# Large Carnivores under Threat:

# Investigating Human Impacts on Large Carnivores in East Africa

**Andrew Paul Jacobson**

**PhD Thesis**

**University College London (UCL)**
**and**
**Institute of Zoology, Zoological Society of London (ZSL)**

**2017**

# Declaration

**I, Andrew Paul Jacobson, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.**

# Abstract

Large carnivores are a polarizing group of species that play an outsized role in relation to their number. They structure ecosystems and feature prominently in human culture. Yet, their place in a rapidly changing world is uncertain. The large carnivore guild in the five countries of East Africa, Burundi, Kenya, Rwanda, Tanzania, and Uganda, is largely intact; however, expanding human populations pose a substantial threat. Interventions are necessary to promote coexistence. To accomplish this, more accurate identification of threats, and improved understanding of species' responses are needed. Primary threats to large carnivores in the region include habitat loss and human-wildlife conflict (HWC). Problematically, identification of human impacted areas from earth observation data can be difficult in heterogeneous savannah habitat, much of East Africa. I create a tool that enables land cover classification using Google Earth's high-resolution imagery. With this tool I develop a data set of human impacted areas for East Africa. To ascertain carnivore response to human dominated lands, I use correlative species distribution modeling (SDM). Yet, there is no clear consensus on proper methods for generating pseudo-absence (PsA) data in these models. I review some existing methods in the context of their ecological meaning, and propose new PsA selection strategies. I then apply two novel and one existing PsA strategy to assess four carnivores' (cheetah, wild dog, leopard, and lion) responses to human land cover and human population densities. Results suggest these carnivores are more susceptible to human land cover than human populations. Finally, I consider existing approaches of using SDM with HWC records to generate spatial risk maps with the goal of alleviating conflict. I draw on the SDM literature to highlight and demonstrate how two commonly overlooked issues in spatial risk modeling can hamper generating useful conclusions. In sum, these efforts represent attempts at improving commonly used methods used to study wildlife distribution and threats, and can be widely applied to other species and systems.

# Dedication

*To Sawyer*

*that he too may one day be able to pursue his childhood passions*

## Acknowledgements

A substantial number of people, and organizations have supported this work and I thank you all for your contributions to this thesis. It would not have been possible without you.

My first thanks and acknowledgements go to my three supervisors, Sarah Durant and Nathalie Pettorelli at the Institute of Zoology, ZSL, and Mat Disney at UCL, for their tireless support and guidance. It is not an overstatement to say that each one has consistently suggested particular improvements to my research and writing habits from Day One until today. I carry their valuable reminders with me. Each has imparted distinct knowledge and I am undoubtedly a better researcher because of my association with them. Mat deserves extra recognition for unscrambling my ecological ramblings and pointing me in useful directions. I owe extra gratitude to Nathalie who provided advice, guidance and reviews with amazing speed and seeming effortlessness. I would also like to particularly thank Sarah for initiating this project. She has been a wonderful mentor and demonstrated how to be an effective carnivore and conservation biologist simultaneously. She did not know when this project started that cheetahs were my favorite animals, and I continue to feel privileged to work beside her to ensure their continued existence. Rosie Woodroffe at ZSL and Philip Lewis at UCL also deserve credit.

I owe tremendous thanks to several professors at my second home, Duke University. I could not have finished this dissertation without Stuart Pimm; quite simply I would not have had a place to sit! Stuart has been more than just generous in including me in his academic family, he provided research guidance and support as well. He continues to impress me with the speed and insightfulness of his responses. Perhaps I owe this thesis to no one more than Luke Dollar. Without his support and encouragement as the Program Director at National Geographic Society's Big Cats Initiative, I never would have pursued carnivore research or a doctorate. Luke is blessed with many talents and works continuously (literally around the clock) to share his talents and knowledge with others. His overflowing and generous spirit has given me opportunities that continue to define who I am. Thirdly, this thesis undeniably bears Dean Urban's imprint. Dean has taught me nearly everything I know (and frustrating what we collectively don't know) about species distribution modeling. Much of the conceptual background of this thesis came out of long discussions with him. That I have no official affiliation with any of these professors only heightens my gratitude for their generosity and support. Finally, I owe the staff at the Nicholas School of the Environment, as well as at ZSL and UCL, a heartfelt thank you for their assistance.

I gratefully acknowledge financial support from the Howard G. Buffett Foundation via the Big Cats Initiative at National Geographic, and the Mohamed bin Zayed Species Conservation Fund.

Split between countries and institutions, I owe thanks to many friends and colleagues across two continents. In the US, I would like to give extra recognition to Jason Riggio, Lisanne Petracca, and members of the Pimm lab. I am blessed to have such a constructive and supportive lab. In the UK, I especially thank Tammy Davies, Judith Reise, Monika Böhm, and Claire Duncan for welcoming me to London and British culture.

Several people and organizations deserve credit for successful field visits in Kenya and Tanzania. Most importantly, I am indebted to Helen O'Neill and Stefanie Strebel for graciously hosting me in Laikipia, as well as the Mpala Research Centre. I also kindly thank Ingela Jansson who squeezed me in to her unbelievable schedule. Staffs at the Ruaha Carnivore Project, and the Tanzania Wildlife Research Institute also deserve thanks.

Special recognition go to Laly Lichtenfeld at the Tanzania People and Wildlife Fund, Amy Dickman at the Ruaha Carnivore Project, Shivani Bhalla at Ewaso Lions, and Alayne Cotterill. I am indebted to them not only for their substantial data contributions, but also for their support, encouragement and engagement with my research. I am inspired by their passion and commitment. Other data contributors are also greatly appreciated, as are those who assisted in conducting manual land cover classifications using GE Grids.

Finally and most importantly, I thank my wife Hannah and the rest of my family. During the struggle and dog days of the PhD, their love and encouragement kept me going. Hannah has been steadfast in her love and abided more ramblings and rants about distribution modeling and carnivores than either of us wanted. She has been my unwavering star, undimmed in even the darkest nights. My parents Stan and Janet taught me to strive, to follow my passions, and how to persevere through adversity. Rolf, Steve and Dixie also provided emotional support and are important pillars in my life. And the last thanks goes to the least (or at least smallest) of us, my son Sawyer. His joy, love and spirit are an endless well that I have constantly returned to during 2016.

# List of acronyms

| | |
|---|---|
| IUCN | International Union for the Conservation of Nature |
| GE | Google Earth |
| GR | Game reserve |
| HPD | Human population density |
| HWC | Human-wildlife conflict |
| LULC | Land use land cover |
| NP | National park |
| NR | National reserve |
| NCA | Ngorongoro Conservation Area |
| NDVI | Normalized difference vegetation index |
| PsA | Pseudo-absence data |
| PA | Protected area |
| RF | Random forest |
| SDM | Species distribution modeling |
| TRI | Terrain ruggedness index |

Directions and distances

| | |
|---|---|
| km | Kilometer |
| m | Meter |
| NE | Northeast |
| NW | Northwest |
| SE | Southeast |
| SW | Southwest |

## List of common and scientific names

| | |
|---|---|
| Cheetah | *Acinonyx jubatus* |
| Impala | *Aepyceros melampus* |
| Springbok | *Antidorcas marsupialis* |
| Pronghorn antelope | *Antilocapra americana* |
| Coyote | *Canis latrans* |
| Black-backed jackal | *Canis mesomelas* |
| Elk | *Cervus canadensis* |
| Gray wolf | *Canis lupus* |
| Maned wolf | *Chrysocyon brachyurus* |
| Civet | *Civettictis civetta* |
| Wildebeest | *Connochaetes taurinus* |
| Spotted hyena | *Crocuta crocuta* |
| Thomson’s gazelle | *Eudorcas thomsonii* |
| Giraffe | *Giraffa camelopardalis* |
| Mountain gorilla | *Gorilla beringei* |
| Wolverine | *Gulo gulo* |
| Brown hyena | *Hyaena brunnea* |
| Striped hyena | *Hyaena hyaena* |
| African elephant | *Loxodonta africana* |
| African wild dog | *Lycaon pictus* |
| European badger | *Meles meles* |
| Striped skunk | *Mephitis mephitis* |
| White-tailed deer | *Odocoileus virginianus* |
| Chimpanzee | *Pan trogolodytes* |
| Lion | *Panthera leo* |
| Jaguar | *Panthera onca* |
| Leopard | *Panthera pardus* |
| Tiger | *Panthera tigris* |
| Snow leopard | *Panthera uncia* |
| African buffalo | *Syncerus caffer* |
| Common eland | *Taurotragus oryx* |
| Asiatic black bear | *Ursus thibetanus* |
| Brown bear | *Ursus arctos* |
| Red fox | *Vulpes vulpes* |

## Table of contents

## List of tables

## List of figures

# Chapter 1

## Large carnivores in a changing world: a general introduction

A lioness regards the briefly lush landscape of northern Kenya. Photo courtesy of Ewaso Lions.

# Chapter 1 Large carnivores in a changing world: a general introduction

## 1.1 Carnivores in context

Carnivores conjure powerful imagery and feelings in humans. They are some of the most evocative of all groups of species and figure prominently in the heraldry and literature of many cultures (Kruuk 2002). Owing to their wide-ranging and predatory behaviour, their interests often come into conflict with our own (Linnell, Swenson, and Andersen 2001). This has contributed to a history of antagonism and persecution against carnivores. As a group, carnivores are threatened, and some species are critically endangered (Ripple et al. 2014). With some exceptions, humans increasingly prescribe carnivores' distribution. Yet, as carnivores are being lost, their ecological importance is increasingly recognized. Predation is a structuring element in the theory of evolution by natural selection, and indeed the entire field of ecology (Holt 1977; Sih et al. 1985; Holt and Polis 1997; Sih, Englund, and Wooster 1998). In addition, large carnivores are crucial for the maintenance of biodiversity and ecosystem function; trophic cascades are documented for seven of the 31 largest terrestrial carnivores (Ripple et al. 2014).

In colloquial terms, carnivore means meat-eater. However, in this thesis, I will focus on all species in the mammalian order Carnivora but exclude the marine members (Figure 1.1). The remaining 245 species in Carnivora naturally occupy all continents except Antarctica and Australia (although they were introduced on the latter continent some 5,000 years ago). Evidence suggests the historical order was larger, more diverse and more fearsome (Dalerum et al. 2009). Extant Carnivora are a monophyletic group, descending from a single Miacidae ancestor (Bininda-Emonds, Gittleman, and Purvis 1999). Typically, they are at or near the top of the food chain and thus are few in number or at low densities. Large carnivores (those with adult body mass >20 kg; see Dalerum et al. 2008) especially tend to have significant energetic constraints, slow life histories and may traverse entire landscapes in search of prey (Ripple et al. 2014).

**Figure 1.1 Some terrestrial members of the order Carnivora, from top left to bottom right they are: tiger, brown bear, spotted hyena, Eurasian wolf, civet, wolverine (photo courtesy of Wikimedia Commons).**

From an ecological perspective, carnivores regulate ecosystems and maintain biodiversity through "top-down" effects (Elton 1927; Terborgh et al. 2001; Ripple et al. 2014). Through fear of predation, carnivores can exert physical and psychological impacts on herbivores and secondarily on plants. For instance, grey wolves can limit the population size and elicit strong behavioural changes in cervids (like elk). Therefore wolves influence can elk, who in turn impact the plant community, ultimately influencing stream bank erosion, and even herpetofauna and invertebrate abundance and richness (Ripple and Beschta 2004; Beschta and Ripple 2009). Trophic cascades can operate in other systems as well. In savannah ecosystems of Africa, large carnivores can influence tree community composition (Ford et al. 2014). Poorly-defended tree species (i.e. those with fewer thorns) were less common in areas of low-predation risk as herbivory was more intense whereas they were more prevalent in areas of high-predation risk (Ford et al. 2014). Thus, through top-down effects carnivores can impact ecosystem processes such as carbon storage, stream morphology, crop production, and disease dynamics (Ripple et al. 2014). Yet, the impact of carnivores in human-dominated landscapes is less certain, as human presence can attenuate their ecological effects (Kuijper et al. 2016).

Effects of carnivores are not limited to herbivores; carnivores influence each other as well (Polis, Myers, and Holt 1989; Palomares and Caro 1999; Linnell and Strand 2000). Carnivores can influence other carnivore population dynamics, constrain habitat selection, reduce prey encounter rates, and limit spatial distribution (Creel, Spong, and Creel 2001). The loss of apex predators can lead to "mesopredator release," an increase in the abundance, distribution and/or niche space of smaller carnivores (Crooks and Soulé 1999). The increase in abundance of mesopredators can have ripple effects throughout the ecosystem. For instance, grey wolves are apex predators and can displace coyotes in North America (Berger and Gese 2007; Berger, Gese, and Berger 2008). While wolves frequently take adult ungulates, coyotes preferentially predate on pronghorn antelope fawns. Locations with high wolf densities had lower coyote densities and increased pronghorn survival rates (Berger and Gese 2007; Berger, Gese, and Berger 2008). However, wolves will also mate and hybridize with coyotes, creating a complex relationship between the two species (Lehman et al. 1991).

