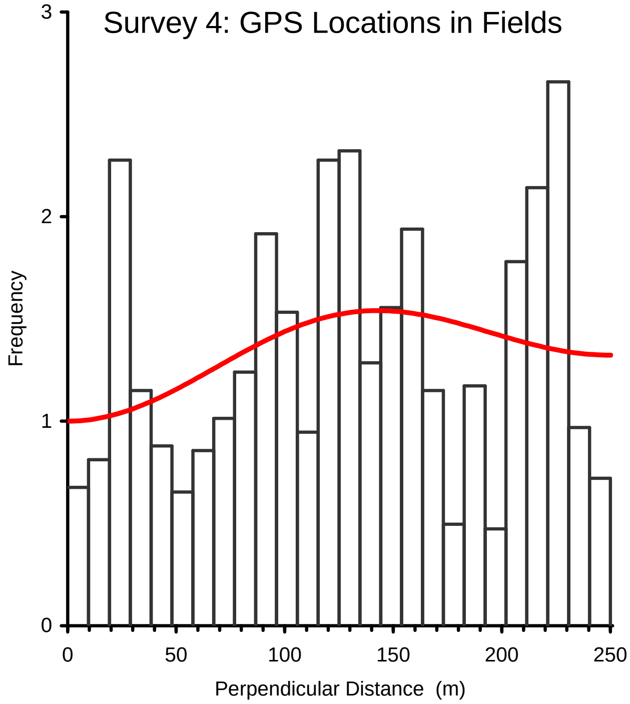


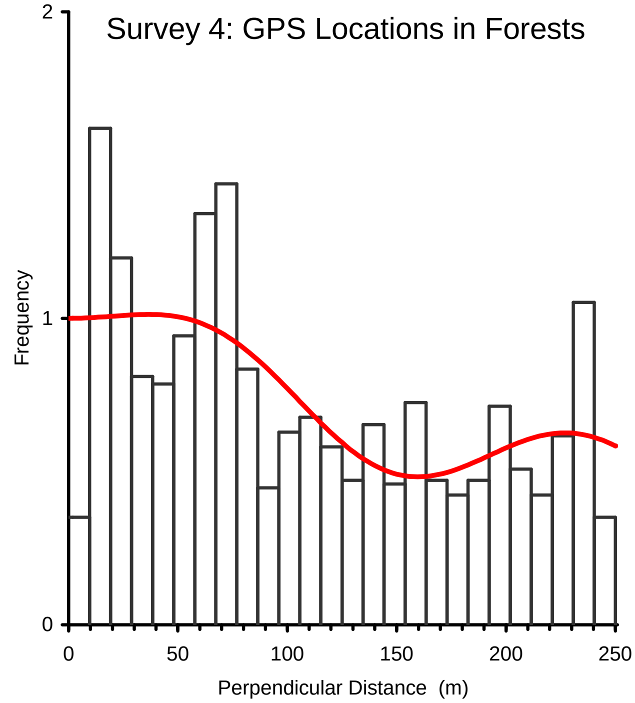
(b) Distribution of GPS locations in open areas and forested areas relative to perpendicular distance from transects during the third distance sampling survey, November 20-25, 2009.



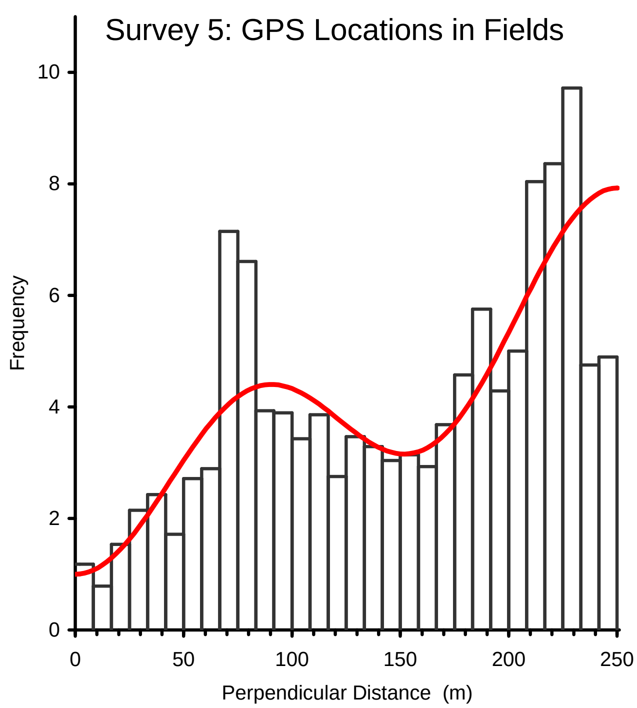


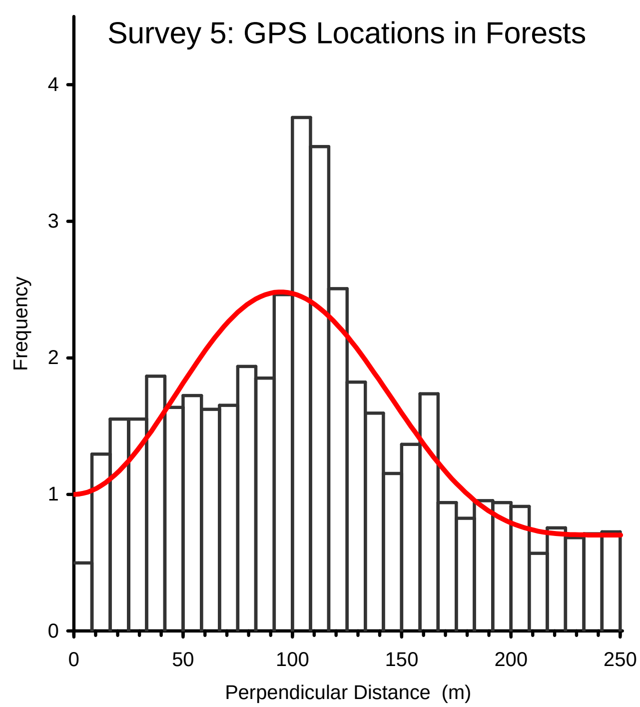
(c) Distribution of GPS locations in open areas and forested areas relative to perpendicular distance from transects during the fourth distance sampling survey, January 5-8, 2010.



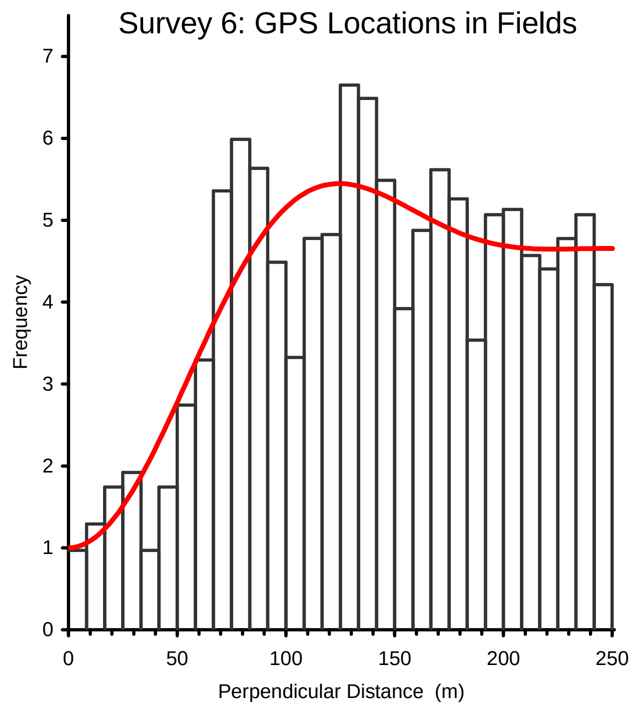


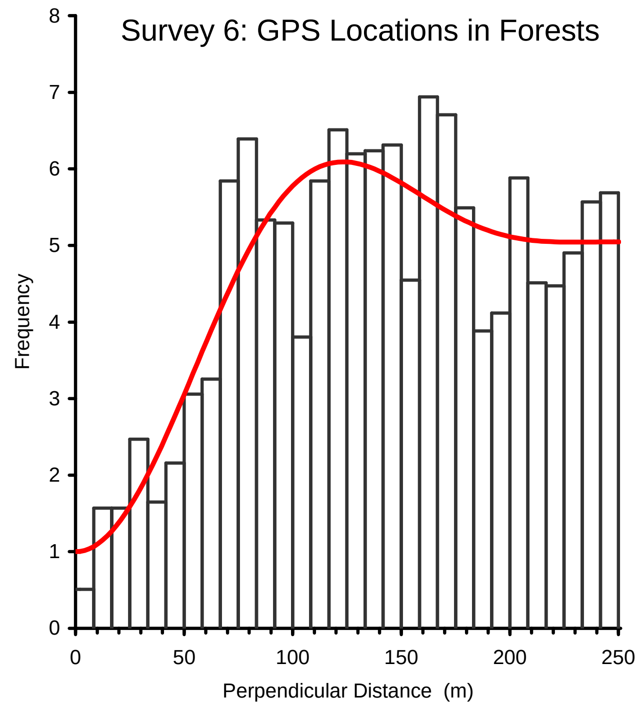
(d) Distribution of GPS locations in open areas and forested areas relative to perpendicular distance from transects during the fifth distance sampling survey, April 1-4, 2010.