From a human viewpoint, carnivores are possibly the most polarizing group of all species, either loved or loathed (Kruuk 2002). Carnivores can be big, bold, intelligent, and social; attributes loved by many humans. They also embody strength, ferocity, and resilience. They feature prominently in our cultures, embedded in our heraldry, songs, stories, and traditions (Kruuk 2002). However, their cultural references are not always positive; carnivores are also associated with evil and destruction and often treated with hatred and disdain (Kellert et al. 1996). Carnivores may be killed without mercy, trapped, poisoned, gassed or gunned down from helicopters. Justification for the removal of carnivores, if needed, is the protection of species valuable to us, or fear of physical injury or death (Treves and Karanth 2003).

Beyond their cultural influence, carnivores play important roles in socio-economics. Carnivores are a top reason to visit a protected area (Lindsey et al. 2007; Okello, Manka, and D'Amour 2008; Grünewald, Schleuning, and Böhning-Gaese 2016), and are favorites at zoos (Carvell et al. 1998). They are some of the most expensive sport hunting packages in Africa and some of the most sought after trophies (Baldus and Cauldwell 2004; Lindsey et al. 2006; Lindsey et al. 2012).

So carnivores, and in particular larger carnivores, are incredibly important taxa in both their anthropological and ecological roles. However, human and carnivore interests often intersect, and often violently.

## 1.2 Threats to carnivore persistence

Rapidly expanding human populations and increasing per capita consumption are undoubtedly the ultimate factors in biodiversity decline (Pimm et al. 2014; Ripple et al. 2014). These drive expansions of agriculture and livestock, as well as energy use, and infrastructure projects. Climate change is already affecting biodiversity and carnivores, and will become more important in the future (Thomas et al. 2004; Loucks et al. 2010; Ogutu et al. 2016). Crop and pasturelands have exceeded 40% of the terrestrial surface area (Foley et al. 2005), so habitat loss and fragmentation are leading threats to biodiversity. However, additional factors such as poverty, poor governance and corruption, have the capacity to ameliorate or worsen human population growth impacts on biodiversity (Nelson, Nshala, and Rodgers 2007; Caro and Davenport 2015; Ogutu et al. 2016).

Yet, the more proximate causes of carnivores decline are threefold: habitat loss and fragmentation, prey loss, and direct killing/utilization. Other threats that may be locally important or for particular species include disease, genetic poverty, tourism, and road/railroad kills (Ray, Hunter, and Zigouris 2005). These threats will not be discussed here in order to focus only on the top three threats to most large carnivores.

### 1.2.1 Habitat loss and fragmentation

Habitat loss is a primary agent in the current extinction crisis and has consistent, negative impacts on biodiversity (Vitousek et al. 1997; Fahrig 2003). Habitat loss is a direct change in the composition of the elements of landscape which cause it to become unsuitable (Reid, Thornton, and Kruska 2004). It can be immediate, but it can also proceed from the cumulative effects of habitat degradation. Habitat loss is a scale and species dependent issue (Reid, Thornton, and Kruska 2004). Habitat for a dung beetle is likely to be very different from the habitat requirements for a cow or a lion for instance. It can also be difficult to link specific mortality events with habitat loss, as the proximate cause of death may be something else, such as starvation or road mortality. Regardless, habitat loss has serious detrimental implications for a variety of direct and indirect biodiversity factors. It can affect species richness, abundance, and diversity as well a populations growth rate (Fahrig 2003; Reid, Thornton, and Kruska 2004). Loss can also alter species interactions, affect breeding success, modify dispersal success, and impact foraging success rates (Fahrig 2003). For these reasons, habitat loss needs to stay above an extinction threshold, the minimum amount of habitat required for a population of a certain species to persist in a landscape (Fahrig 2002).

A recent review of seven large, up to 35 years-long fragmentation projects reveal fragmentation's clear negative impacts on biodiversity, and the structure and function of the remaining fragments (Haddad et al. 2015). While some argue that the configuration (i.e. spatial pattern) of habitat needs to be understood separately from fragmentation's impact on area (i.e. amount of habitat) (Fahrig 2003), I follow the definition that fragmentation "is the division of habitat into smaller and more isolated fragments separated by a matrix of human-transformed land cover" (Haddad et al. 2015). Fragmentation impacts biodiversity via three primary methods, a reduction in fragment area, increased isolation, and greater edge length (Lindenmayer and Fischer 2013; Haddad et al. 2015). Fragmentation can also have delayed effects, including extinction debt and ecosystem function debt (Haddad et al. 2015). Therefore, fragmentation degrades ecosystems with impacts on species persistence and richness, nutrient cycling, and trophic dynamics, and these impacts can accumulate over time (Haddad et al. 2015).

There are multiple ways in which habitat loss and fragmentation threaten carnivores specifically, and together they are implicated as a major factor in the decline of many carnivores (Nowell and Jackson 1996). For instance, loss of habitat may reduce prey abundance or alter prey composition (Oehler and Litvaitis 1996). Fragmentation also results in greater contact between carnivores and humans, or domestic animals. Contact increases opportunities for human-carnivore conflict and disease transmission to carnivores (e.g. canine distemper virus affecting lions and other carnivores in Serengeti National Park (NP) in 1994) (Roelke-Parker et al. 1996). Impact may be species-specific (Crooks 2002) but is particularly problematic for wide-ranging species such as the African wild dog and cheetah (Woodroffe and Ginsberg 1998). Thus, habitat loss, fragmentation and degradation negatively affect carnivores and can directly reduce their survival rates.

### 1.2.2 Prey loss

Nearly all members of Carnivora eat meat, and many large terrestrial carnivores are hyper-carnivores, aka obligate meat-eaters, and derive at least 70% of their diet from meat (Van Valkenburgh 2007). Carnivores tend to be limited by food resources (Carbone and Gittleman 2002). In some species, there is a tight relationship between food resources and carnivore densities, such as for the lion (van Orsdol, Hanby, and Bygott

1985; Loveridge and Canney 2009), tiger (Karanth et al. 2004), cheetah (Laurenson 1995; Gros, Kelly, and Caro 1996), and leopard (Stander et al. 1997). That is, large carnivore biomass is strongly correlated with prey biomass. Therefore, the loss of prey represents a clear and present threat to carnivores.

Reduction in prey biomass affects overall carnivore abundance and density, primarily via impacts on reproduction, and mortality (Fuller and Sievert 2001). For instance, age of first reproduction, pregnancy rates, and litter sizes could all be negatively affected by decreased food resources (Fuller and Sievert 2001). Of course, survival of all age classes may also decrease as a consequence of food shortages. The youngest age class may be particularly vulnerable (Fuller and Sievert 2001).

Prey loss is a widely recognized issue for carnivores, but may be particularly severe in Africa, Asia and South America (Ripple et al. 2014). Several large carnivores are recognized as having >50% of their prey classified as declining and threatened (Wolf and Ripple 2016). In particular, wildlife populations are declining rapidly across much of Africa (Newmark 1996; Western, Groom, and Worden 2009; Bouché et al. 2012; Durant et al. 2014; Ogutu et al. 2016). Threats for herbivores are often the same as for carnivores, such as habitat loss and disease (Wolf and Ripple 2016). However, bushmeat hunting, the illegal acquisition and trade of wild meat, more seriously threatens herbivore populations (Lindsey et al. 2013b; Ripple et al. 2016). While carnivores are typically not targeted in bushmeat hunting, the rate of accidental snaring and capture can still have substantial population-level effects (Becker et al. 2013A).

### 1.2.3 Direct killing and utilization

A third primary threat to carnivores is killing and utilization, focused in sub-Saharan Africa and Asia (Ripple et al. 2014). Humans have used carnivore meat, fur and other body parts for ages. Possibly the most widespread and important use of carnivores is for their fur; indeed, some are even termed "fur-bearers" (Kruuk 2002). The scale of this trade is immense; in 1995 the fur retail trade was ~$6 billion in the European Union and $1.2 billion in the US (Kruuk 2002). Until recently, the vast majority of fur was from wild populations, but this has now almost completely transitioned to captive animals (Sillero-Zubiri and Laurenson 2001). This transition is beneficial as the fur trade harvest can have a serious detrimental impact on wild carnivore populations (Sillero-Zubiri and Laurenson 2001). For large carnivores, the fur trade was and continues to be focused more exclusively on spotted cats, such as the leopard (Nowell and Jackson 1996). But carnivores are killed for many other reasons, including for their meat, bones, glands, and with a diversity of end uses including fashion, perfume, and medicine (Kruuk 2002).

The trafficking of carnivore parts for medicine, such as use of the tiger and lion in East Asian medicines (Nowell and Jackson 1996; Bauer et al. 2015), is part of a rapidly expanding illegal wildlife trade. The global trade in illegal wildlife and wildlife products was recently calculated at between $7.8 and $10 billion annually, making it the 4th largest global illegal trade (Haken 2011; World Bank Group 2016). In response to this, international donors spent over $1.3 billion to combat illegal wildlife trade since 2010. The majority of this was spent in sub-Saharan Africa and primarily to prevent poaching (World Bank Group 2016).

Finally, legal killing via problem-animal control and sport hunting can also be substantial causes of mortality. In problem-animal control, individual animals may be targeted for elimination (or removal) by authorities if they are identified as causing conflict. While carnivores are often legally protected around the world, there are exceptions when the animal threatens life, livestock or property. In that instance, a person can kill the responsible carnivore on the spot. Yet, unless the individual(s) involved are neutralized immediately, it can be difficult to accurately identify and kill the offending individual(s) at a later date. The impact of problem animal control on carnivore mortality at larger extents is unknown but it can limit populations at a local level (Woodroffe and Frank 2005; Balme, Slotow, and Hunter 2009). Translocation of problem animals is also commonly practiced in some parts of the world, although without extremely rigorous standards in place, the practice is rarely effective at reducing conflict and often leads to the individual's death (Fontúrbel and Simonetti 2011; Weise et al. 2015).

Sport hunting for carnivores is big business, generating substantial economic revenue (Lindsey, Roulet, and Romanach 2007; Lindsey et al. 2012). In Africa, trophy hunting of all species generates gross revenue of >$200 million per year (Lindsey, Roulet, and Romanach 2007). Hunting of large carnivores, like lion and leopard, comprise some of the most popular sport hunting packages (Baldus and Cauldwell 2004; Lindsey et al. 2006; Lindsey et al. 2012) and lions attract nearly the highest mean price of all trophy species, from $24,000 – $71,000 (Lindsey et al. 2012). Despite or indeed because of these economic incentives, sport hunting is often poorly regulated, conducted without knowing the sustainability of the harvest, and importantly, can lead to population decline (Lindsey, Roulet, and Romanach 2007; Packer et al. 2010; Hunter et al. 2012).

## 1.3 Carnivore and human interactions

### 1.3.1 Impacts on humans

Carnivores generate widespread human-wildlife conflict (HWC). Thirgood et al. (2005) reviewed HWC more broadly and identified five of the most common forms of conflict. Four of them are more relevant to carnivores: predation upon livestock, predation upon game species, attacks on humans, and disease transmission. Each type of conflict has undoubtedly existed for millennia, and will be briefly examined.

Predation on livestock or other domestic animals is pervasive where large carnivores and livestock intersect. Livestock have attenuated threat responses (Shrader et al. 2008) and hence can be particularly easy for carnivores to kill. Some domestic animals, like dogs, may even be preferentially preyed upon by large felids (Athreya et al. 2014). Thus, it is not surprising that predation on livestock is the most common form of conflict with carnivores and that it occurs throughout the world (Sillero-Zubiri and Laurenson 2001; Thirgood, Woodroffe, and Rabinowitz 2005). A wide range of carnivores is involved targeting a range of livestock including maned wolf attacks on poultry (Consorte-McCrea 2013), wolverine attacks on reindeer and sheep (Landa et al. 1997), Asiatic black bear attacks on apiaries (Huygens and Hayashi 1999) and snow leopard predation on yaks and horses (Oli, Taylor, and Rogers 1994). Predation pressures may be intense; Nepalese villagers blamed 63% of livestock losses on carnivores (Jackson et al. 1996). Livestock may also compose a substantial portion of a predator's diet; domestic livestock accounted for 70% of prey biomass for snow leopards in Pakistan (Anwar et al. 2011). Yet, carnivore predation typically causes from <1% to 10% annual losses to livestock, and often less than disease (Thirgood, Woodroffe, and Rabinowitz 2005). Despite this, predation can cause significant cultural and economic repercussions (Barua, Bhagwat, and Jadhav 2013; Dickman et al. 2014). In particular, surplus killing, where more livestock are killed than the predator can eat, can devastate a herd and stir animosity (Kruuk 1972). Yet, generally speaking, carnivores do not prefer livestock and attack them at rates less than predicted according to their abundance on the landscape (Odden et al. 2008).