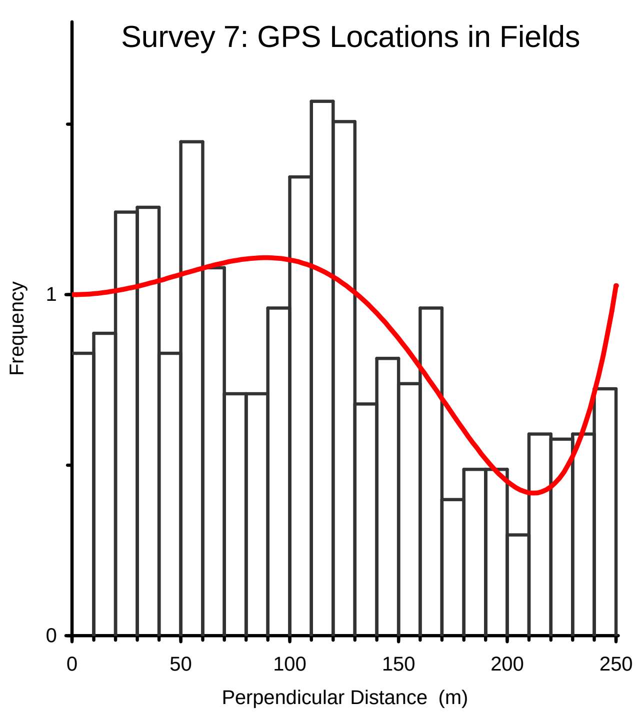


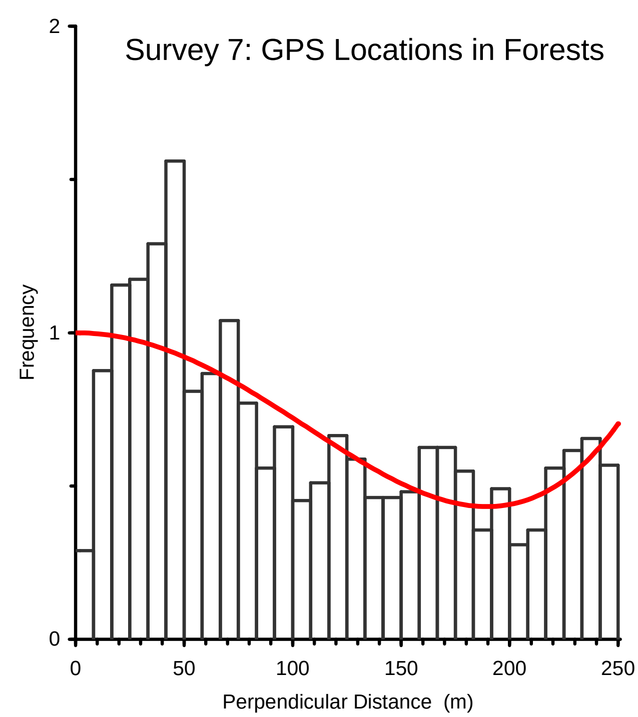
(e) Distribution of GPS locations in open areas and forested areas relative to perpendicular distance from transects during the sixth distance sampling survey, August 25-30, 2010.



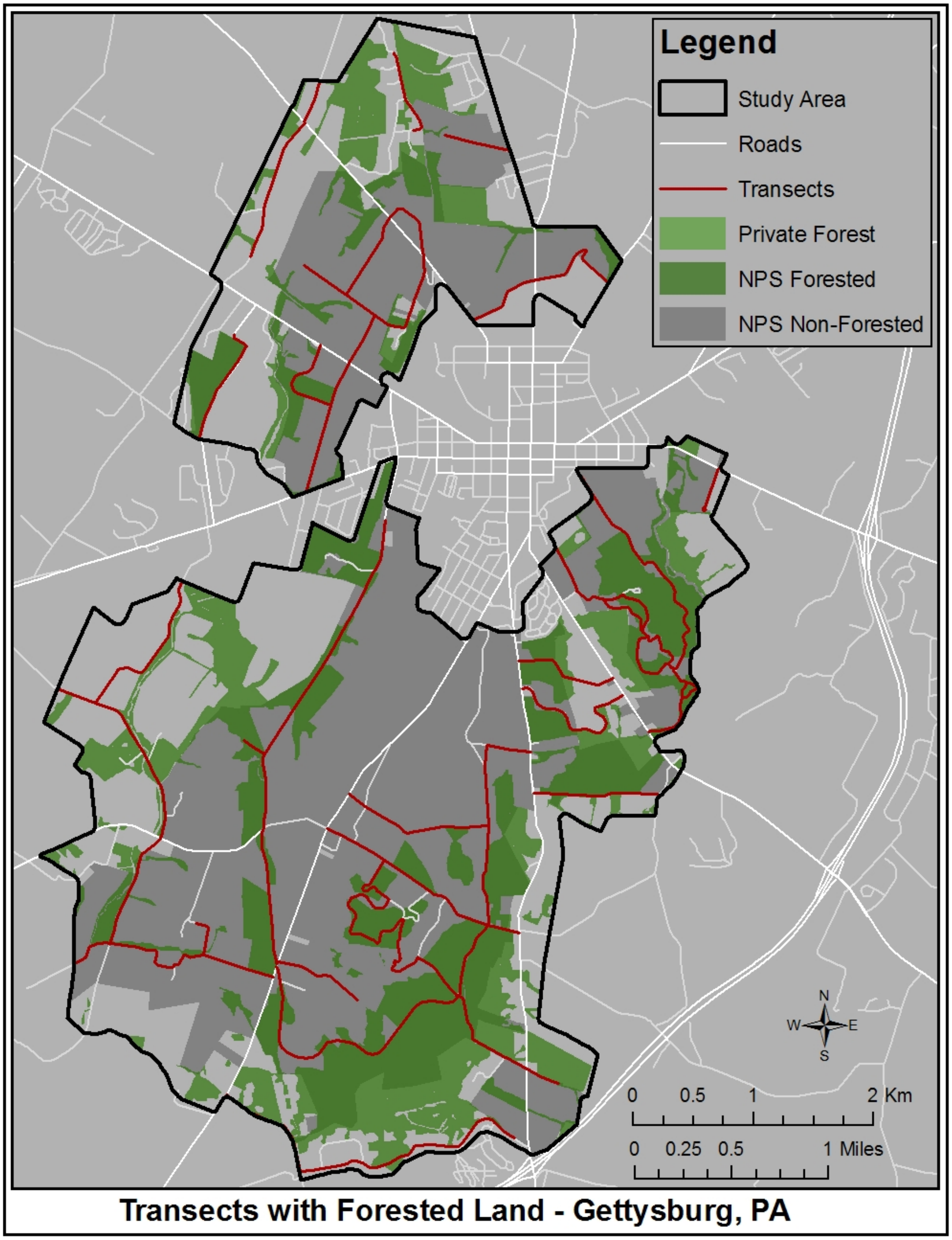


(f) Distribution of GPS locations in open areas and forested areas relative to perpendicular distance from transects during the seventh distance sampling survey, November 15-19, 2010.

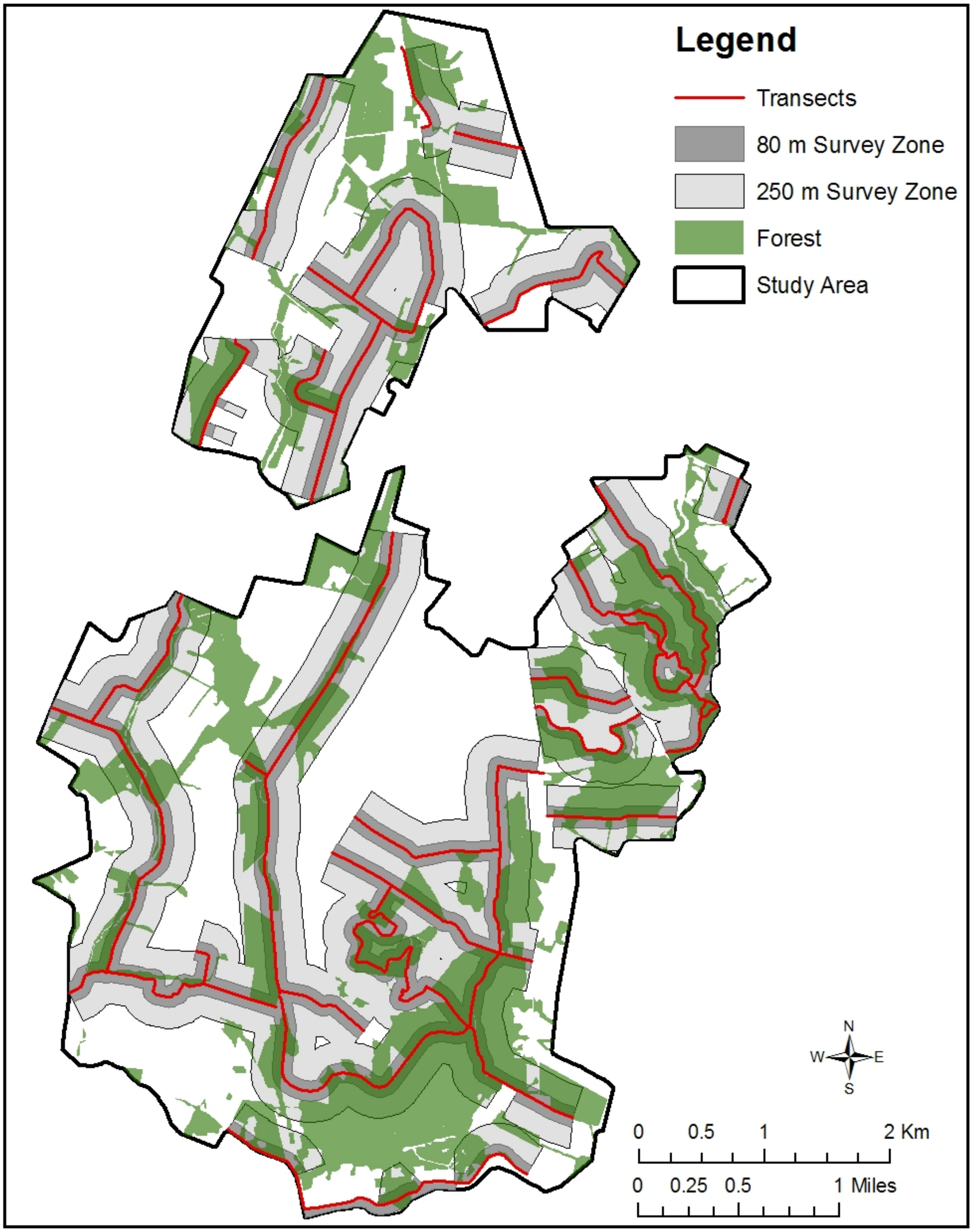




APPENDIX B: Forested Land in the Study Area



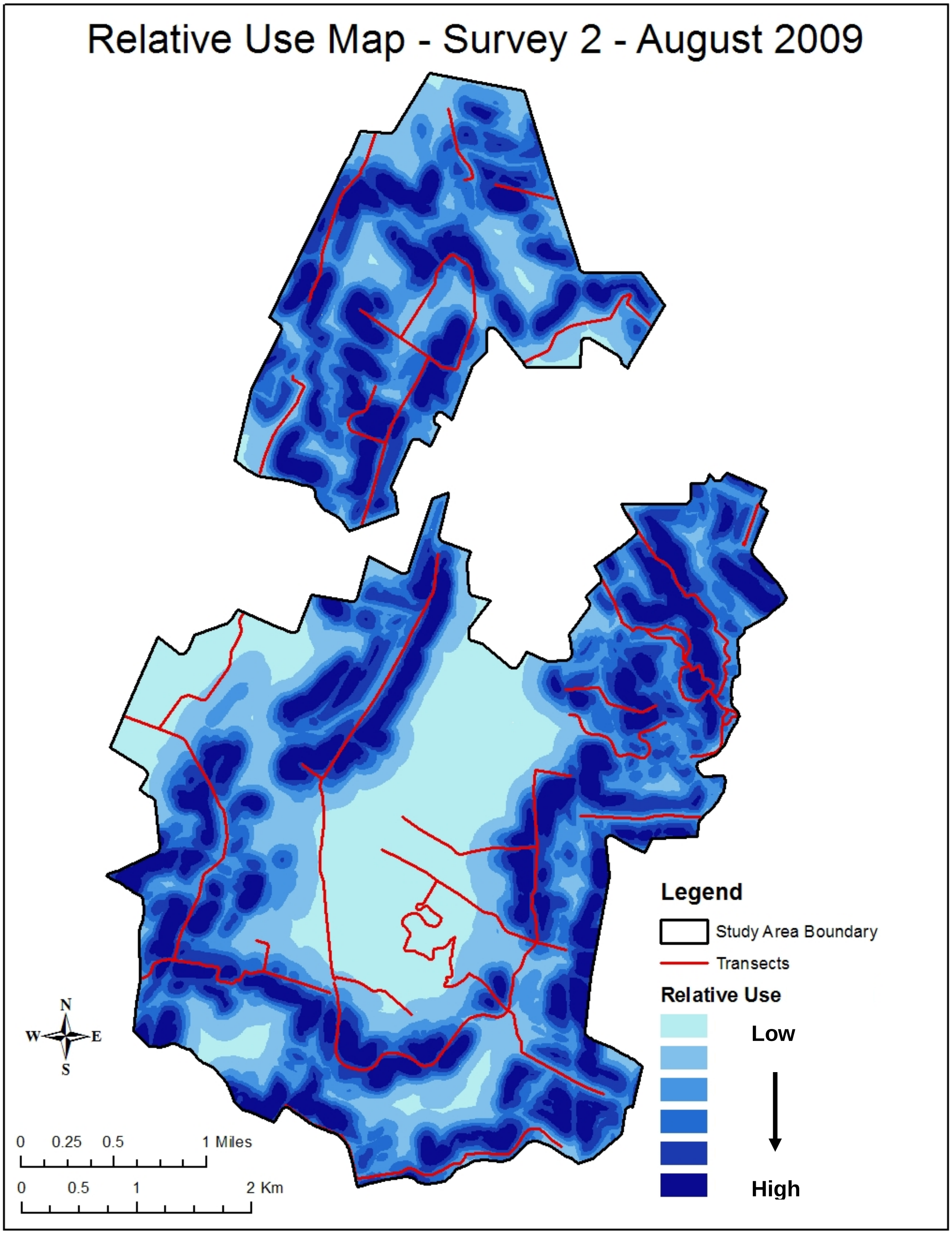
APPENDIX C: Survey Zones



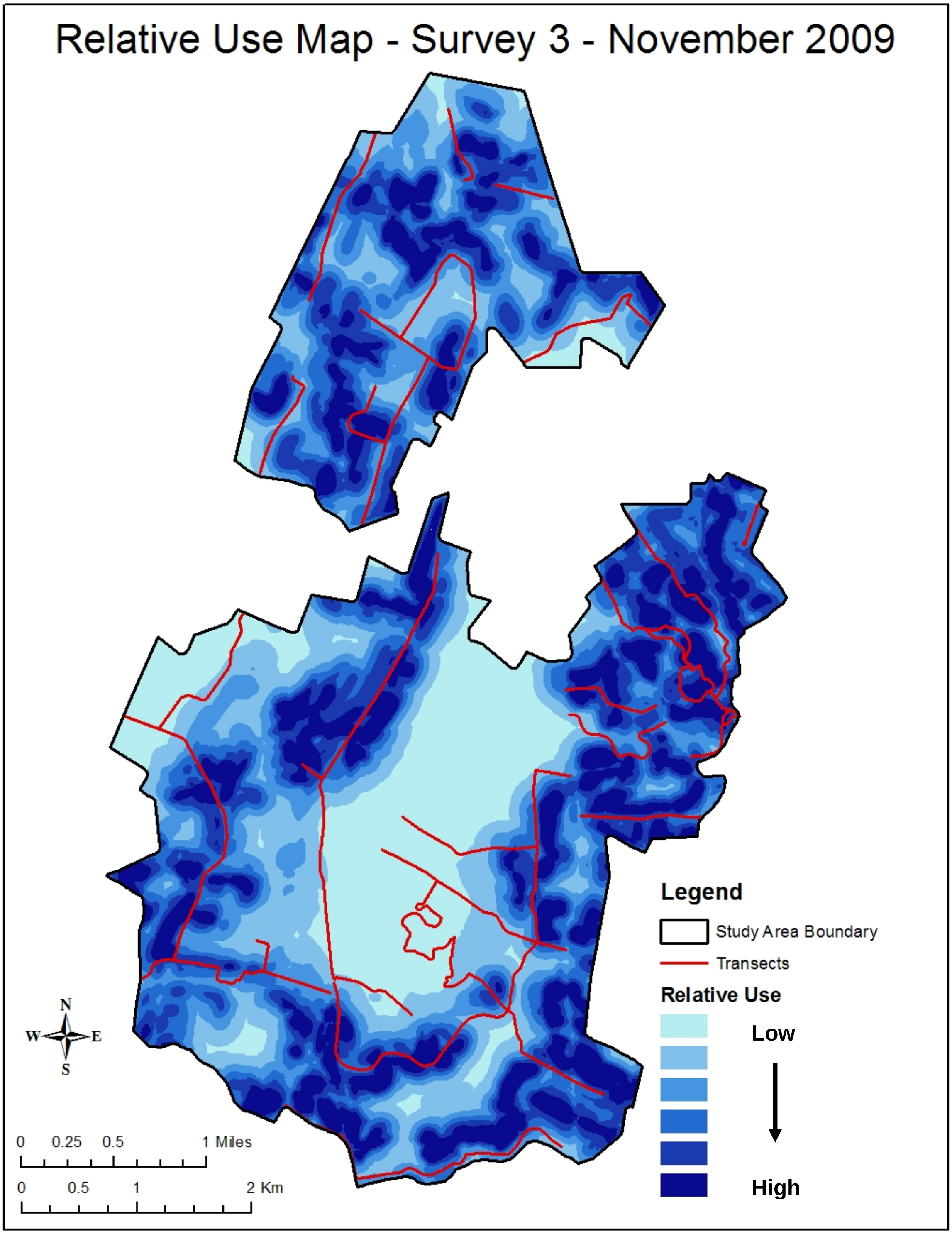
Survey Zones with Forested Land - Gettysburg, PA

APPENDIX D: Resource Selection Function Maps

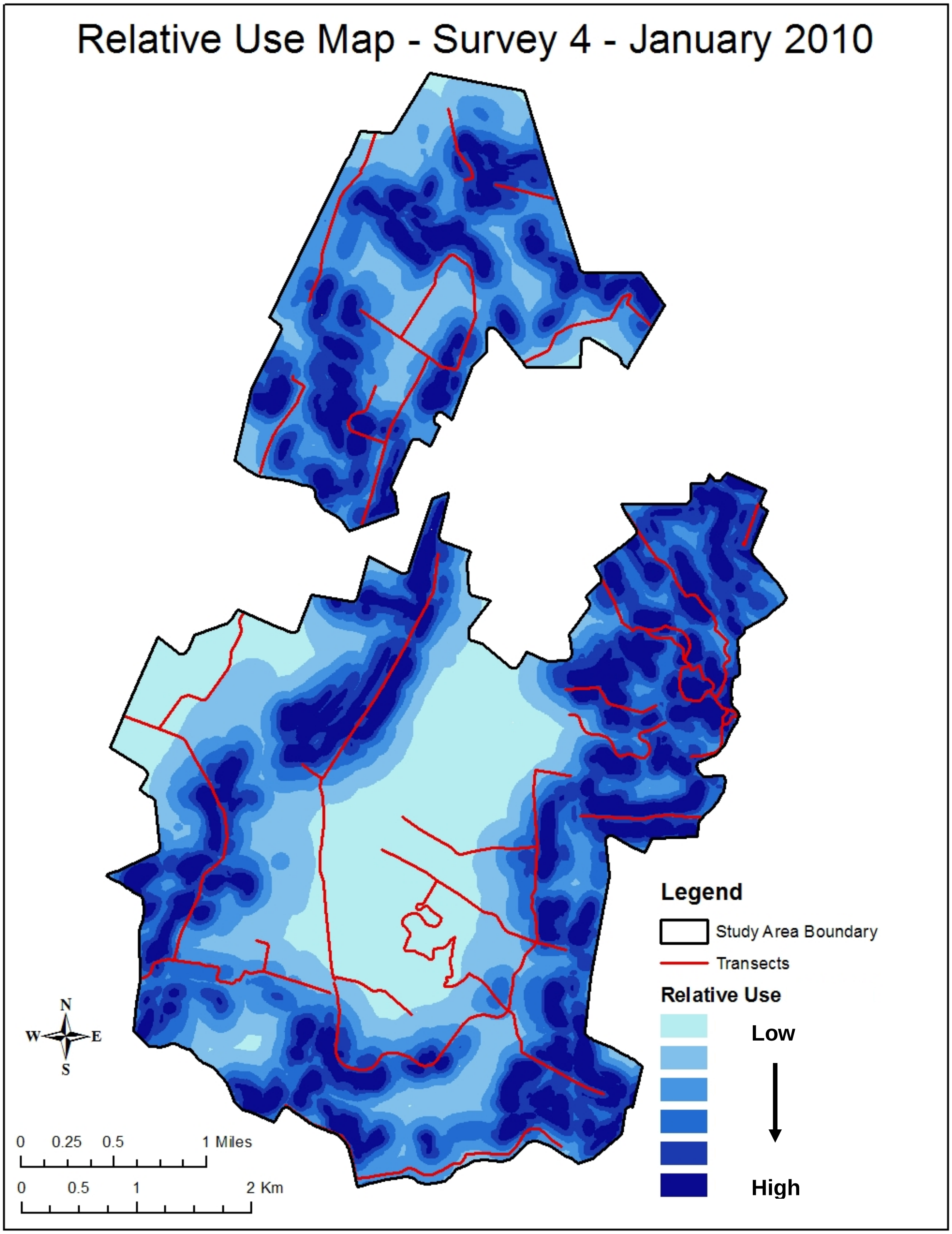
(a)