Predation on game is another form of conflict albeit one that is harder to fault the carnivore. Humans hunt many species across the world, from large herbivores to small birds. These species are components of a predator's diet but are also valued by humans as food, medicine, status symbols, for sport, or other reasons. Hunters desire these species to be plentiful, healthy and in good condition, and for other attributes like antler size. Therefore, hunters blame carnivores if they feel that carnivores keep herbivore populations or just the desired population subset low. In the UK, humans exterminated several carnivore species in royal hunting reserves to protect deer populations for hunting (Sillero-Zubiri and Laurenson 2001). Recent evidence has suggested that predation on ungulates keeps ungulate abundance lower than when predation is absent, and hence reduces the number of ungulates available for hunting (Gasaway et al. 1992; Thirgood, Woodroffe, and Rabinowitz 2005). There is also evidence that smaller generalist carnivores can reduce breeding success and breeding densities of birds and lagomorphs, and in turn limit the numbers available for hunting (Marcstrom, Kenward, and Engren 1988; Tapper, Potts, and Brockless 1996; Thirgood, Woodroffe, and Rabinowitz 2005). Thus, it is clear that the presence of carnivores can negatively impact hunting objectives.

A variety of species of carnivores attack humans throughout the world (Löe and Röskaft 2004; Quigley and Herrero 2005). Overall, attacks by animals comprise a tiny percentage of human deaths (Löe 2002). Human deaths from carnivores seem to have declined substantially in the 20th century relative to earlier time periods (John D C Linnell et al. 2002). Yet, the overall number of deaths by large carnivores in the 20th century is >15,000, and is probably an underestimate (Löe and Röskaft 2004). Tigers, leopards, wolves, lions and brown bears were the only species to kill >100 people, with tigers responsible for the most deaths by far (Löe and Röskaft 2004). For obvious reasons, attacks on humans are particularly significant and unlike any other form of conflict. Both adults and children may be killed. Not all attacks are predatory; some are defensive or due to rabies (Linnell et al. 2002). Regardless of the reason, attacks are traumatic events, can inspire widespread fear among a community, and spur spontaneous carnivore hunting parties (Jhala and Sharma 1997).

Disease transmission is the fourth primary form of conflict caused by carnivores. Carnivore species can be hosts and transmitters of disease that ultimately threaten or kill livestock, domestic animals, and humans (Thirgood, Woodroffe, and Rabinowitz 2005). Rabies is a classic example. Many carnivores are proven hosts for rabies such as the red fox, striped skunk, and black-backed jackal. The World Health Organization has estimated that rabies kills 50,000 people a year, and up to 10,000 in India alone (WHO 1998). Of course, rabies also threatens domestic animals and livestock. Macdonald (1980) estimated that rabies costs the livestock industry in Latin America $400 million a year. Other examples include canids that are reservoirs of the tapeworm *Echinococcus multilocularis*, and felids that transmit the protozoan *Toxoplasma gondii* (Thirgood, Woodroffe, and Rabinowitz 2005). A hotly contested current issue in the UK is the role of

European badger transmission of tuberculosis to cattle. Over the course of a decade, >310,000 cattle were slaughtered due to the disease and costs to the British taxpayer exceeded £500 million (DEFRA 2014). Widespread badger culls are being implemented by the government, despite a lack of scientific evidence demonstrating a likely benefit to disease control in cattle (DEFRA 2014).

These descriptions of four types of HWC focus on the direct costs of wildlife on humans, however, wildlife may also impose substantial indirect costs (Norton-Griffiths and Southey 1995; Thirgood, Woodroffe, and Rabinowitz 2005; Barua, Bhagwat, and Jadhav 2013). Indirect costs include opportunity costs of time, money, and resources spent attempting to prevent wildlife damage. For instance, time spent protecting livestock can limit time spent at school or harvesting crops (Norton-Griffiths and Southey 1995). In addition, there are other hidden costs of HWC that can include diminished psychosocial wellbeing, disruption of livelihoods, and food insecurity that are typically poorly understood or quantified (Barua, Bhagwat, and Jadhav 2013).

HWC threatens the lives and livelihoods of millions of people throughout the world but also threatens wildlife persistence and undermines conservation actions (Sillero-Zubiri, Sukumar, and Treves 2007). Conflict, and in particular, actual or threatened livestock depredation, is a primary driver in the hatred of carnivores (Sillero-Zubiri and Laurenson 2001). However, there is almost never a simple, linear relationship linking damage, attitudes and actions (Dickman and Hazzah 2016). In some cases, animosity seems roughly proportional to the damage caused by a species, while in most cases, these are not proportional and people may be unusually tolerant or intolerant towards the species (Dickman and Hazzah 2016). Attitude towards a species can relate to a person's actions towards it. Hazzah et al. (2017) found that a negative attitude towards lions was the strongest predictor of actual lion-killing behaviour among Maasai in southern Kenya, but in many cases the attitude may not be the best barometer of actions taken (Dickman and Hazzah 2016). Understanding (in)tolerance of local people and relationship to human actions is complex. Direct killing can result with carnivores either pre-emptively killed or killed in retaliation. Reducing depredation and other types of conflict may reduce animosity towards carnivores and alleviate killing (Dickman 2010). But other forms of conflict resolution, often between groups of people, and increasing tolerance are also necessary.

### 1.3.2 Impacts on carnivores

Humans negatively impact carnivores. Most obviously, humans kill and hunt carnivores such that their distribution and populations are greatly reduced with few exceptions. But, human activities impact carnivores in other ways as well, such as: altered demographic structure, modified interspecific competition, disrupted foraging, reduced breeding capability, and a multitude of other behavioural changes due to the risk of human-caused mortality (Creel, Spong, and Creel 2001; Oriol-Cotterill et al. 2015b). Through targeted killing of individuals of specific age or gender, humans can alter the normal demographic structure of a population. For instance, hunters targeting male lions in Zambia led to substantial depletion of adult males in three protected areas adjacent to hunting reserves (Becker et al. 2013b). Interspecific competition is modified when some species are better able to adapt to human-modified conditions and have altered abundance compared to natural ecosystems. For instance, black-backed jackals and brown hyena compete for scavenging opportunities in South Africa and the loss of additional scavenging opportunities when lions were absent led to reduced brown hyena and increased jackal densities (Yarnell et al. 2013). Introduced carnivores may compete with native carnivores, but also threaten them through the risk of hybridization and disease transmission (Lescureux and Linnell 2014). For instance, Farris et al. (2016) found that feral cats and dogs are excluding and ultimately completely replacing native carnivores in rainforests of Madagascar.

In terms of behavioural impacts, Oriol-Cotterill et al. (2015b) laid out a framework of potential changes in carnivores due to the threat of human-caused mortality. This includes altered spatio-temporal use of the landscape, habitat selection, movement patterns, foraging patterns, group size, sub-adult dispersal, vigilance, prey selection and feeding behaviour. Broadly speaking, carnivores may shift their spatio-temporal use of the environment to lessen their chances of interaction with humans with the result that landscape use is primarily explained by human densities, distribution and activities rather than explained by ecological theory (Oriol-Cotterill et al. 2015b). For instance, in high-risk areas, carnivores may shift their foraging and movement patterns to minimize overlap with humans, and select habitat that provides good refuge from people during times when people are active (see Schuette, Creel, and Christianson 2013). This raft of behavioural changes may ultimately decrease a species' fitness (Oriol-Cotterill et al. 2015b).

### 1.3.3 Benefits derived from each other

Although the above discussion focuses on the negative impacts carnivores and humans impose on each other, there are also benefits. A major benefit humans provide to carnivores is alternative food sources and scavenging opportunities, but these attractants are often associated with higher mortality risk (Wilson et al. 2006; Bateman and Fleming 2012; Athreya et al. 2014; Newsome et al. 2014). Another benefit is reduction or elimination of interspecific competitors (Bateman and Fleming 2012). Carnivores in turn can benefit humans. Less directly, carnivores regulate ecosystems through top-down effects (Ripple et al. 2014). More directly, carnivores can play a role in human health such as by reducing the abundance of white-tailed deer and other species. Reduced deer numbers leads to fewer deer-vehicle collisions (Gilbert et al. 2016), and may reduce the prevalence of Lyme disease in humans (as overabundant deer may help spread the disease) (Velasquez-Manoff 2016). In both ways, carnivores can save human lives and money. However, the ecosystem service benefits that humans derive from carnivores are typically not well described, and the benefits carnivores receive are generally outweighed by the negative impacts.

When examining the positive and negative impacts of carnivores on humans, impacts typically do not co-occur on the same group, i.e. the impacts fall on different groups of the public (Thirgood, Woodroffe, and Rabinowitz 2005). The group(s) that benefit from carnivore presence may not be the one(s) shouldering the burden. Costs may be highly concentrated, with local hotspots of depredation or crop raiding (Naughton-Treves 1997). While local villagers in Botswana spend time and resources attempting to prevent livestock depredation, not always successfully, it is the hunting concession operator who receives the financial benefit of trophy hunting, and local elites who can control or restrict the flow of resources to the community (Lewis and Jackson 2005).

## 1.4 Carnivore vulnerability

Given the myriad ways in which carnivores and humans interact and interfere with each other, it is perhaps unsurprising that carnivores are one of the most challenging taxonomic groups to conserve. Carnivora, species with large body size, and habitat specialists were recently identified as groups with greater sensitivity to habitat fragmentation (Keinath et al. 2016). Carnivores, and particularly large carnivores, also exhibit many biological traits that are linked with higher vulnerability to extinction (Cardillo et al. 2005; Cardillo et al. 2008). These traits include body size, range size, densities, and reproductive parameters (Cardillo et al. 2005; Cardillo et al. 2008), but these often co-vary and can make it difficult to ascribe causation from a single biological trait. Yet, carnivores are known for their tremendous adaptability and as a mammalian order are relatively less threatened than many other orders (Gittleman et al. 2001). Their behavioural flexibility may help some of them to persevere in the face of human expansion better than other taxa. Carnivores, and large carnivores in particular, face a daunting line-up of threats and their persistence is in no way guaranteed.

## 1.5 The power of maps in addressing threats to carnivores

Mapping is often a critical component in the field of biology and has a variety of practical conservation uses for carnivores; see Rondinini and Boitani (2012), Kanagaraj et al. (2013), and García-Rangel and Pettorelli (2013) for overviews of how mapping is involved in carnivore research and conservation. Some purposes of mapping in conservation will be briefly described below, and include: mapping species distributions, spatial modeling, connectivity modeling, and conflict and mortality mapping.

Mapping species can take many forms and at different scales, conducted with a variety of data types and for a variety of end uses (Rondinini and Boitani 2012). For instance, mapping is useful at each step of Johnson's (1980) hierarchy of habitat selection: 1st order (species geographical range), 2nd order (home range), 3rd order (use of habitat components within home range), and 4th order (habitat elements used from within those available at that site). An important scale for mapping but missing from Johnson's hierarchy is the regional (or population) scale where metapopulation processes operate (Hanski 1998). Some of the conservation purposes for mapping distributions include identifying species core areas (Linkie et al. 2006), reserve planning and establishment (e.g. Land of the Leopard NP, Russia) (MNRERF and WWF 2014), priority setting (Durant 2007), and population monitoring (e.g. tracking changes in extent of occurrence/area of occupancy for e.g. the IUCN Red List) (Durant et al. 2015). Spatial modeling (i.e. through species distribution modeling (SDM)/ecological niche modeling) can be useful to understand species' suitable habitat (Swanepoel et al. 2013), relative environmental preferences and tolerances (Pettorelli et al. 2009; Pettorelli et al. 2010), to ascertain risk from potential threats (e.g. climate change) (Forrest et al. 2012), and recovery opportunities (e.g. reintroductions) (Hebblewhite et al. 2011). SDM and habitat mapping can be used in connectivity modeling with the goal of maintaining or restoring population connectivity (Wegmann et al. 2014; Cushman et al. 2015). Finally, mapping HWC or species mortalities can be useful to understand patterns and to determine if hotspots exist (Kushnir et al. 2014; Miller 2015). Once mapped, researchers can prioritize locations for interventions to prevent HWC or mortality events. Thus, spatial analysis and mapping are powerful tools to inform a wide variety of conservation, management, and policy decisions.

Reserve planning and establishment falls within the larger field of zoning and land use planning. Zoning regulates the types of activities and uses that are allowed on a property. Land use planning is a political process used to regulate land uses in an efficient manner to reduce potential land use conflicts. Zoning has been widely used in biodiversity conservation via the creation of protected areas and buffer zones (Linnell et al. 2005). Yet, this activity could be more widely applied at the landscape scale for carnivore conservation specifically, by separating carnivores from activities that lead to HWC (Linnell et al. 2005). Hence activities such as ranching might not be allowed in carnivore zones, or ranching in carnivore zones may be subject to different rules than in areas where large carnivores are not tolerated.

## 1.6 Research approaches to investigating threats to carnivores

Two of the most serious threats to carnivore, and indeed most wildlife, persistence are habitat loss and HWC. Habitat degradation, such as via selective logging or over-grazing, can lead to habitat loss, but is difficult to monitor remotely. Mapping the intensity and extent of habitat loss and HWC, and researching species' response to them is necessary in identifying vulnerability to extinction and designing appropriate conservation strategies. Below I lay out some of the common approaches biologists use.