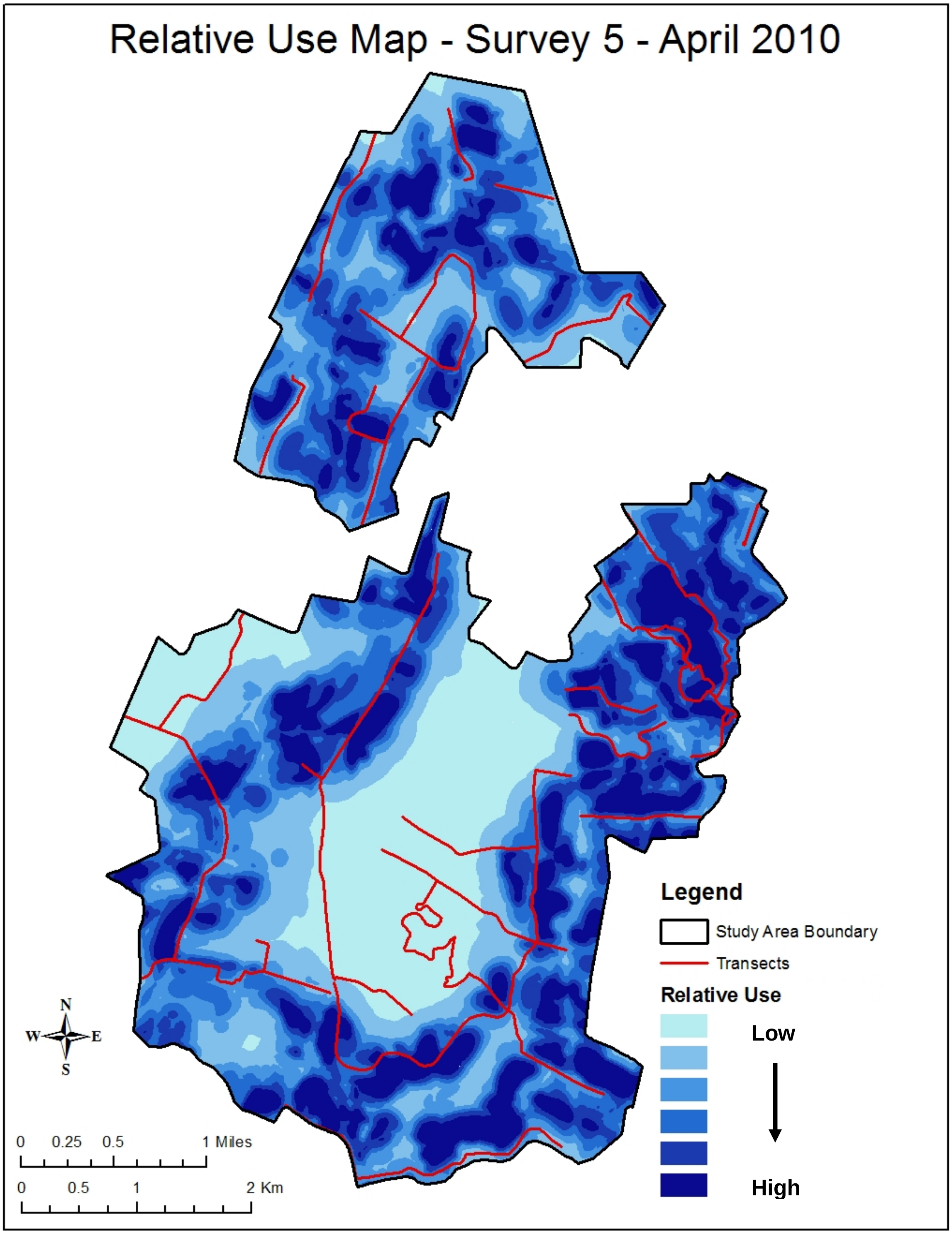
(b)


Relative Use Map - Survey 3 - November 2009
Legend
Study Area Boundary
Transects
Relative Use
Low
High
N
W
E
S
0 0.25 0.5 1 Miles
0 0.5 1 2 Km

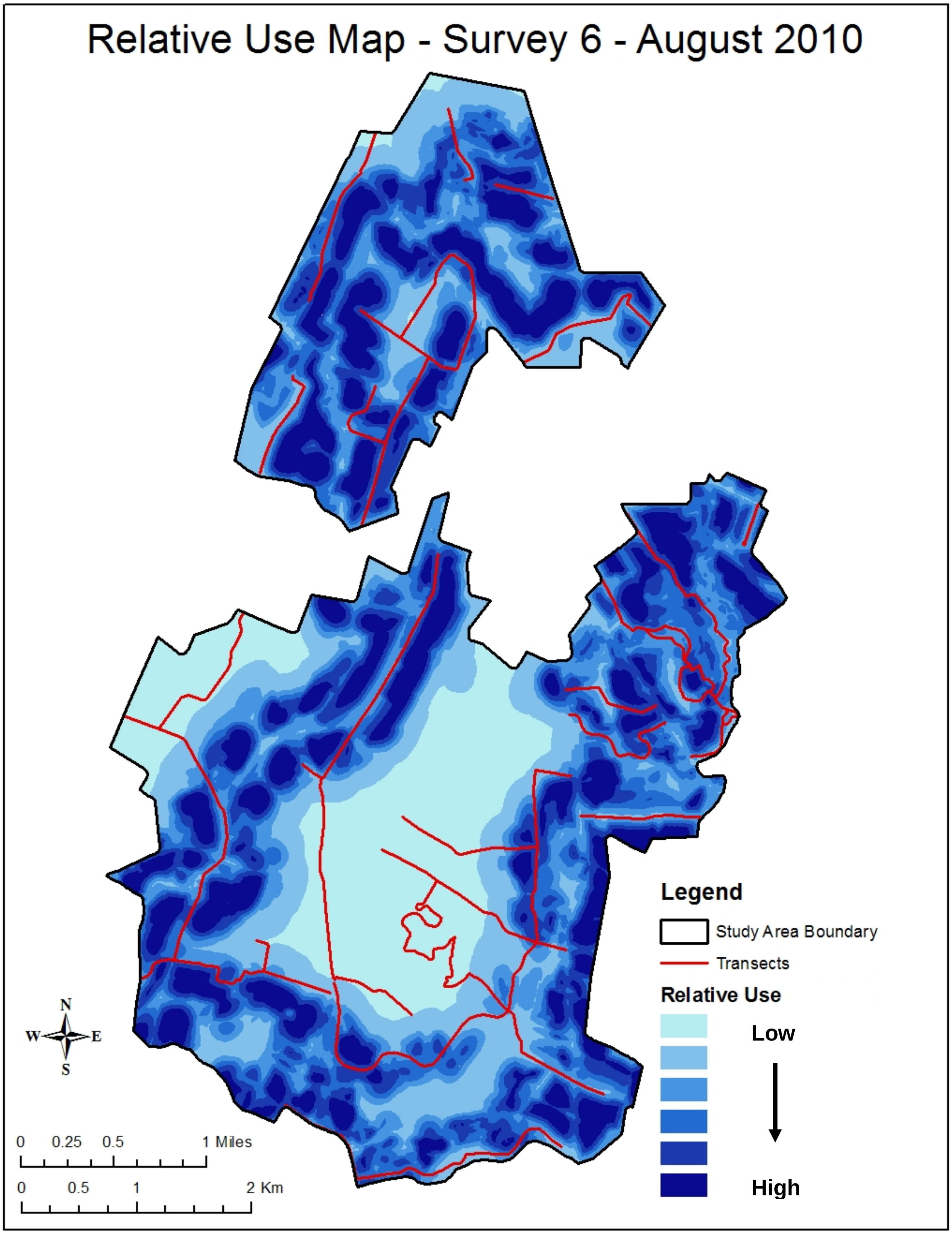
(c)


Relative Use Map - Survey 4 - January 2010
Legend
Study Area Boundary
Transects
Relative Use
Low
High
N
W
E
S
0 0.25 0.5 1 Miles
0 0.5 1 2 Km

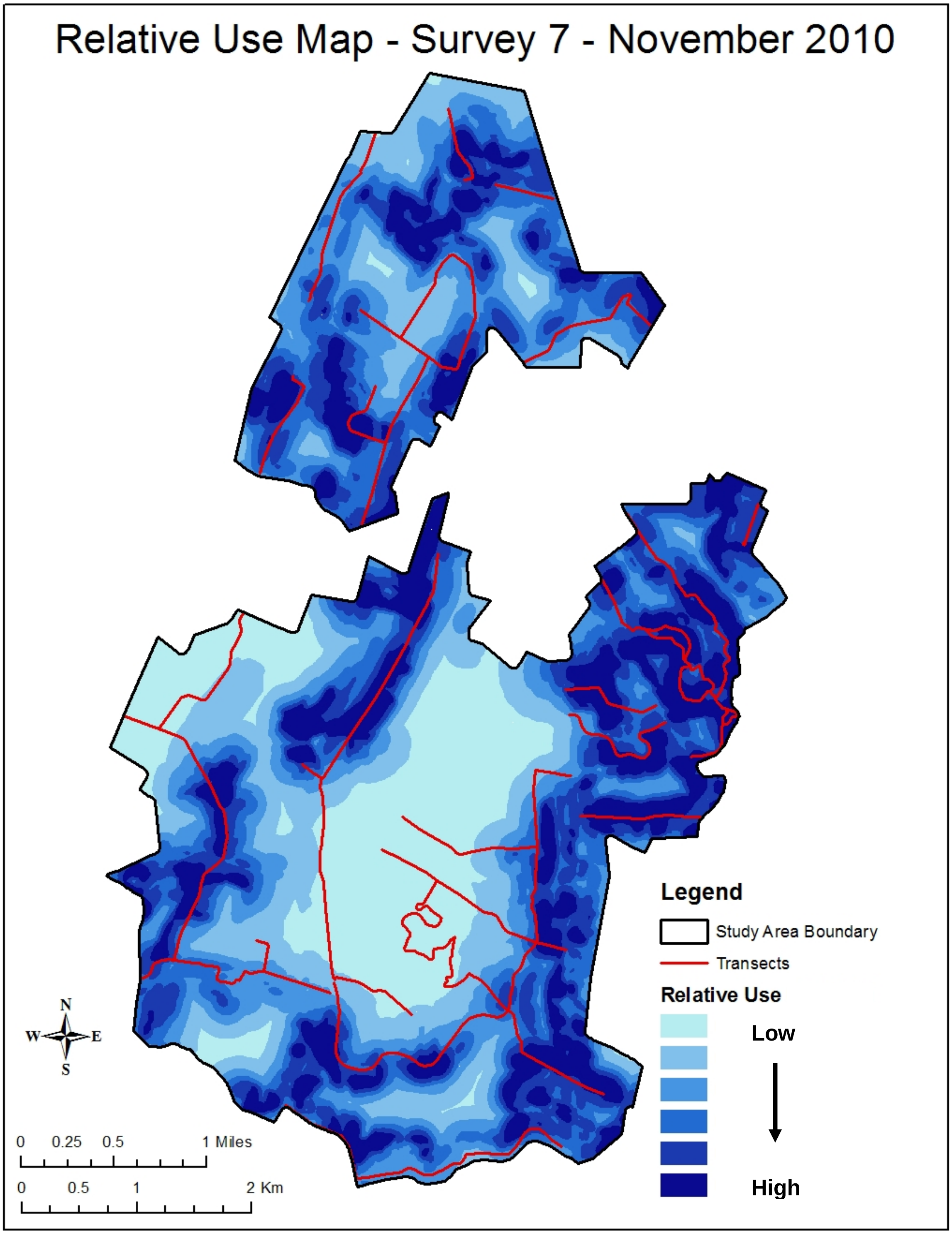
(d)