### 1.6.1 Mapping habitat loss

The first step in researching the impacts of habitat loss is identifying the difference between habitat and non-habitat for the species of interest. For some species, e.g. forest-dependent birds, identifying non-habitat can be easy. But for most species, this is challenging and requires detailed observational data and biological knowledge. Identifying and delineating human-dominated lands and land cover classes can aid in this question.

Earth observation data can be used to describe how humans use land (land use) and the physical features that cover the Earth's surface (land cover). To create land use land cover (LULC) data sets, traditional remote sensing classification approaches cluster spectral signatures and then discriminate between land cover types with 'known' spectral responses (Pfeifer et al. 2012c). Emerging techniques, such as object-based classification, reduce reliance on accurately grouping and discriminating spectral properties by allowing probabilistic class descriptions (Blaschke 2010). LULC data sets identify the location and extent of human activities. Yet, unambiguous identification of urban and agricultural areas is challenging and varies greatly across data sets (Potere and Schneider 2007; Fritz et al. 2011; Vancutsem et al. 2012). Land cover classification may be relatively easy where vegetated landscapes are homogenous and extensive, e.g. some croplands. However, in heterogeneous landscapes with smaller, patchier agricultural fields, discriminating cropland from natural land cover using traditional classification algorithms can be difficult (Tchuenté, Roujean, and De Jong 2011; Vancutsem et al. 2012). Africa's croplands are often small and embedded in savannah, a naturally challenging and heterogeneous land cover; together this creates a particularly difficult classification challenge (Pfeifer et al. 2012a; Watson et al. 2015). Thus, estimates of croplands in Africa can vary tremendously; Fritz, See, and Rembold (2010) compared three data sets and found the total cropland area for Africa varied from <1,000,000 km$^2$ to nearly 7,000,000 km$^2$. Hannerz and Lotsch (2008) also found significant differences; in 1/3rd of African countries, disagreement in cropland extent among different land cover products was >25% of the total country area. In a global comparison of urban extent, total urban area varied by greater than an order of magnitude among six different land cover products (Potere and Schneider 2007). This level of variation can hinder confident application of the data for conservation and research applications.

### 1.6.2 Using species distribution models to research threats

Once the extent of human land cover is more positively identified, then research into species-specific responses to habitat loss can proceed. There are multiple approaches to this and the approach at least partially relies on the type of species data available. A deductive habitat suitability model can be used to estimate habitat potentially used by a species. An example of this is a habitat suitability index, an analytical, species-specific function derived from expert knowledge that does not require presence or movement data (e.g. Rondinini, Stuart, and Boitani 2005). Correlative, or inductive, approaches to habitat modeling use species occurrence and/or movement data in a model to predict distribution or occupancy based on a variety of environmental and human predictors. With presence-only data there are limited options that also tend to have limited power (generative approaches like envelope models). Presence data can be contrasted with other point data; pseudo-absence (PsA) data if no absence data are available (i.e. discriminative approaches like generalized linear models). Ideally, absence data are available to contrast with presences. The above are termed species distribution models or environmental niche models. Standardized surveys documenting presence and absence, and that minimize sampling bias and address imperfect detection, can achieve more accurate estimates of occupancy from these models (Guillera-Arroita et al. 2015). If telemetry data are available (or rigorous sign survey data), resource selection functions can compare used with available habitat (Manly et al. 2002). Finally, with repeated independent survey data that document detection-non-detection, imperfect detectability can be incorporated into occupancy modeling (MacKenzie et al. 2006). Other emerging approaches that extend beyond correlative modeling include spatially-explicit individual-based models and mechanistic models (Shenk and Franklin 2001). Together these methods can be used to understand how individuals or species respond to threats such as changing land cover and habitat loss.

I focus on species distribution models, and specifically on the presence-PsA approach as this is the most common type of distribution modeling due to the difficulty of obtaining absence data (Ahmed et al. 2015; Guillera-Arroita et al. 2015). Yet, when applying this method, the crucial question of how to generate the PsA data remains. This is important because many different PsA generation methods have been proposed, and the different methods result in substantially different model results including model predictions (Chefaoui and Lobo 2008), variable importance (Stokland, Halvorsen, and Støa 2011), and response curves (Thuiller 2004; Lobo and Tognelli 2011). Indeed, the different approaches will shape a model's research aims and ecological inferences (Elith et al. 2011). However, there is no consensus approach in the SDM literature and the discussion surrounding PsA generation have largely been from a statistical perspective.

### 1.6.3 Using SDM for spatial risk predictions of human-wildlife conflict

Despite this challenge, distribution models using presence-PsA data are also used to explore the second primary threat to large carnivores, HWC. Much research on HWC is descriptive, simply indicating species involved, rates, timing and locations. But it is also common to assess factors that are related to depredation events. This can be done in a number of ways, from simple approaches like finding the correlation between depredation events and predictors (e.g. Van Bommel et al. 2007) to more rigorous approaches that use a matched-pairs design and regression or classification trees to compare locations with attacks to nearby locations without attacks (e.g. Mech et al. 2000; Kolowski and Holekamp 2006). In order to spatially predict risk however, there are three common approaches, spatial interpolation (such as Kriging), spatial association (cluster techniques like Moran's I), and correlation modeling (Miller 2015). Correlation modeling techniques (i.e. species distribution modeling) were the most common approaches in a recent review (Miller 2015). However, just the same as in the broader SDM literature, there is insufficient guidance in properly generating PsA data and setting the analysis extent. This point is not explicitly considered in the existing depredation risk modeling literature, but it should be.

## 1.7 Research aims

As shown above, some of the most common approaches to researching primary threats to carnivores have substantial existing methodological challenges. Overall this dissertation contributes to the development of conservation science through refining and reframing some common methods used to study wildlife distributions and their threats using case studies of large carnivores in East Africa.

The specific aims of this project are:

i. To improve identification of habitat loss by creating a new tool for classifying land cover, and to use this tool to generate a new assessment of human-dominated lands in East Africa.
ii. To reframe the discussion of pseudo-absence generation strategies in species distribution modeling to recognize the different strategies' influences on ecological inference.
iii. To identify the most important predictors of presence for four large carnivores (cheetah, African wild dog, leopard and lion) and environmental factors correlated with their decline in East Africa.
iv. To explore the validity of using risk modeling approaches to map human-large carnivore conflict.

## 1.8 Thesis structure

This thesis is divided into the following seven chapters.

- Chapter 1 provides an overview of carnivore importance, common interactions between carnivores and human, primary threats to carnivores, common methods used to analyse these threats, challenges associated with these methods, and the specific aims of this dissertation.
- Chapter 2 describes the study area (East Africa), wildlife resources, specific regional threats to wildlife, and ends with general descriptions of the five primary large carnivores involved in the dissertation.
- Chapter 3 introduces a new tool to conduct land cover classification that uses manual identification of human land cover using Google Earth data to develop a new assessment of human-dominated lands in East Africa.
- Chapter 4 demonstrates how previously proposed methods of generating PsA data alter the ecological interpretation of distribution model results with a case study exploring range contraction of the cheetah. I gather >10,000 cheetah presence records and generate PsA data using five different approaches. I expand on the concepts behind delimitation of background extent and PsA generation, and use the human land cover data set developed in Chapter 3 as an input.
- Chapter 5 applies insight on PsA generation strategies gained from Chapter 4 to examine predictors correlated with large carnivore presence and decline. I test three hypotheses:
    - H1: increasing anthropogenic impacts threaten large carnivore persistence;
    - H2: carnivores with larger home ranges are more sensitive to habitat loss;
    - H3: habitat loss is a greater factor impacting carnivore distributions than human population density.
- Chapter 6 uses depredation records in northern Tanzania and the Maxent modeling software (Phillips, Anderson, and Schapire 2006) to demonstrate how livestock depredation risk from large carnivores can change when adjusting the background study extent and accounting for changes in factors associated with livestock depredations (i.e. differences in event location, predator, season, and livestock involved).
- Chapter 7 presents the limitations and implications of the study results, highlights the challenges for carnivore conservation in East Africa, and suggests how these results can be incorporated into conservation efforts.

# Chapter 2

# Study area

Sunrise over the Great Ruaha River. © Andrew Jacobson

# Chapter 2 Study area

## 2.1 Africa

Africa is the second largest continent with the second largest human population. Africa's human population is about 1.1 billion and is projected to reach over two billion by 2050 (UN 2013). Africa's share of the world population has risen substantially since 1900 and will continue to do so. Much of the world's human population growth between 2013 and 2050 will occur in high-fertility countries, which are primarily concentrated in Africa. Populations of many countries in Africa are projected to triple by 2100 (UN 2013).

As the population expands, African populations will remain young and Africa as a whole will gain an increasing share of the global workforce. By 2050, its labour force will be over one billion people, larger than China or India (Ahlers et al. 2014). Africa is composed primarily of low and lower-middle income countries (Ahlers et al. 2014). While African economies stagnated in the 1970s and 1980s, it has become the 2nd fastest growing region in the world, with gross domestic product growing at 4.5% between the mid-1990s and 2010 (Ahlers et al. 2014).

East Africa (defined below) is the focus of this thesis, with three of four research chapters using the full extent of the region. The fourth chapter focuses exclusively on northern Tanzanian rangelands, within the larger East Africa region.

## 2.2. East Africa

### 2.2.1. Location, history and socio-economics

The East African Community (28.86° – 41.89° E, 4.63° N – 11.75° S), a geopolitical definition and hereafter shortened to East Africa, is composed of five countries: Tanzania, Kenya, Uganda, Rwanda and Burundi (Figure 2.1). Neolithic hunting and foraging people, represented today in Tanzania by the Hadza and Sandawe, gave way to Bantu expansion in the Great Lakes region around 1000 BC. The Bantu introduced agriculture and absorbed or displaced the previous cultures. The Bantu organized into clans and eventually coalesced into kingdoms. Although the Portuguese first reached East Africa (Mombasa) in 1498, European influence did not begin extending inland from the coasts until the mid 1800s. At the Berlin Conference of 1884-1885, East Africa was split between British East Africa encompassing Kenya and Uganda, and German East Africa including

Rwanda, Burundi and the mainland part of Tanzania (known as Tanganyika). With Germany's defeat in World War I, German East Africa broke apart; the British took control of Tanganyika and Belgian authorities oversaw Rwanda and Burundi. Rwanda and Burundi became United Nations trust territories after the League of Nations disbanded in 1946, while Kenya and Uganda were maintained as the Colony and Protectorate of Kenya. All nations gained independence early in the 1960s: Tanganyika in 1961, Burundi, Rwanda and Uganda in 1962, with Kenya (and Zanzibar) in 1963. In 1964, the two independent states of Tanganyika and Zanzibar joined together as the United Republic of Tanzania. With a long history of cooperation, the five nations revived an older partnership and became the East African Community in 2000. South Sudan officially joined the Community in September 2016 after this analysis was completed and was not included.

Despite the interwoven political history and on-going partnership, the five countries are quite distinct (Table 2.1). Rwanda and Burundi are small, landlocked nations with dense human populations owing to their rich volcanic soils and ample rainfall. On the other hand, Tanzania is almost 40 times larger than Rwanda with a much lower human population density and substantially more land remaining in a natural state. Kenya is similar to Tanzania with a large land base and low population density, although much of its population is crammed into the more habitable southern part of the country. Uganda in some ways is a blend of the other countries; of intermediate size, density and remaining natural land.

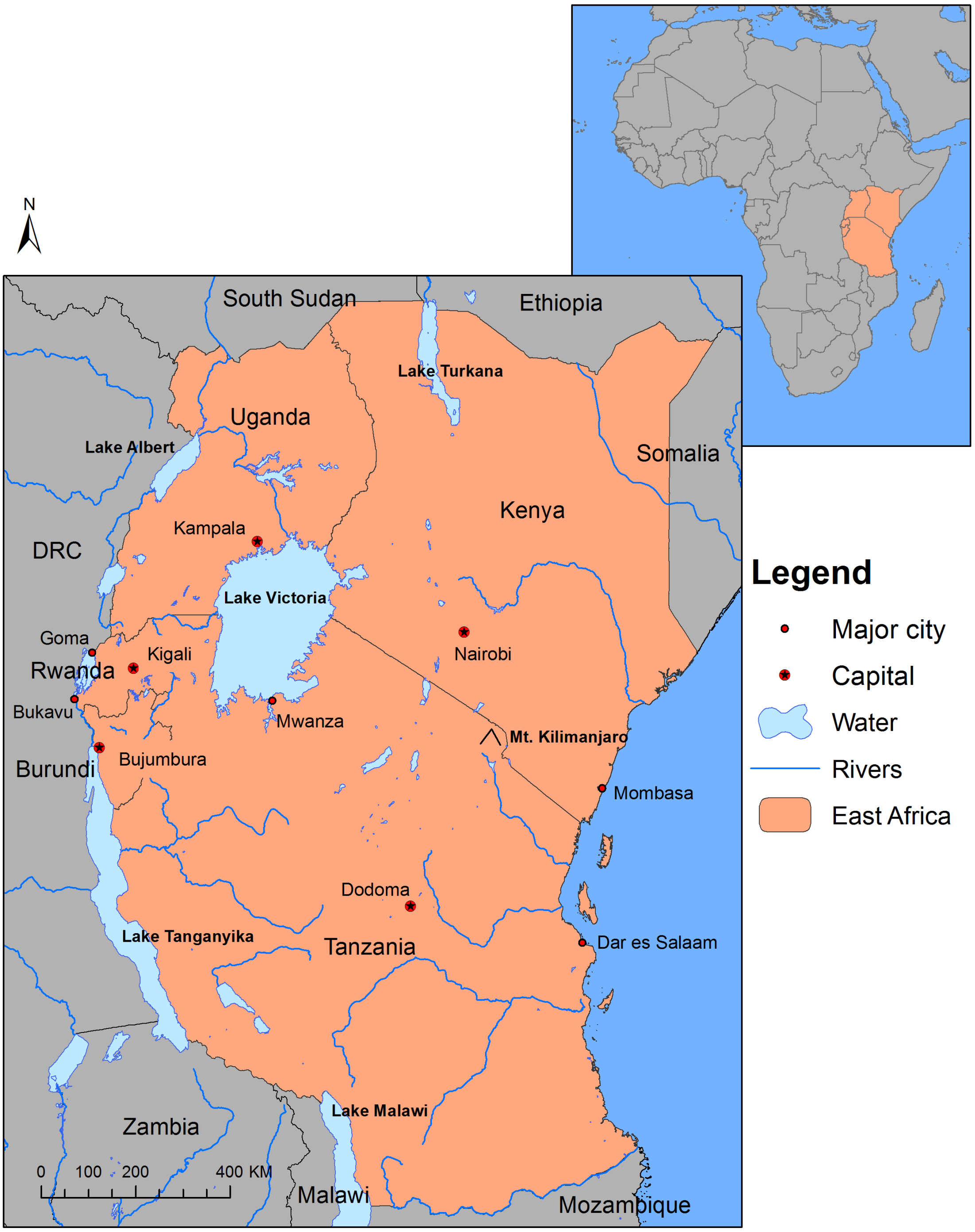


**Figure 2.1** **The five countries of the East Africa Community and some landmarks. Cities are either the capital of a country or have a population >500,000. DRC - Democratic Republic of the Congo.**

**Table 2.1 Summary statistics of countries in the East Africa Community.**

| | Total area[1] ($km^2$) | Human pop'n[1] in millions (growth rate) | Human pop'n density[1] ($people/km^2$) | % natural[2] | % within protected areas[3] |
|---|---|---|---|---|---|
| Rwanda | 26,338 | 13.0 (2.5%) | 500 | 18 | 8.6 |
| Burundi | 27,830 | 11.1 (3.3%) | 405 | 14 | 3.7 |
| Uganda | 240,038 | 38.3 (3.2%) | 180 | 57 | 9.4 |
| Kenya | 580,367 | 46.8 (1.8%) | 80 | 83 | 7.3 |
| Tanzania | 947,300 | 52.5 (2.8%) | 56 | 68 | 20.3 |
| Total | 1,822,000 | 161.7 | | 70 | 14.4 |

1. CIA 2016
2. Jacobson et al. 2015
3. Jason Riggio et al. unpublished data. IUCN categories 1-4 and the NCA. In sum this represents 47 National Parks, 26 Game Reserves, 23 National Reserves, 12 Wildlife Reserves, 5 Nature Reserves and 1 Conservation Area.

East African countries are among the poorest in the world in terms of per capita income. Yet, all economies, with the possible exception of Burundi, have recently enjoyed relatively high economic growth rates. Indeed Kenya, the most prosperous of the five countries, and the economic and transportation hub of eastern Africa, recently passed a World Bank threshold to be defined as a lower middle income country. Despite this, the vast majority (~70-80%) of all countries' populations are involved in the agricultural sector, and primarily in small-scale rain-fed subsistence farming (Gelorini and Verschuren 2012). In drier regions of East Africa where rain-fed agriculture is uncertain, transhumant pastoralism is common. That is, people whose livelihoods primarily depend on domestic animals and who seasonally migrate with their livestock between different pastures (Fratkin 2001). Although farmers or others may raise livestock, and commercial ranchers raise livestock for market, pastoralists use their herds for daily subsistence. The largest pastoral groups of East Africa include the Maasai in southern Kenya/northern Tanzania, and Turkana in northern Kenya.

### 2.2.2. Land use, protected areas and wildlife utilization

In terms of land use, pastoralists occupy up to 70% of the total land of Kenya, 50% of Tanzania and 40% of Uganda (Fratkin 2001). Agriculture is another major land use. Over 80% of densely settled Rwanda and Burundi are converted to agriculture or urban areas

(Jacobson et al. 2015). Agriculture is more patchily distributed in Kenya, Tanzania, and Uganda, although it is concentrated in suitable areas, such as around Lake Victoria. All countries are urbanizing quickly with rates of urbanization >4% annually (CIA 2016). Hence, urban areas and settlements represent an increasing share of land use. Wildlife protection is also a substantial land use although this varies widely across countries.

East Africa has a rich history of wildlife protection with the establishment of game reserves (GR) in British East Africa (now Kenya) in 1896. There is now an extensive network of protected areas that contain some of the world's most iconic parks, such as Serengeti National Park (NP)/Maasai Mara National Reserve (NR), Mt. Kilimanjaro NP, and Amboseli NP (Figure 2.2). The amount of protection varies widely between the countries with upwards of 20% protected in Tanzania, and only 3.7% in Burundi (Jason Riggio et al. unpublished data). These percentages would be larger if forest reserves, wildlife management areas or other less restrictive protected areas (i.e. IUCN categories >4) are included. These areas are excluded here as the other categories of protection (IUCN categories <4; e.g. NPs, GRs, NRs, Wildlife Reserves, Nature Reserves and Conservation Area) are fully protected, do not permit settlement (with the exception of Ngorongoro Conservation Area, NCA) and only allow limited resource extraction. Forest reserves and wildlife management areas are less protected, and can permit extensive habitat conversion and/or settlement. Indeed, these designations and related levels of enforcement matter as stricter protected areas in East Africa can better prevent habitat loss and wildlife declines (Stoner et al. 2007a; Pfeifer et al. 2012b; Green et al. 2013).

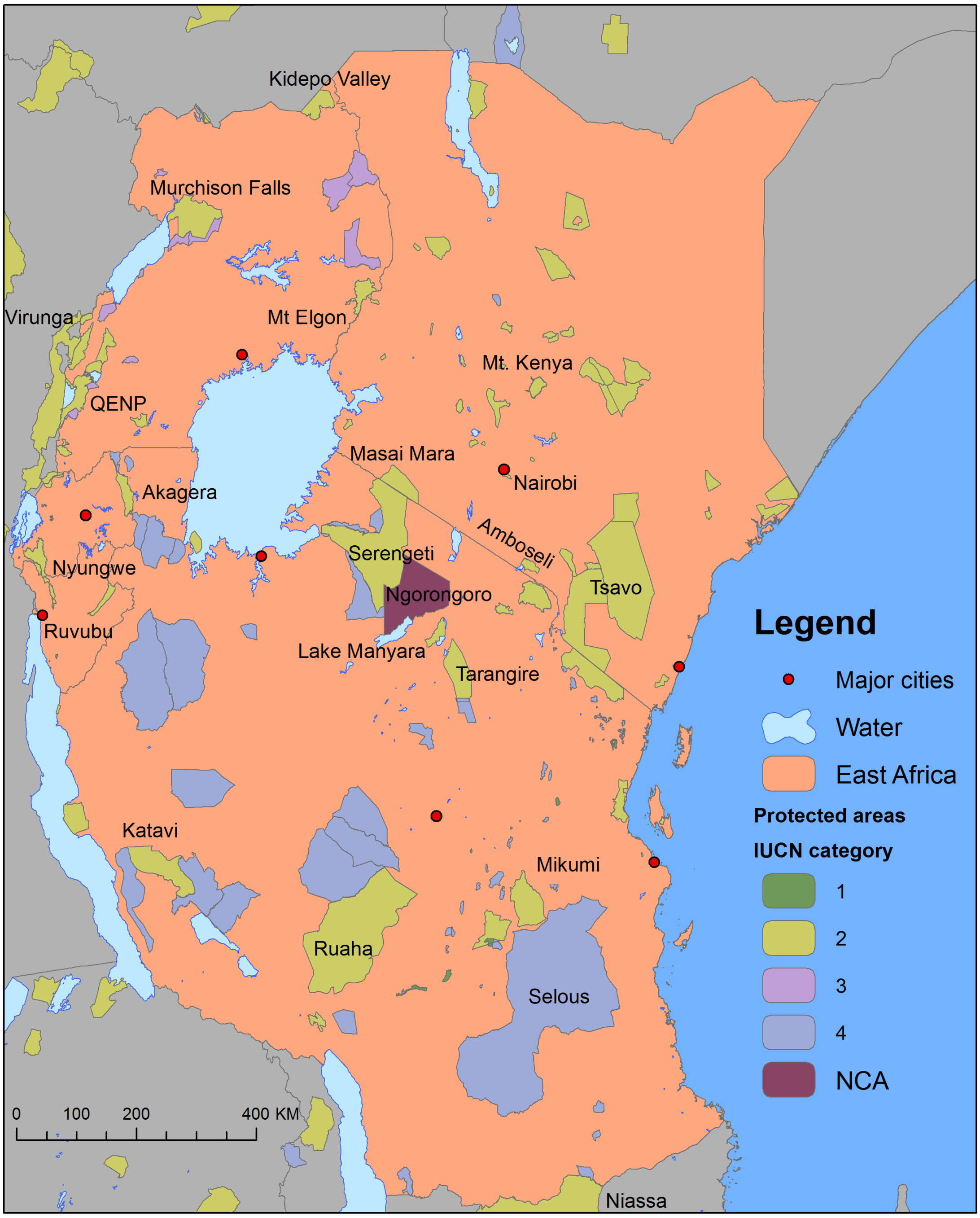


**Figure 2.2 Protected areas of East Africa. Those shown have the strictest protection categories, with 1 more strict and 4 less strict. QENP - Queen Elizabeth NP.**

An important difference in protection categories between East African countries, and wildlife utilization more broadly, is that Tanzania currently allows trophy hunting. Trophy hunting began with the colonial powers and many of the early reserves were set aside by Europeans as good places to hunt, i.e. they held substantial concentrations of large mammals (Caro 2003). (Alternatively, some East African protected areas were established to protect populations of iconic species, such as mountain gorillas in Volcanoes NP, Rwanda.) Shortly after the beginning of colonial rule by both German and British authorities, wildlife regulations were instituted that centralized control of wildlife resources and restricted local use (Kock 1995; Nelson 2007; Nelson, Nshala, and Rodgers 2007). Game Departments and hunting reserves were established while traditional methods of hunting were disallowed. Thus, the states took ownership of wildlife while local use became largely prohibited. These practices continued throughout the colonial and early post-independence periods. In Tanzania, while many of the early game reserves were converted to national parks (which ban trophy hunting) between the 1950s and 1970s (Caro 2003), a substantial portion of Tanzania's remaining protected areas permit hunting (see 2.2.6). On the other hand, Uganda banned trophy hunting in 1979 although it is currently experimenting with limited hunting of select herbivores and leopard (Lindsey, Roulet, and Romanach 2007). In addition, Kenya banned trophy hunting in 1977 and hunting is not currently practiced in Burundi or Rwanda (although it historically occurred in Akagera NP).

State protected areas are the cornerstone of biodiversity protection efforts in East Africa, but there are also limited community and private-led wildlife protection efforts. In Tanzania, some local communities' have agreements with private tourism operators to retain areas where wildlife is the primary land use (Sachedina and Nelson 2010). There are also group ranches or communally owned pastoral areas (a.k.a. conservancies), and in some cases, private livestock ranches, which can also be wildlife-friendly and contribute substantially towards wildlife conservation (i.e. Mpala Ranch and Shompole Group Ranch in Kenya). Finally, there are still some remote unprotected and largely undisturbed landscapes that may be important for wildlife but fall under none of the above categories.

### 2.2.3 Climate for protection

Related to the process of centralizing government control over wildlife resources, the establishment of protected areas by colonial governments typically led to the eviction of rural communities or loss of access to grazing land/resources (Neumann 2000). Indeed, every country but Burundi (which may have been data deficient), has documented evictions of people during protected area establishment (Brockington and Igoe 2006). Although many East African protected areas were gazetted in the early-to-mid 20th century, evictions have continued into the 2000s in Tanzania (Benjaminsen and Bryceson

2012). Therefore, the establishment of protected areas is insinuated with on going land tenure insecurity and the underdevelopment of rural areas (Neumann 2000; Benjaminsen and Bryceson 2012). This conflict has resulted in a land rights and justice movement primarily among pastoralist groups in Tanzania (Neumann 2000). Indeed, each country in East Africa but Burundi has downsized or de-gazetted protected areas (Mascia and Pailler 2011).