Relative Use Map - Survey 5 - April 2010
Legend
Study Area Boundary
Transects
Relative Use
Low
High
N
W
E
S
0 0.25 0.5 1 Miles
0 0.5 1 2 Km

(e)


Relative Use Map - Survey 6 - August 2010
Legend
Study Area Boundary
Transects
Relative Use
Low
High
N
W
E
S
0 0.25 0.5 1 Miles
0 0.5 1 2 Km

(f)


Relative Use Map - Survey 7 - November 2010
Legend
Study Area Boundary
Transects
Relative Use
Low
High
N
W
E
S
0 0.25 0.5 1 Miles
0 0.5 1 2 Km

APPENDIX E: Resource Selection Function Parameter Estimates

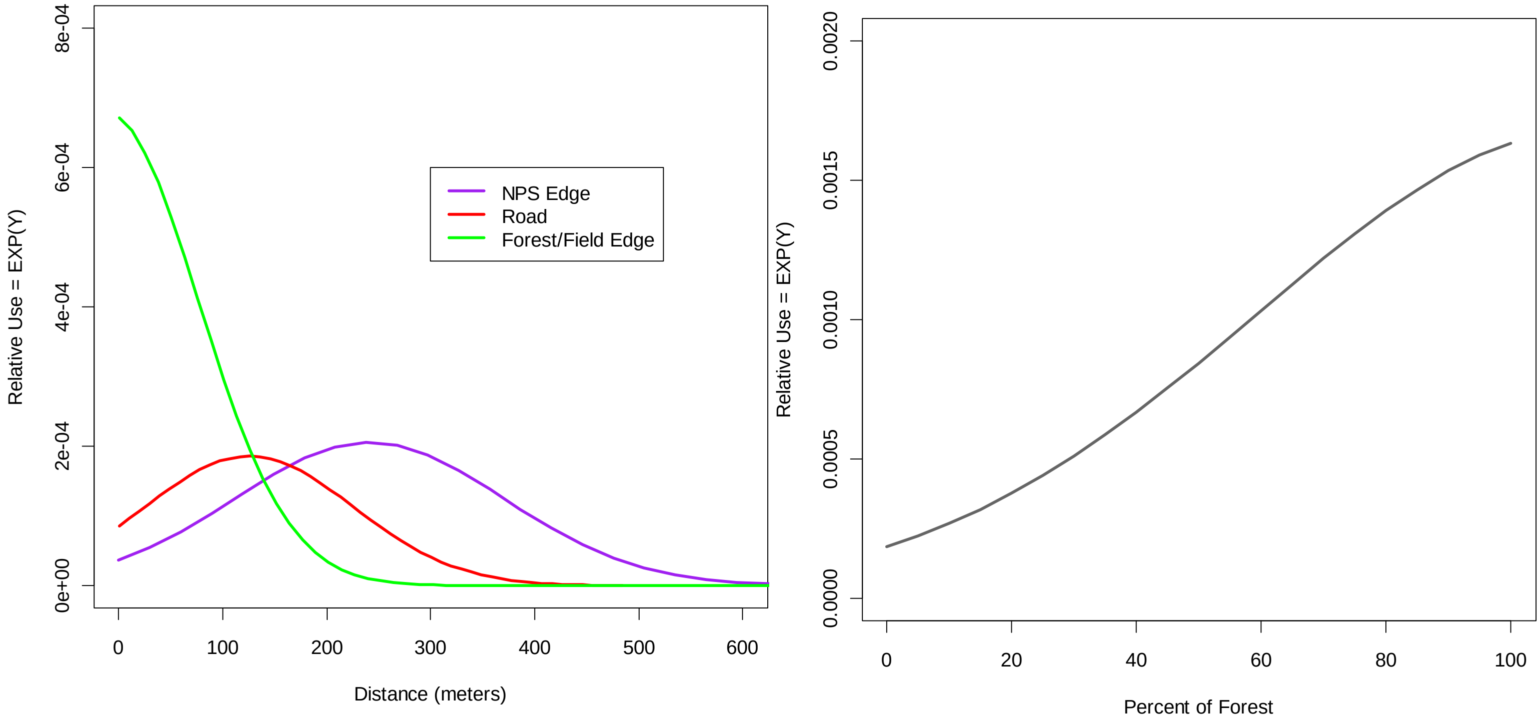
Appendix E. Weighted and standardized parameter estimates and associated standard errors from the zero-inflated negative binomial model 6 for all surveys, Gettysburg, Pennsylvania, 2009-2010. Also, *n* is the number of collared deer used in each analysis.

	April 2009 (*n*=29)		August 2009 (*n*=23)		November 2009 (*n*=17)		January 2010 (*n*=17)		April 2010 (*n*=26)		August 2010 (*n*=22)		November 2010 (*n*=17)	
Cov[a]	β_i	SE	β_i	SE	β_i	SE	β_i	SE	β_i	SE	β_i	SE	β_i	SE
I	-8.592	0.676	-7.026	0.573	-7.150	1.539	-7.444	0.712	-8.993	1.276	-6.530	0.861	-7.500	1.369
PF	3.886	0.331	-0.262	0.361	2.925	0.494	0.357	0.316	1.997	0.283	1.012	0.235	2.469	0.486
PF^2	-1.711	0.276	0.235	0.306	-2.067	0.505	0.047	0.302	-0.535	0.291	-1.867	0.276	-1.666	0.484
FFE	-2.339	0.209	-1.240	0.420	-1.889	0.287	-1.321	0.591	-0.898	0.4360	-0.357	0.541	-0.531	0.595
FFE^2	-1.057	0.213	-1.163	0.222	-1.239	0.365	-0.736	0.309	-0.468	0.337	-0.447	0.382	-0.313	0.551
RD	0.037	0.449	-0.240	0.582	0.004	0.696	-0.231	0.527	0.179	0.542	0.358	0.696	-0.640	0.549
RD^2	-0.463	0.261	-0.922	0.289	-0.524	0.332	-0.472	0.377	-0.334	0.377	-0.359	0.370	-0.587	0.417
NPS	-1.044	0.375	-1.121	0.585	-1.224	0.560	-2.648	0.536	-1.398	0.649	-1.814	0.523	-2.319	0.418
NPS^2	-2.630	0.184	-3.363	0.372	-1.747	0.421	-2.260	0.438	-2.512	0.258	-2.918	0.280	-1.427	0.310

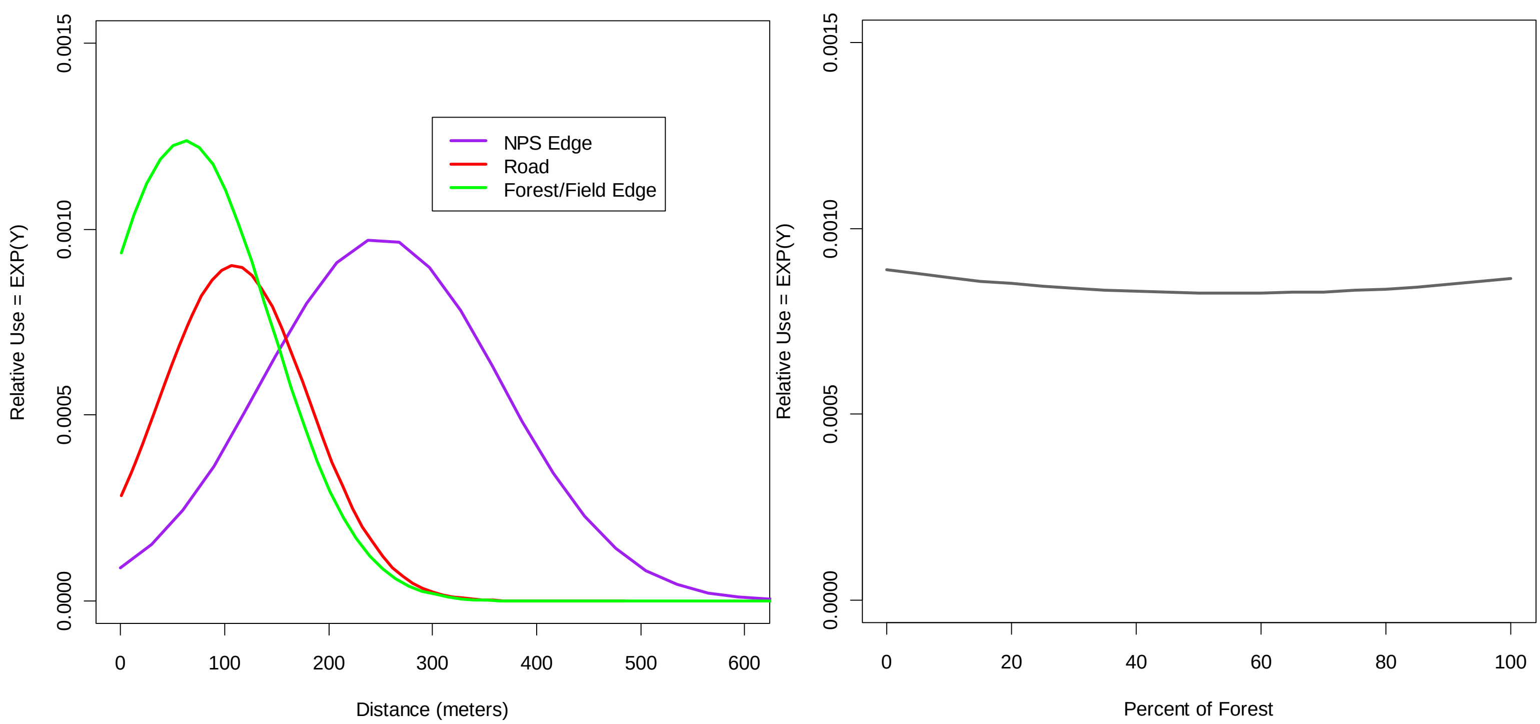
[a] Covariates: I = Intercept, PF = Percent Forest, FFE = distance to nearest forest-field edge, RD = distance to nearest road, and NPS = distance to nearest National Park Service owned land boundary.