Recognizing this conflict between local communities and wildlife protection, a new discourse developed in the 1980s and 1990s that emphasized local participation in conservation (Benjaminsen and Bryceson 2012). This movement argued that local people should participate in protected area management, and benefit from conservation. In East Africa, possibly the best example of this shift was the development in Tanzania of a new Wildlife Policy in 1998 that allowed for the designation of wildlife management areas (Benjaminsen and Bryceson 2012). These new land use types were promoted, in which local communities would manage and benefit from wildlife on their land. Many wildlife management areas were developed (>35,000 km$^2$), primarily in buffer zones of or corridors between state protected areas (Caro and Davenport 2015). Although they were possibly never given the chance to succeed, their effectiveness in protecting wildlife and improving rural livelihoods is mixed or too early to tell (Caro and Davenport 2015). Indeed, the new Wildlife Act of 2009 reformulated wildlife policy and re-consolidated governmental control of wildlife at the expense of local communities (Benjaminsen and Bryceson 2012).

Against this backdrop, state protected areas still represent a substantial portion of land use in most East African countries (Table 2.1). They also generate substantial economic revenue for the government; tourism in Tanzania generated $1.95 billion in 2014 (Caro and Davenport 2015) and >$800 million in Kenya in 2010 (Homewood, Trench, and Brockington 2012). Tanzania in particular has continued to invest in the protected area system. In the last 20 years, Tanzania has gazetted six new NPs, enlarged five existing NPs, upgraded several GRs to NPs, and established a new protected area category (nature reserves) to better protect forest reserves (Caro and Davenport 2015). Yet to a large extent, African protected areas are underfunded and under tremendous pressure (Lindsey et al. 2014; Lindsey et al. 2016). Without greater funding for law enforcement, management, and infrastructure development, mammals including large carnivores like lions, are declining within protected areas (Craigie et al. 2010; Packer et al. 2013; Bauer, Chapron, et al. 2015). Indeed, more effective management (i.e. through increased funding) would likely be more beneficial for biodiversity than protected area expansion (Costelloe et al. 2016).

### 2.2.4. Climate and geography

East Africa is heterogeneous from an environmental and geographic standpoint. The region has tremendous variation in rainfall (172-2625 mm) (Hijmans et al. 2005), with most precipitation occurring in October-December and March-May (Gelorini and Verschuren 2012). The Great Lakes and mountainous regions receive higher rainfall levels, while northern and eastern Kenya are the driest. The Great Lakes of the East African rift system are a tremendous natural resource; they contain ~25% of the world's non-frozen, surface freshwater. These lakes include the world's 2nd largest lake by volume, Lake Victoria, and the 2nd deepest lake, Lake Tanganyika. Lakes Victoria, Edward and Albert are the start of the (White) Nile River. The region varies greatly in elevation, from sea level to Mount Kilimanjaro, the highest peak on the continent at 5882 m. The Virunga and Ruwenzori mountain ranges are on the western edge of East Africa, while scattered volcanoes form the Eastern Arc Mountains in Tanzania and Kenya. Hence, the region also varies greatly in temperature with mean average temperatures ranging from -4° to 29° C (Hijmans et al. 2005).

While East Africa contains six terrestrial biomes (plus freshwater), one predominates - tropical and subtropical grasslands, savannahs, and shrublands (Figure 2.3) (Olson et al. 2001). Moist broadleaf forests, and deserts and xeric shrublands, are the two next most extensive biomes. Relatively little of East Africa has >20% forest cover and this is mostly located along the eastern coast, along the western boundary of East Africa and scattered across areas of higher elevation (DiMiceli et al. 2011). In a transition zone in central Tanzania, acacia savannah blends into miombo woodland further south.

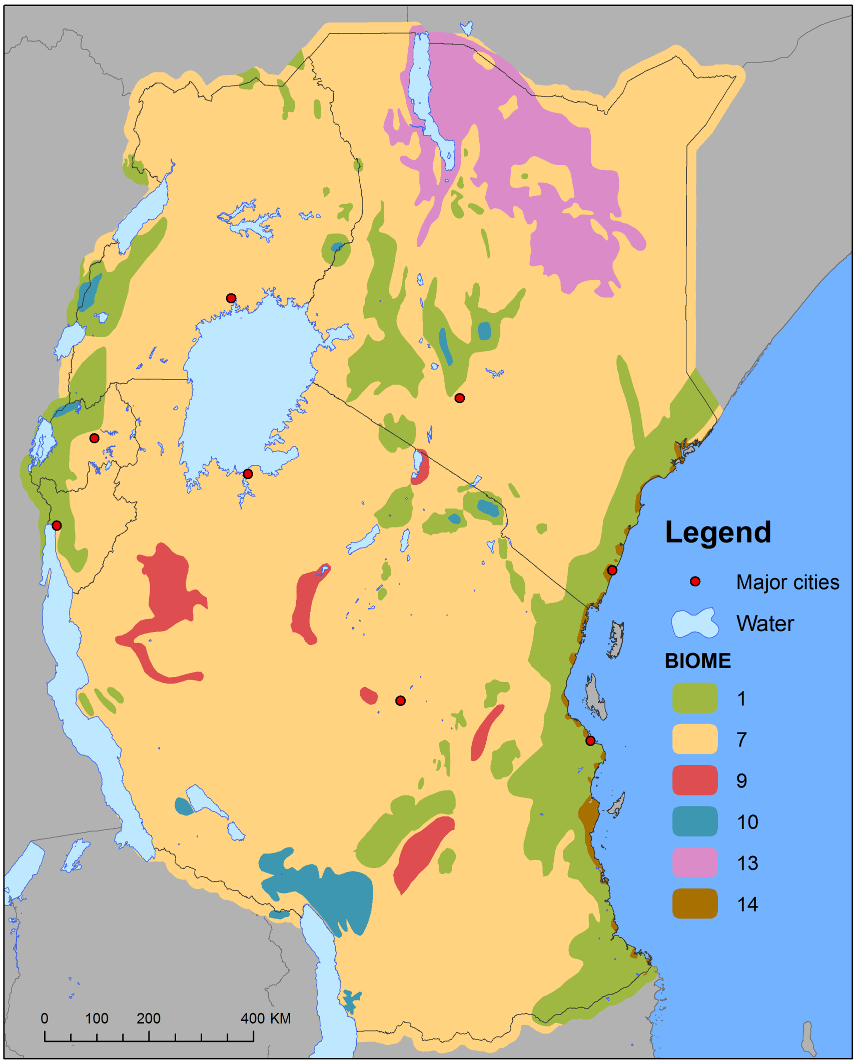


**Figure 2.3 Distribution of seven biomes (including water) in East Africa: 1 - tropical and subtropical moist broadleaf forests; 7 - tropical and subtropical grasslands, savannahs and shrublands, 9 - flooded grasslands and savannah; 10 - montane grasslands and shrublands; 13 - deserts and xeric shrublands; 14 - mangroves.**

### 2.2.5 Wildlife resources

East Africa is tremendously important in regards to biodiversity and wildlife. Many of the world's most iconic species including carnivores such as the lion and cheetah, great apes including gorilla and chimpanzee, as well as elephants, and giraffe reside here. The Eastern Arc Mountains and coastal forests are a biodiversity hotspot and contain some of the highest rates of endemic plants and vertebrates in the world (Myers et al. 2000). Tanzania itself may be the most important country for conservation in Africa overall, with substantial wildlife populations and as a hotspot of vertebrate endemism (East 1999;

Jenkins, Pimm, and Joppa 2013; Caro and Davenport 2015). In terms of herbivores, Tanzania alone held >50% of the African buffalo population, nearly half of all giraffe *Giraffa* spp., ~40% of the common eland, and >70% of the wildebeest population (East 1999).

Tanzania and East Africa as a whole hold globally important carnivore populations. Sub-Saharan Africa is one of the only places left in the world which has much of its large carnivore guild intact (Dalerum et al. 2009). The large carnivore guild here, remnants of Pleistocene mega fauna, have been primarily lost on all other continents. The African large carnivore guild consists of one canid, three felids, and three species of hyena: the African wild dog, lion, leopard, cheetah, spotted hyena, striped hyena, and brown hyena. East Africa holds roughly a third of all cheetahs, nearly a half of all African wild dogs, and over a half of all African lions (Woodroffe and Sillero-Zubiri 2012; Riggio et al. 2013; Durant et al. 2015). Large carnivore populations in East Africa are centred on protected areas with significant extension into multiple use landscapes. The Selous GR holds the largest single population of lions and second largest wild dog population on earth (Woodroffe and Sillero-Zubiri 2012; Riggio et al. 2013). The largest single population of cheetahs in the region is in the Serengeti/Mara/Tsavo (Durant et al. 2015). However, wildlife populations inside and outside protected areas across East Africa have been declining (Stoner et al. 2007a; USAID 2011; Ogutu et al. 2016).

Ogutu et al. (2016) found that wildlife populations in Kenya, which has the most comprehensive data, declined on average by 68% between 1977 and 2016. These data covered 18 species of herbivore. An earlier analysis of Kenyan wildlife found that wildlife populations in national parks and reserves were declining at a comparable rate to non-protected areas, and to the country overall (Western, Russell, and Cuthill 2009). While Tanzanian wildlife populations are also declining (Newmark 1996; Stoner et al. 2007b; Chase et al. 2016), it seems that protection there is having some impact on slowing wildlife declines. Stoner et al. (2007a) found that of 23 herbivore species, more species were faring better in national parks than in areas with little to no protection. Yet, even within many Tanzanian parks, wildlife populations were in decline. Surveys suggest large mammals in Uganda have also steeply declined since the 1960s, although many appear to have stabilized more recently (USAID 2011). Little data is available from Rwanda and Burundi and they are not expected to hold significant populations of large mammals.

### 2.2.6 Threats

Of the four primary threats to wildlife described in section 1.4, all four are important in East Africa. Human populations in East Africa are expanding particularly rapidly, driving agricultural and settlement expansion and causing habitat loss and fragmentation. Around Serengeti NP in Tanzania, natural areas are being converted to agriculture at rates up to 2.3% a year (Estes et al. 2012). Large-scale infrastructure projects in the region are also impending drivers of habitat loss (Laurance et al. 2015). Logging for the charcoal trade can also degrade habitat (Arnold, Köhlin, and Persson 2006; Zulu and Richardson 2013). Tanzania and Kenya are identified as two of five biodiverse countries globally with the highest projected suitable habitat loss by 2050 (Visconti et al. 2011).

Pastoralism is a dominant land use in much of East Africa, and when well-managed or at low intensity it can be ecologically neutral or even positive (Darkoh 2003; Reid, Thornton, and Kruska 2004; Fynn et al. 2016). Livestock grazing is a disturbance and promotes heterogeneity in the landscape, particularly at old corral sites that become nutrient-rich hotspots (Reid, Thornton, and Kruska 2004). However, as grazing land is lost to cultivation, and human and livestock populations increase, migration becomes more difficult and the pressure of land ownership, sedentarization and fencing becomes stronger (Western, Groom, and Worden 2009). Grazing pressures that used to vary over time and space become concentrated, with serious, negative implications for biodiversity (Darkoh 2003; Western, Groom, and Worden 2009). Yet, conversion of pastoral lands to cropland is also a major threat to carnivores in the region (Reid, Thornton, and Kruska 2004). Carnivore species richness was similar between pastoralist and protected areas but much lower in cultivated areas (Msuha et al. 2012) demonstrating the potential compatibility with pastoralism and the limited value of croplands for carnivores. Additionally, of seven carnivore species surveyed in Tanzania, all displayed high sensitivity to land conversion and avoided cropland areas (Pettorelli et al. 2010).

While carnivore densities can be similar in pastoral landscapes to protected areas (see Schuette, Creel, and Christianson 2013), pastoralist reliance on livestock can make carnivores particularly vulnerable if depredations occur. Pastoralists in the region are typically politically marginalized and cash poor (Homewood and Rodgers 1991). Therefore, human wildlife conflict (HWC) can be influential in pastoralist attitude and action towards carnivores, even though other causes such as disease may be greater mortality factors to livestock (Dickman et al. 2014). Pastoralist killings (including via poisoning) can be the largest source of mortality locally for some carnivores (Kissui

2008a). Therefore, while pastoralist coexistence with carnivores is possible it cannot be guaranteed.

Prey decline is also a growing threat to East African carnivores. Wildlife populations including prey are declining rapidly as detailed above. An important contributor to prey decline is via bushmeat hunting. Bushmeat hunting has emerged as the primary threat to wildlife in parts of Kenya (Okello and Kiringe 2004). In Tanzania over 2,000 tons of bushmeat are confiscated annually with a value of >$50 million (Lindsey et al. 2013b). This is obviously a fraction of the true scale of the trade. It should be noted however that hunting is rarely practiced by pastoral groups; the Maasai only traditionally hunted for ritual purposes or in times of famine (Homewood and Rodgers 1991).

Finally, direct killing and utilization is also an issue. Several pastoral tribes, notably the Maasai, traditionally hold lion hunts as rites of passage (Goldman, de Pinho, and Perry 2013). Another form of killing, and one that is relatively better studied than say problem animal control, is trophy hunting. Tanzania is the only country to allow trophy hunting of carnivores in the East Africa Community (with the exception that a few leopard can be hunted in Uganda). Of the large carnivores of interest (see next section), the lion, leopard and spotted hyena can be hunted in Tanzania. Hunting reserves comprise >13% of Tanzania and lions can be hunted in >33% of their range in Tanzania (Lindsey, Balme, Funston, et al. 2013a). Yet, hunting as practiced in Tanzania has been shown to be unsustainable for lions and leopards, leading to population decline (Packer et al. 2009; Packer et al. 2010).

## 2.3 Carnivore profiles

Of six large carnivores in East Africa, I ran distribution models on four of them. I neglected the striped and spotted hyena. Chapter 4 uses the cheetah as a case study. Chapter 5 models the distribution of the cheetah, wild dog, leopard and lion. In Chapter 6, I develop depredation risk models with HWC data in which the leopard, lion and hyena are the primary large carnivore perpetrators. The brown hyena does not occur in the study area and while the spotted is probably a greater threat to livestock in the region ((Ogada et al. 2003; Kolowski and Holekamp 2006; Lyamuya et al. 2014), the distinction between striped and spotted hyenas when reporting HWC events is often overlooked. The HWC data I used did not distinguish between the species and thus I refer generically to hyena.

### 2.3.1 Cheetah

**Figure 2.4. Two cheetahs on a termite mound in Tarangire NP. © Andrew Jacobson**

An often-solitary yellow cat with black spots, the cheetah can be confused with the leopard. Distinguishing the cheetah from the leopard are the black teardrops that run from its eyes to its mouth, a more pronounced chest cavity, and slightly taller and more slender build. The cheetah is between 110-140 cm in length and up to about 90 cm in height (Caro 1994). With males slightly larger and heavier than females, cheetahs range in weight from 35-65 kg (Sunquist and Sunquist 2002).

The social structure of cheetahs is unique among felids. Males typically hold small territories, sometimes only 30 km$^2$ in the Serengeti, while females are not territorial and range across larger areas encompassing multiple male territories (Caro 1994). Females are solitary or accompanied by their dependent young. Males can be solitary, but also form small stable coalitions of related or unrelated males (Caro 1994). Females wander through male territories, breeding with multiple males. In less productive habitats, home ranges of both genders may expand to >1,500 km$^2$ (Marker 2002; Belbachir et al. 2015).

The cheetah is known as the world's fastest land mammal. Cheetahs accelerate quickly and can reach speeds >100 kph (Sharp 1997), but can be outmanoeuvred by prey. A typical hunt begins with a stalk then chase while a minority of hunts skip a stalk altogether (Hilborn et al. 2012). Cheetahs prefer to kill the most available prey present at a site within a prey body mass range of 23-56 kg, such as impala, springbok or Thomson's gazelle (Hayward, Hofmeyr, et al. 2006b). In coalitions, cheetah may hunt larger prey like wildebeest or zebra. Cheetah feed primarily on flesh due to their slender teeth and jaw. Focusing on young ungulates (fawns) or hares helps cheetah have a relatively high hunting success rate (Hilborn et al. 2012). Cheetah rarely scavenge (Sunquist and Sunquist 2002) but are themselves susceptible to kleptoparasites such as the lion and spotted hyena (Hayward, Hofmeyr, et al. 2006b). Cheetahs have no defence against the lion, always abandoning their kills when faced with the larger felid (Hunter, Durant, and Caro 2007).

Reliant on eyesight for hunting and yet sensitive to heat, cheetah are usually diurnal, with peak activity in morning and evenings (Caro 1994; Cozzi et al. 2012). Nocturnal activity can vary with the moon cycle, peaking on moonlit nights (Cozzi et al. 2012), and they may be more nocturnal in specific geographic regions such as the Saharan desert (Belbachir et al. 2015).

Formerly extant across Africa, much of the Middle East and into India, the cheetah is currently restricted to Africa except for a small population in Iran (Durant et al. 2016). Cheetahs are listed by the IUCN as vulnerable and currently reside in ~3,100,000 km$^2$, or 9% of their historical distribution. Generally thought to be a desert and savannah-specialist, the cheetah inhabits a range of ecosystems, including desert, savannah, bush and wood lands (Myers 1975; Sunquist and Sunquist 2002). The species is absent from more mountainous regions, African tropical forests and mangrove habitats (IUCN/SSC 2007a). Densities are low, ranging from 0.02 to as low as 0.0002 per km$^2$ (Durant et al. 2016). The population is tentatively estimated at 7,100 individuals (Durant et al. 2016). While strongholds are in Botswana, Namibia, Kenya and Tanzania, there is widespread population decline and Durant et al. (2016) recommended the species be uplisted to Endangered.

Primary threats to cheetahs are habitat loss and fragmentation, HWC, illegal wildlife trade, and overharvesting (Durant et al. 2016). Interspecific competition can be a major constraint, with lions and spotted hyenas an important regional cause of mortality for cubs and juveniles (Laurenson 1994; but see Mills and Mills 2013). Cheetahs are sometimes described as "fugitive species," avoiding their primary competitors, the lion and spotted hyena, at a fine scale in space and time (Durant 1998; Durant 2000; Broekhuis et al. 2013; Swanson et al. 2014).

A relatively minor threat to livestock, cheetah prey on smaller stock such as sheep and goats during the daytime and rarely attack the boma at night (Maddox 2003; Woodroffe et al. 2007b; Dickman 2008). Cheetahs can be important threats to sheep and goats while grazing, but this may vary regionally (Ogada et al. 2003; Kolowski and Holekamp 2006). Cheetahs are not a threat to people.

### 2.3.2 African wild dog

**Figure 2.5 Several members of an African wild dog pack near a den in Ol Pejeta Conservancy, Kenya. © Andrew Jacobson**

The wild dog is lean and tall, with outsized ears, and a striking coat pattern. Their pelage is a variable pattern of black, yellow and white blotches with a black muzzle and white tail. They stand about 65-75 cm at the shoulder, and weigh from 18-28 kg with males slightly larger than females (Creel and Creel 2002).

Intensely social, wild dogs live in packs of up to 30 adults and yearlings (Macdonald and Sillero-Zubiri 2004) although 6-7 adults per pack is common (Creel and Creel 2002). The simplest pack structure is a set of related males and a set of related females, unrelated to the other (Creel and Creel 2002). Both males and females within a pack have dominance hierarchies. Described as obligate cooperative breeders, typically only one female breeds per pack and the entire pack helps care for the young (Creel and Creel 2002). Five adults is suggested as the minimum number in a pack to have good chance at successful recruitment (Courchamp, Clutton-Brock, and Grenfell 2000). Wild dogs typically disperse from their natal group as yearlings or two-year-olds, leaving in single-sex groups (McNutt 1996). Both sexes may emigrate, and they may disperse great distances looking for the opposite gender (Fuller et al. 1992; McNutt 1996). Dispersal within populations of 20-40 km is normal although straight line distance of up to 520 km are recorded (McNutt 1996; Masenga et al. 2016). Obligatory, and potentially long-distance dispersal of young wild dogs may compromise population persistence due to high likelihood of disperser death (Leigh et al. 2012), although it may also allow new population establishment (e.g. Laikipia) (Woodroffe 2011) and reduce chances of inbreeding depression.

Known for extremely large home ranges, they are more restricted during the denning season, returning to the den multiple times a day. Without the den as a focal area, wild dog home ranges may be up to 1,500 km$^2$ (Childes 1988) (roughly similar to the cheetah), although they averaged 440 km$^2$ in the Selous GR (Creel and Creel 2002) and 423 km$^2$ in Laikipia (Woodroffe 2011).

Wild dogs are primarily crepuscular (Cozzi et al. 2012), balancing the need for light (relying on sight and smell) and cool temperatures while hunting (they are threatened by overheating) (Hubel et al. 2016a). While hunting in packs, the level of cooperation, and style and energetic cost of hunting has recently come into question. Previously, wild dogs were thought to be coursing predators, using stamina and cooperation to outrun prey over medium to long distances. This had high energetic costs and left them susceptible to loss of food by competitors (Gorman et al. 1998). However, this style of hunt may be more common in short-grass plains habitats. In wooded habitats with more abundant and easily captured medium-sized prey, wild dogs may rely on short opportunistic hunts. This method of hunting requires less cooperation (spacing of dogs helps to flush more prey) and less energetic expenditure that suggests they are less threatened by kleptoparasitism (Jongeling and Koetsier 2014; Hubel et al. 2016a; Hubel et al. 2016b). Meat is bolted at the kill, with large quantities consumed very quickly (Creel and Creel 2002). Abundant, small and medium-sized ungulates are preferred prey like Thomson gazelle with a bimodal body mass range of 16-32 kg and 120-140 kg (Hayward et al. 2006c). Wild dogs rarely scavenge or drive other predators off kills (Kingdon 1977; Creel and Creel 2002).

Once nearly continuously distributed throughout sub-Saharan Africa except for tropical forests and deserts (Fanshawe, Frame, and Ginsberg 1991), wild dogs are resident in less than 10% of their historical range (IUCN/SSC 2007b; IUCN/SSC 2007a; IUCN/SSC 2012).

Described as habitat generalists, wild dog occupy short grass plains, semi-desert, bushy savannahs, and upland forest (Macdonald and Sillero-Zubiri 2004). However, they have also been observed in mangroves, montane forest, and montane moorland (IUCN/SSC 2007a). Listed by the IUCN as Endangered and in decline, approximately 6,600 individuals live in the world, with possibly as few as 1,400 breeding individuals (Woodroffe and Sillero-Zubiri 2012). Wild dogs live at relatively low population densities, even for large carnivores (Macdonald and Sillero-Zubiri 2004), The highest densities they achieve currently, such as in northern Selous, Tanzania, is around 4 individuals per 100 km$^2$ (Creel and Creel 2002). More typically, they live at densities of 1-2 per 100 km$^2$. Biomass of wild dogs is generally 1-2 orders of magnitude less than the biomass of lions or spotted hyenas and their densities correlate negatively with those of lion and spotted hyena (Creel and Creel 2002).

Intraguild predation via lions is the largest source of natural mortality in some populations (Woodroffe and Ginsberg 1999; Woodroffe et al. 2007a). Older reports seem to suggest greater parity in competition between wild dogs and other members of the large carnivore guild; more recent narratives suggest the wild dog as a "fugitive" species needing its extensive home range to avoid competitive effects from lion and spotted hyena (Kingdon 1977; Webster, McNutt, and McComb 2012). Similar to the cheetah, they use fine-scale avoidance of competitors (Webster, McNutt, and McComb 2012). However, more so than the cheetah, wild dogs also tend to avoid areas of higher lion densities utilizing both spatial and temporal partitioning (van Dyk and Slotow 2003; Vanak et al. 2013; Swanson et al. 2014).

However, anthropogenic mortality is the primary cause of death in most populations, even those inside protected areas (Woodroffe et al. 2007a). Problematically, human-caused deaths appear additive to natural causes, possibly undermining population persistence (Woodroffe et al. 2007a). Historical targeted killing is the primary reason for their endangerment today (Woodroffe and Ginsberg 1999) and in some areas wild dogs may be among the least tolerated or liked carnivores (Dickman 2008; Thorn et al. 2013). However, wild dogs suffer under a range of anthropogenic threats beyond direct killing, including road kill, and snaring (van der Meer et al. 2013). Diseases like canine distemper virus are an important source of mortality as well, implicated in some population declines (Creel and Creel 2002) and may be exacerbated by contact with domestic dogs.

Wild dog can be locally important predators of livestock but generally do not pose the greatest threat to livestock in an area. Certain factors may make wild dogs greater threats to livestock such as when prey densities are low or during the denning season when dogs are more geographically restricted (Woodroffe et al. 2005). In other areas wild dogs are present at such low densities they are scarcely mentioned as problem animals (Maddox 2003; Kissui 2008a). Overall, wild dogs typically attack during the daytime while livestock are grazing, focusing on sheep and goats, and very rarely attack cattle. Finally, wild dogs are not a threat to people.

### 2.3.3 Leopard

**Figure 2.6 A leopard pauses atop a rocky outcrop in Ruaha NP. © Andrew Jacobson**

Secretive, solitary, and spotted, the leopard is hard to detect in even minimal cover. With its reticent nature, dietary flexibility and generalist habitat requirements, the leopard has the largest geographic range of any felid (Bailey 1993; Sunquist and Sunquist 2002). The leopard ranges throughout Africa and Asia but I will focus on the leopard in sub-Saharan Africa. Leopards are medium-sized carnivores, with males larger than females and weighing between 34-69 kg (Macdonald and Loveridge 2010). Other than a female with her young, or temporarily during courtship, leopards are solitary (Sunquist and Sunquist 2002). Leopards are typically nocturnal although can be diurnal as well (Bailey 1993). Their activity pattern seems context dependent and can be more nocturnal around human presence or more diurnal to avoid competitors like lions or tigers (Sunquist and Sunquist 2002).