APPENDIX F: Resource Selection Function Parameter Plots

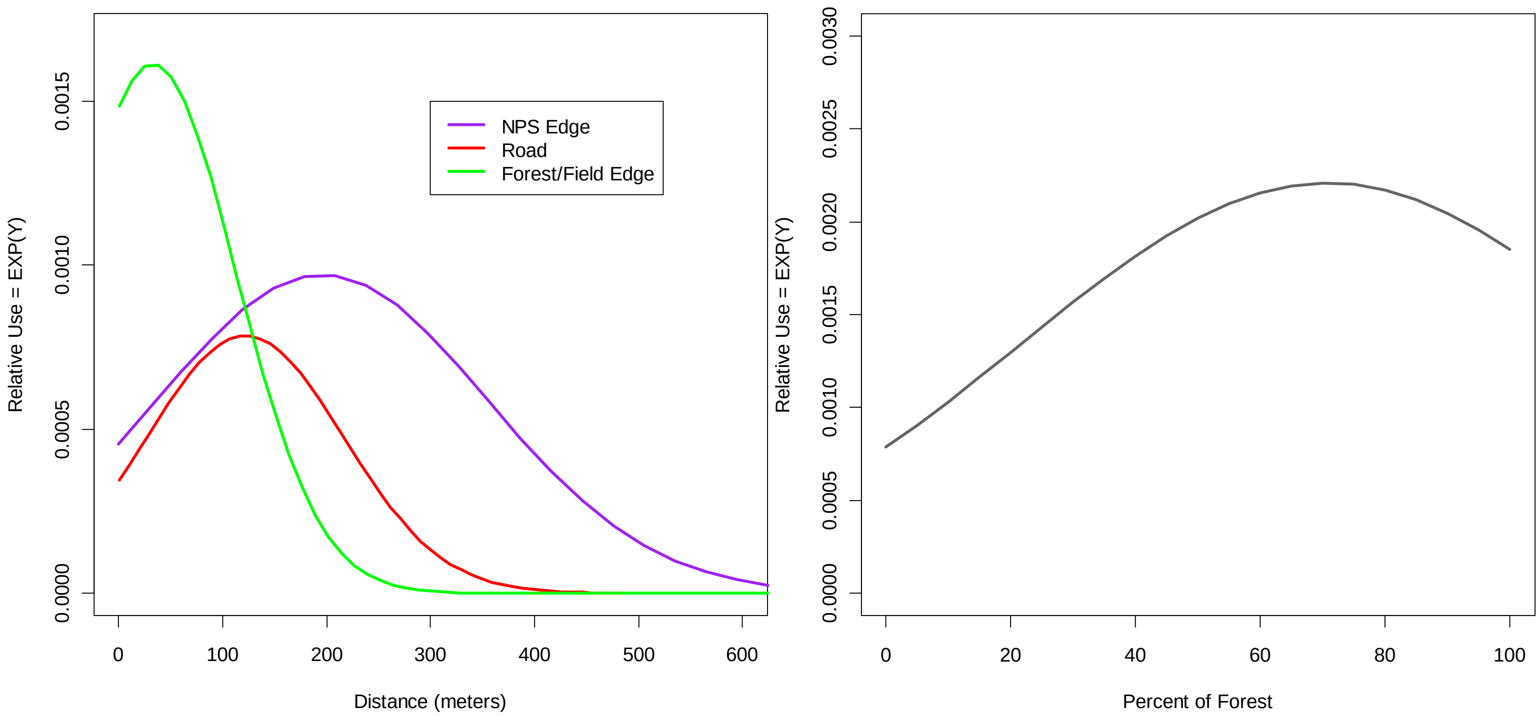
(a) Plotted parameter values from Appendix E, April 2009



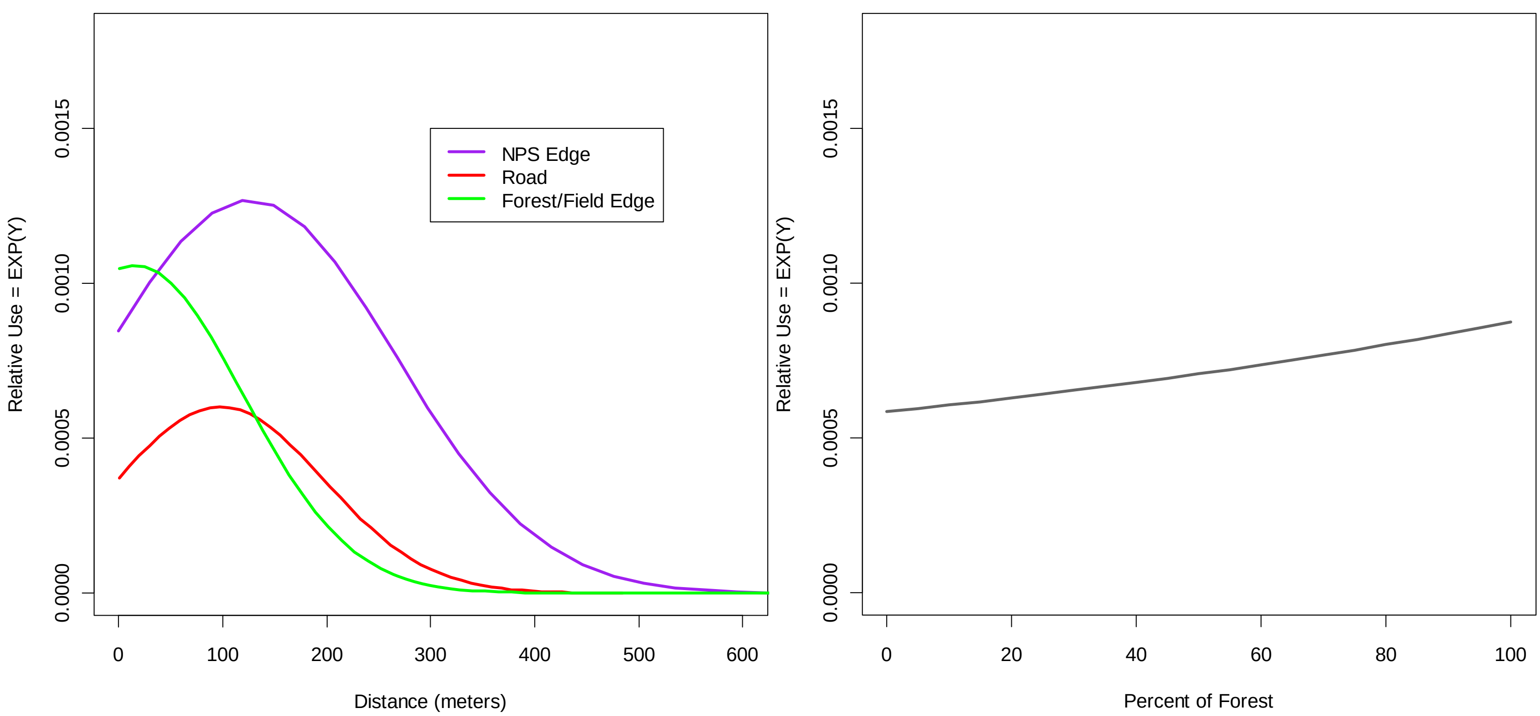
(b) Plotted parameter values from Appendix E, August 2009



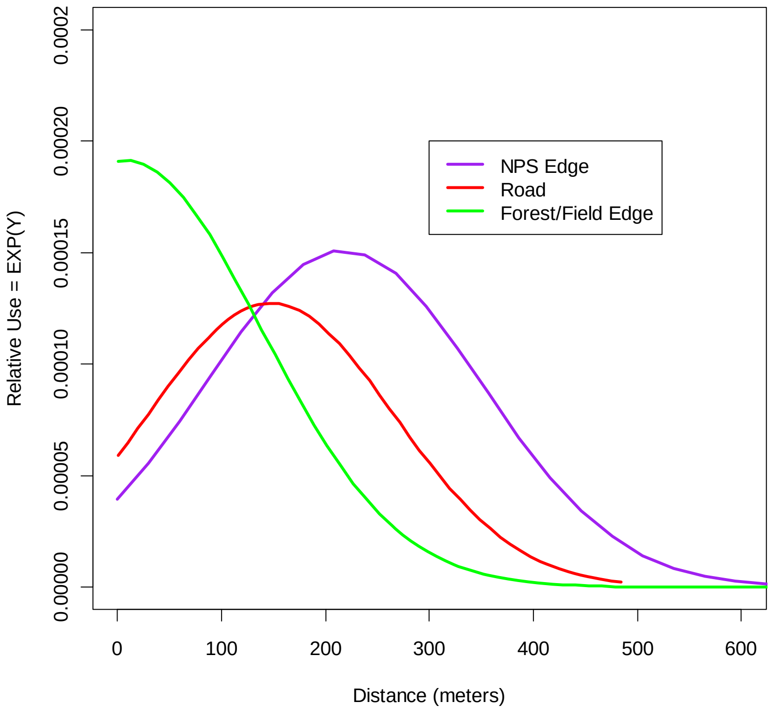
(c) Plotted parameter values from Appendix E, November 2009

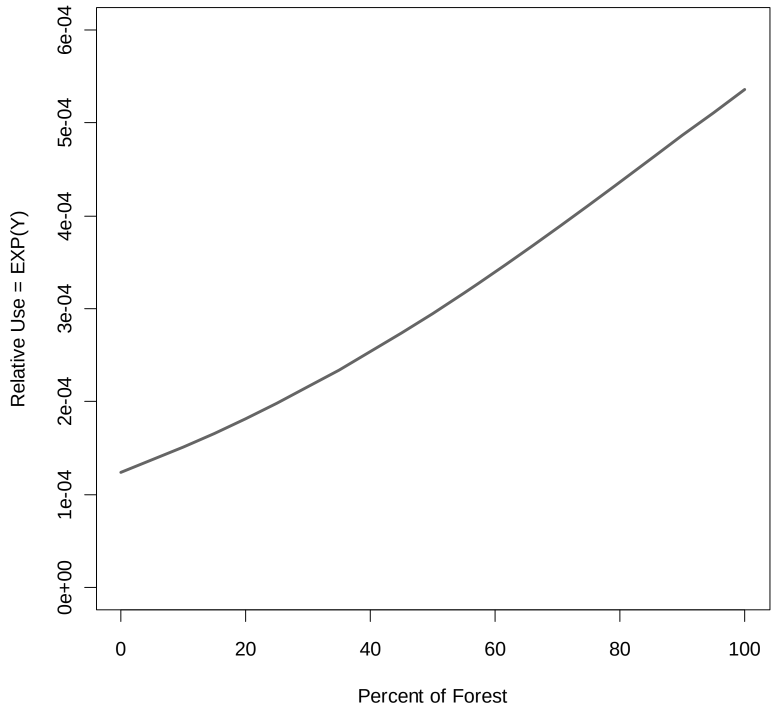


(d) Plotted parameter values from Appendix E, January 2010

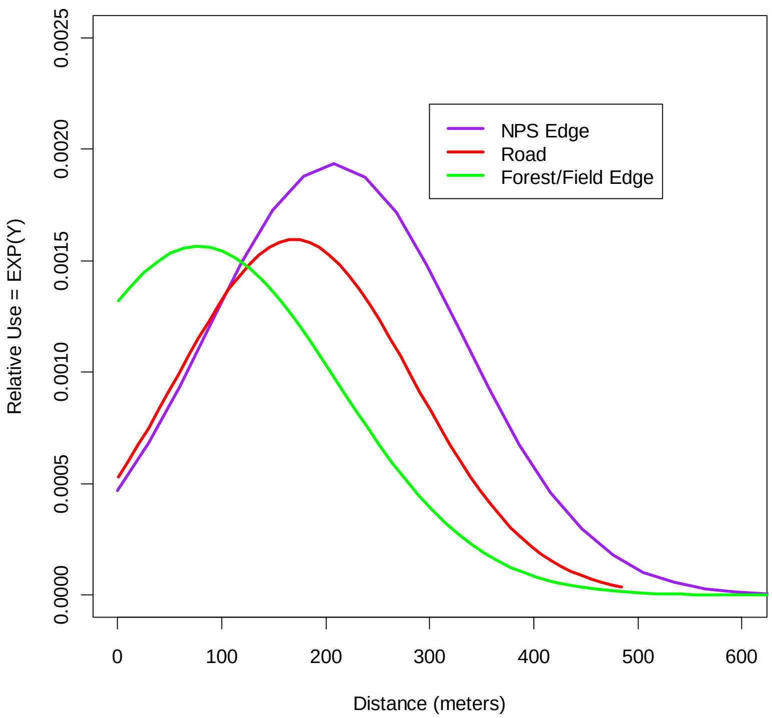


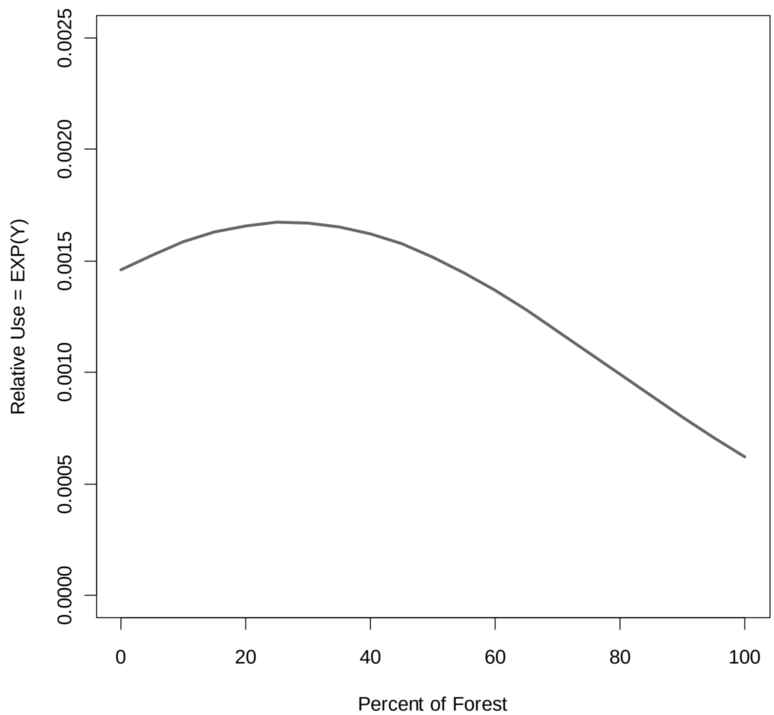
(e) Plotted parameter values from Appendix E, April 2010



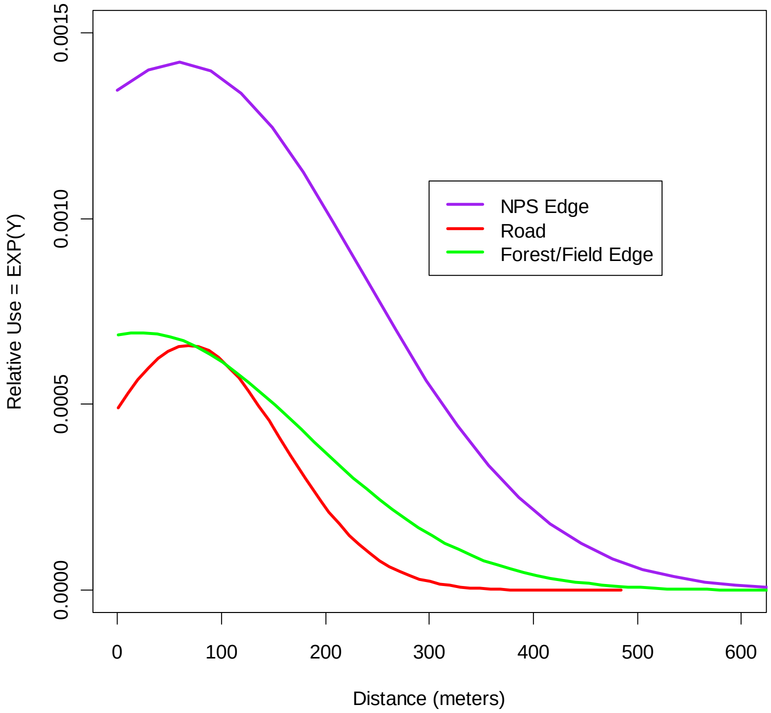


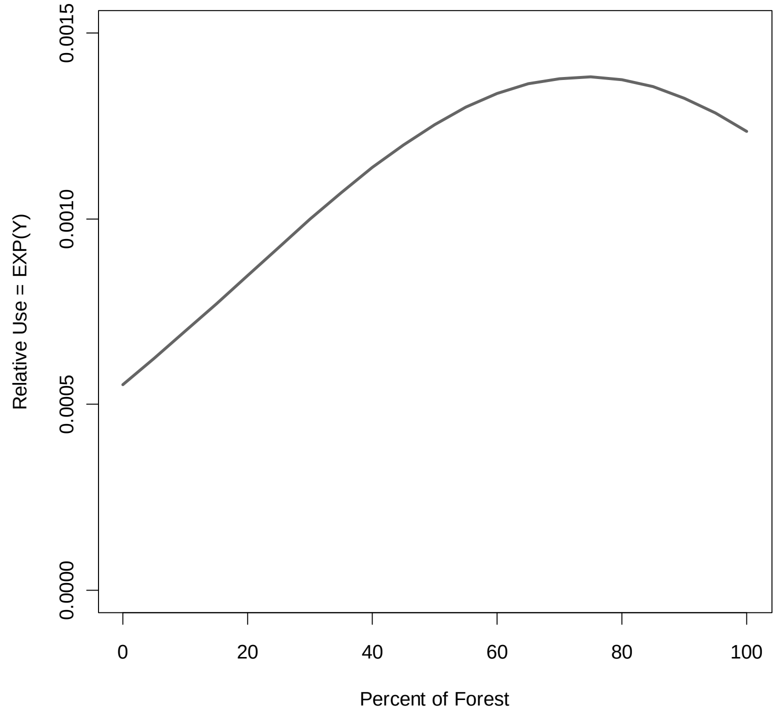
(f) Plotted parameter values from Appendix E, August 2010





(g) Plotted parameter values from Appendix E, November 2010





APPENDIX G: Distance Sampling Observations

(a) Summary of observations in the 250 m survey zone for each survey, Gettysburg, Pennsylvania, 2009-2010. From left to right: The actual number of deer seen in fields and forests, the number of groups seen in fields (n_{Field}) and forests (n_{Forest}), the number of complete rounds of all transects, the average number of groups seen in fields per round ($\overline{n}_{Field}$), and the average number of groups seen in forests per round ($\overline{n}_{Forest}$).

Survey	$Deer_{(Field)}$	$Deer_{(Forest)}$	n_{Field}	n_{Forest}	# Rounds	$\overline{n}_{Field}$	$\overline{n}_{Forest}$
Apr 2009	239	122	64	42	3	21	14
Aug 2009	323	23	131	12	3	44	4
Nov 2009	246	65	112	31	2.8	40	11
Jan 2010	270	62	78	21	2.5	31	8
Apr 2010	257	60	66	24	3	22	8
Aug 2010	480	24	154	13	3	51	4
Nov 2010	288	103	127	57	3	42	19

(b) Summary of observations in the 80 m survey zone for each survey, Gettysburg, Pennsylvania, 2009-2010. From left to right: The actual number of deer seen in fields and forests, the number of groups seen in fields (n_{Field}) and forests (n_{Forest}), the number of complete rounds of all transects, the average number of groups seen in fields per round ($\overline{n}_{Field}$), and the average number of groups seen in forests per round ($\overline{n}_{Forest}$).

Survey	$Deer_{(Field)}$	$Deer_{(Forest)}$	n_{Field}	n_{Forest}	# Rounds	$\overline{n}_{Field}$	$\overline{n}_{Forest}$
Apr 2009	89	107	24	35	3	8	12
Aug 2009	169	20	72	11	3	24	4
Nov 2009	139	65	67	31	2.8	24	11
Jan 2010	82	48	28	15	2.5	11	6
Apr 2010	127	60	32	24	3	11	8
Aug 2010	225	24	81	13	3	27	4
Nov 2010	153	97	69	54	3	23	18

APPENDIX H: Distance Sampling Mean Cluster Sizes

(a) Summary of mean cluster sizes for fields and forests in the 250 m survey zone for each survey, Gettysburg, Pennsylvania, 2009-2010. Standard errors were calculated using non-parametric bootstrapping.

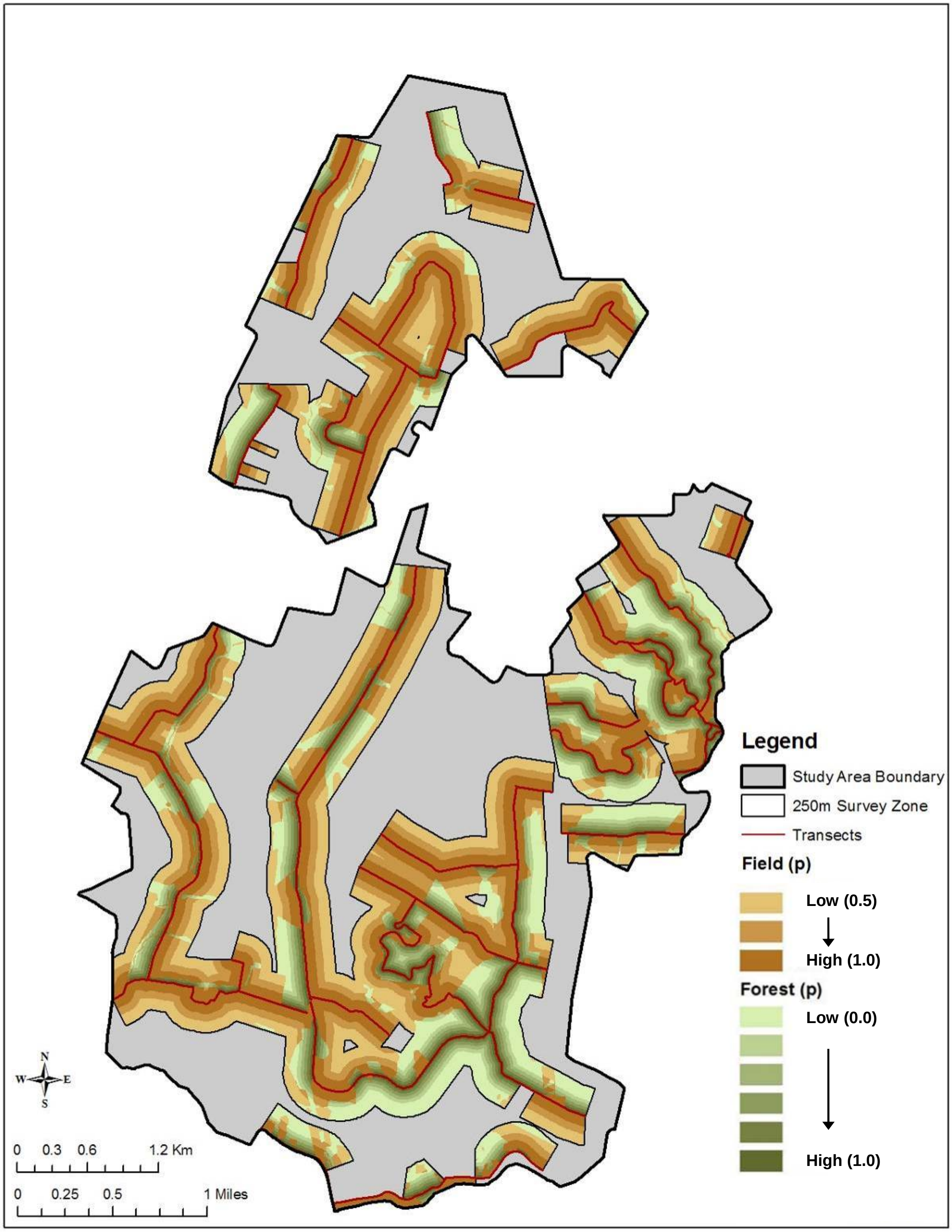
Survey	$E(S)_{Field}$	$SE[E(S)_{Field}]$	$E(S)_{Forest}$	$SE[E(S)_{Forest}]$
Apr 2009	3.7	0.35	2.9	0.22
Aug 2009	2.5	0.15	1.9	0.30
Nov 2009	2.2	0.14	2.1	0.16
Jan 2010	3.5	0.29	3.0	0.33
Apr 2010	3.9	0.39	2.5	0.39
Aug 2010	3.1	0.17	1.9	0.28
Nov 2010	2.3	0.11	1.8	0.14

(b) Summary of mean cluster sizes for fields and forests in the 80 m survey zone for each survey, Gettysburg, Pennsylvania, 2009-2010. Standard errors were calculated using non-parametric bootstrapping.

Survey	$E(S)_{Field}$	$SE[E(S)_{Field}]$	$E(S)_{Forest}$	$SE[E(S)_{Forest}]$
Apr 2009	3.7	0.66	3.1	0.25
Aug 2009	2.4	0.20	1.8	0.31
Nov 2009	2.1	0.16	2.1	0.16
Jan 2010	2.9	0.38	3.2	0.39
Apr 2010	4.0	0.67	2.5	0.37
Aug 2010	2.8	0.21	1.9	0.28
Nov 2010	2.2	0.15	1.8	0.14

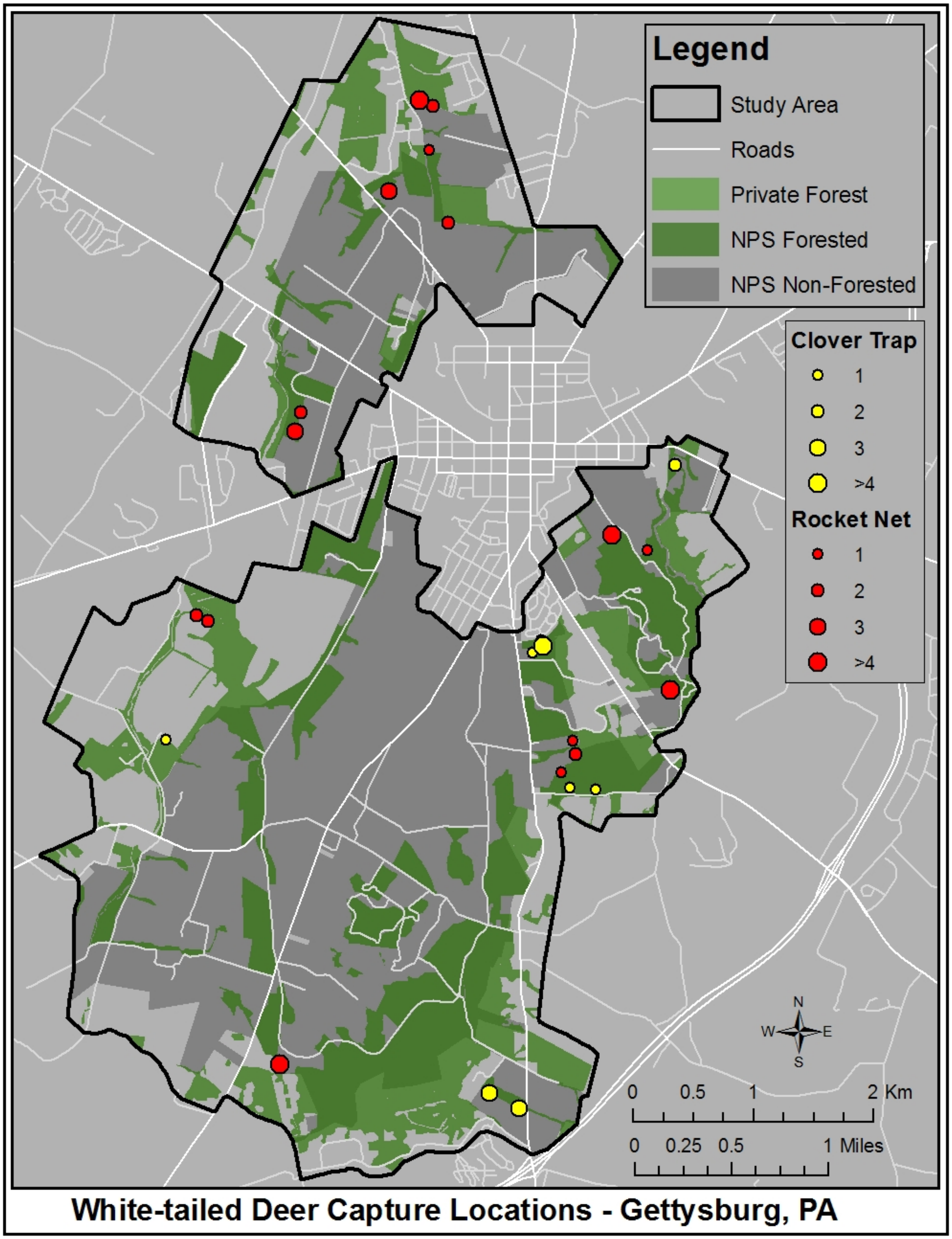
APPENDIX I: Detection Probability Calculation Grid

Shown for January 2010 survey, where p = detection probability for fields and forests.



APPENDIX J: Capture Locations of White-tailed Deer

Number of deer captured at successful clover trap and rocket net sites by symbol size.



APPENDIX K: Example R Code for Zero-inflated Negative Binomial Model

Below is R code for model 1 for one deer, using GPS locations for that one deer during the April 2010 distance sampling survey.

```
>df3<-read.table("April_10_Input_File.txt", na.strings="NA",header=T)

# Center each covariate so that 0 is now the grand mean
>df3$dist_road.c<-(df3$Dist_Road.m-mean(df3$Dist_Road.m))
>df3$dist_NPS_edge.c<-(df3$Dist_NPS_Edge.m-mean(df3$Dist_NPS_Edge.m))
>df3$dist_forest_field_edge.c<-(df3$Dist_Forest_Field_Edge.m-
mean(df3$Dist_Forest_Field_Edge.m))

# Normalize each covariate by dividing by its standard deviation
>df3$norm_dist_road.c<-(df3$dist_road.c/sd(df3$dist_road.c))
>df3$norm_dist_NPS_edge.c<-
(df3$dist_NPS_edge.c/sd(df3$dist_NPS_edge.c))
>df3$norm_dist_forest_field_edge.c<-
(df3$dist_forest_field_edge.c/sd(df3$dist_forest_field_edge.c))

# Divide covariate Percent Forest by 100 (constrain from 0-1)
>df3$Per_Forest<-(df3$Per_Forest/100)

# Subset data for each deer (shown here for deer with low tag number 7)
>df.LT7<-subset(df3, df3$LT=="7")

# Create the offset term for each deer (shown here for deer with LT 7)
>offset.LT7<-rep(log(sum(df.LT7$GPS_Locs_April_10)), times=3000)

# Model 1 - Percent Forest + Distance to Forest-field edge
>library(pscl)

>m1.7<-zeroinfl(GPS_Locs_April_10~1+Percent_Forest+
norm_dist_forest_field_edge.c|1,subset(df3,df3$LT=="7"),dist="negbin",
offset=offset.LT7)

>summary(m1.7)
Pearson residuals:
    Min      1Q  Median      3Q     Max
-0.1277 -0.1254 -0.1198 -0.1170 38.7838

Count model coefficients (negbin with log link):
                              Estimate Std. Error z value Pr(>|z|)
(Intercept)                    -5.9978     0.3443 -17.418  < 2e-16 ***
Per_Forest                      1.8523     0.3813   4.858 1.19e-06 ***
norm_dist_forest_field_edge.c  -0.3688     0.2226  -1.657   0.0975 .
Log(theta)                     -0.2663     0.3874  -0.687   0.4919

Zero-inflation model coefficients (binomial with logit link):
            Estimate Std. Error z value Pr(>|z|)
(Intercept)   3.2383     0.1896   17.07   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Theta = 0.7662
Number of iterations in BFGS optimization: 60
Log-likelihood: -628.9 on 5 Df
```