The leopard may utilize dense vegetation and broken terrain to aid hunting (Bailey 1993). A hunter that relies primarily on visual and auditory cues, the leopard prefers prey within a range of 10-40 kg (Sunquist and Sunquist 2002; Hayward, Henschel, et al. 2006a). However, the leopard's diet is extremely wide-ranging with over 90 prey species recorded in sub-Saharan Africa from arthropods to adult male eland (Macdonald and Loveridge 2010). Despite using a variety of hunting methods, the leopard is primarily a stalking predator although carrion is consumed as well. A stalk brings the leopard within striking distance of its prey, so the final charge is relatively short; this averages 65 m in the Kalahari desert (Sunquist and Sunquist 2002). Leopards are susceptible to kleptoparasitism from larger carnivores and will move their prey great distances, or cache them in trees to decrease confrontation (Sunquist and Sunquist 2002).

The leopard is very adaptable and tolerates a range of habitats from deserts and mountains to rainforest and coastal scrub (Sunquist and Sunquist 2002). Commonly associated with some type of forest cover, they are found in essentially all habitats with annual rainfall above 50 mm but can navigate areas with even less rainfall by following watercourses (Nowell and Jackson 1996). Leopard densities vary widely from less than one per 100 $km^2$ to over 30 per 100 $km^2$ due to variations in habitat, prey availability, and degree of threat (Jacobson et al. 2016). The highest densities in Africa are recorded from protected East and Southern African mesic woodland savannahs (Macdonald and Loveridge 2010).

Leopard home ranges vary widely, dependent on habitat and prey availability. Home range in mountainous areas may be ~500 $km^2$ or as small as 15 $km^2$ for a female in savannah woodland (Nowell and Jackson 1996). Female leopard home ranges are smaller and partially overlap with male ranges (Sunquist and Sunquist 2002). Leopards travel widely, regularly patrolling the home range. Dispersal timing, distances and patterns are unknown although males are thought to disperse at an earlier age and move farther from natal range than females (Sunquist and Sunquist 2002).

Although tentative population estimates exist for all other regions except Africa (encompassing eight other subspecies), there is no accepted estimate for Africa (Stein et al. 2015). Leopards are extremely difficult to count, and prior African estimates are believed to be wildly inaccurate. However, the leopard is critically endangered in some parts of its range outside Africa, and has lost up to 99% and 86-95% of its range in North and West Africa respectively (Jacobson et al. 2016). The leopard is doing relatively better in other regions of Africa and has lost at most 50% of its range in southern Africa. Leopards can live on the edge or even within urban areas but it is thought competition from humans and from other carnivores are the only ways to limit them (Sunquist and Sunquist 2002).

Primary threats to the leopard include habitat loss, prey base depletion, illegal trade and direct killing due to real or potential HWC (Macdonald and Loveridge 2010; Jacobson et al. 2016). Leopards can be important livestock killers and man-eating is a problem in parts of its range (Sunquist and Sunquist 2002). In terms of livestock, leopards may or may not be important contributors to livestock depredations (Myers 1976; Kolowski and Holekamp 2006; Thorn et al. 2013). However, a proclivity for "surplus killing" can engender negative feelings for this cat among pastoralists (Myers 1976). Leopards seem to prefer sheep and goats over cattle and will attack both at pasture and the boma (Kolowski and Holekamp 2006; Woodroffe et al. 2007b, Dickman 2008). Leopards are a threat to humans; a single leopard in India killed 125 people (Sunquist and Sunquist 2002). However, man-eating seems relatively less common in Africa than elsewhere (Myers 1976), and attacks are less frequent and severe than those by lions (Treves and Naughton-Treves 1999; Dickman 2008).

### 2.3.4 Lion

**Figure 2.7 A male lion stands out in the Tarangire grasslands. © Andrew Jacobson**

The lion is the largest and most dominant carnivore in sub-Saharan Africa (Kingdon 1977) averaging 172 kg for males and 151 kg for females (Kingdon 1977). They can reach over a meter high at the shoulder and three meters in length. Some lions are nomadic but most live in prides. A pride is centred on a single or coalition of male lions with single to multiple adult females and their cubs (Schaller 1972). Nomadic individuals or pairs of lions will traverse large landscapes. Prides, ranging from three to 35 individuals (van Orsdol, Hanby, and Bygott 1985) have well-defined ranges. Sub-adult males will leave the pride before reaching maturity and are generally forced out by the resident (or incoming) male. The young males lead a nomadic life while they mature before attempting to take over a pride of their own (Schaller 1972). Females tend to stay within the pride with their mothers and sisters; these female lineages may stay in the same place for generations (Schaller 1972). Male dispersal can be nearly 350 km although the average distance may be closer to 115 km (Dolrenry et al. 2014).

Known to eat anything from mice and tortoises, to giraffe and elephant, the lion nonetheless prefers prey within a range of 190-550 kg (Kingdon 1977; Hayward and Kerley 2005). Generally the lion is an opportunistic feeder, with smaller prey eaten by single animals and larger prey requiring multiple individuals to hunt. Diet varies seasonally with the abundance of prey. Lions are primarily stalking predators (Schaller 1972) but they also scavenge and steal carcasses. Lions are predominantly nocturnal with two peaks of activity, near sundown and before sunrise, but may be active at any time of day (Schaller 1972; Cozzi et al. 2012).

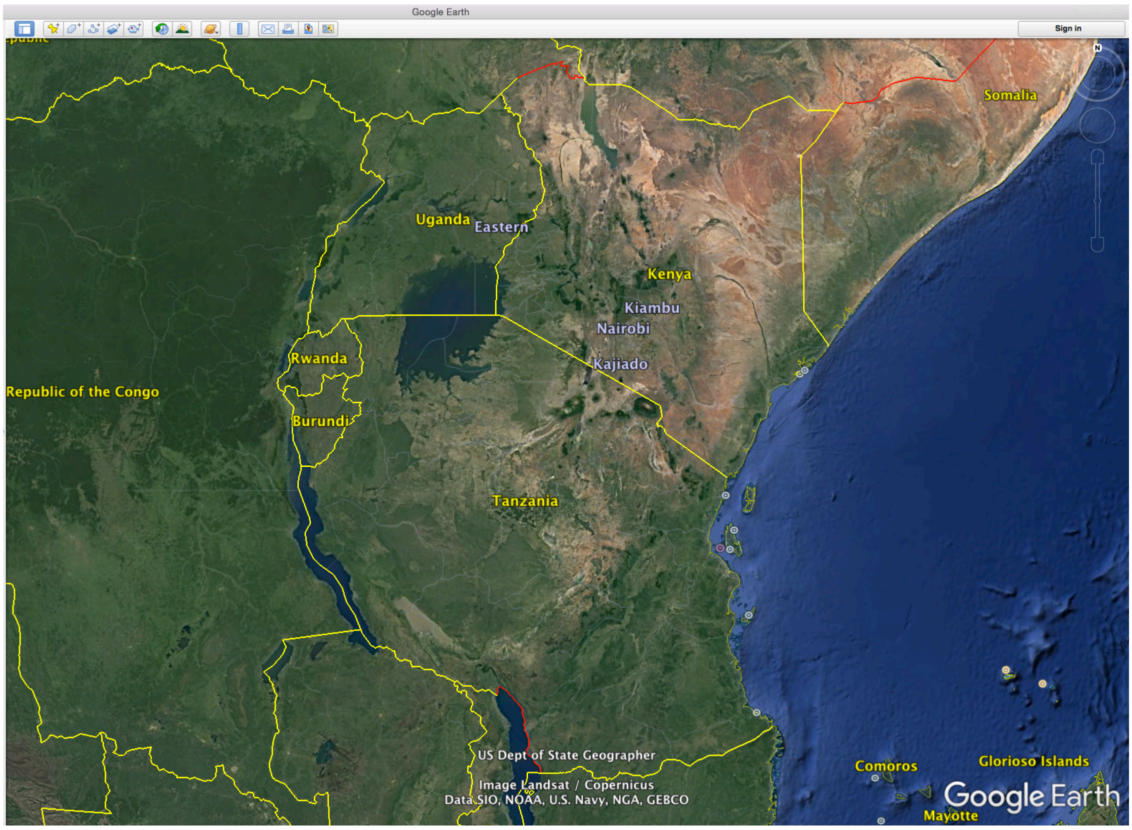
Originally distributed throughout Africa, the Middle East and parts of Asia, the lion is now restricted to sub-Saharan Africa (other than a small population in the Gir forest of India) (Bauer, Packer, et al. 2015). A habitat generalist, the lion can survive in semi-deserts, savannahs, woodlands and even montane areas, although their greatest concentrations mirror concentrations of plains animals (Sunquist and Sunquist 2002). Home range size and pride size are related to lean season prey biomass (van Orsdol, Hanby, and Bygott 1985). These factors in turn are associated with lion density. In more productive habitats of East Africa, home range can be as small as 65 km$^2$ and densities can reach 55 lions per 100 km$^2$ (Sunquist and Sunquist 2002). However, at the other extreme, home ranges were >2,000 km in Etosha NP, Namibia, and pride size averaged only seven individuals with densities around 1.5 – 2 lions per 100 km$^2$ (Sunquist and Sunquist 2002). Lions are widely killed and their range much reduced to roughly 3,400,000 km$^2$ (Riggio et al. 2013). Having lost nearly their entire habitat in West Africa (Henschel et al. 2014), lion range is constricting within East and Southern Africa as well (Riggio et al. 2013).

Fewer than 35,000 lions remain in sub-Saharan Africa, with particularly steep rates of decline outside protected areas (Riggio et al. 2013). Listed by IUCN as Vulnerable with a declining population trend, population strongholds for the species are in Southern and Eastern Africa; Tanzania alone is estimated to hold more than 1/3rd of the world's population (Riggio et al. 2013). Anthropogenic mortality is the greatest threat to lions currently, both inside and outside protected areas, although lions die from a variety of interspecific and intraspecific causes (Ray, Hunter, and Zigouris 2005). Diseases such as canine distemper can also play an important role in population dynamics (Packer et al. 1999)

Major contributors to HWC, lions are threats to livestock and people. Lions can cause significant losses to livestock, including sheep, goats and cattle. They predate livestock both while grazing during the day and while penned at night in the boma (Ogada et al. 2003; Woodroffe et al. 2007b). Some lions are also man-eaters; for instance over 1,000 people were killed in Tanzania during the 1990s and 2000s (Kushnir et al. 2010).

# Chapter 3

## A novel approach to mapping land conversion using Google Earth with an application to East Africa



A screenshot of East Africa from Google Earth.

* This chapter is published as: Jacobson, A., J. Dhanota, J. Godfrey, H. Jacobson, Z. Rossman, A. Stanish, H. Walker, J. Riggio. 2015. A novel approach to mapping land conversion using Google Earth with an application to East Africa. Environmental Modelling and Software 72: 1-9. DOI: 10.1016/j.envsoft.2015.06.011.

This version differs from the published version due to minor changes during the Viva process.

# Chapter 3 A novel approach to mapping land conversion using Google Earth with an application to East Africa

## Abstract

Effective conservation planning relies on accurate identification of anthropogenic land cover. However, accessing localized information can be difficult or impossible in developing countries. Additionally, global, medium-resolution land use land cover products can provide conflicting information and may be insufficient for conservation planning purposes at the scale of a country or smaller. A new tool, GE Grids, is introduced to bridge this gap and to address Research Aim I. This tool creates an interactive user-specified binary grid laid over Google Earth's high-resolution imagery. The grid size can be adjusted by the user and here was set at ~ 1 km$^2$, an appropriate scale for use in the distribution modeling to follow. Using GE Grids, anthropogenic land conversion was identified across East Africa and compared against available land cover products. I found that nearly 30% of East Africa is converted to anthropogenic land cover. In addition, the two highest-resolution comparative datasets have the greatest agreement with the GE Grids' product at the regional extent, despite having as low as 44% agreement at the country level. GE Grids is intended to complement existing remote sensing products at local scales.

## 3.1 Introduction

Satellite remote sensing improves our understanding of earth processes and our monitoring of natural and anthropogenic changes (Pettorelli, Safi, and Turner 2014). However, while the financial burdens of obtaining these data are decreasing, remote sensing products still require specialized, and often expensive, training and software (Stensgaard et al. 2009; Pettorelli, Safi, and Turner 2014). Accessing specialized training and software may present particular hurdles to research and conservation programs, particularly those in developing countries. Error in selecting, downloading, processing, and analyzing remote sensing datasets may furthermore result in inappropriate recommendations and conclusions (Watson et al. 2015), particularly for ecological applications (Kerr and Ostrovsky 2003). Inappropriate analyses may result in missed opportunities, or squandered resources (Wilson et al. 2005). There is thus a need for easily created, inexpensive, locally-accurate products that can confidently be used in conservation planning (Watson et al. 2015).

Land use land cover (LULC) data sets describe how humans use land (land use) as well as the physical features that cover the earth's surface (land cover). Identifying the location, intensity, and extent of human activities is essential to conservation planning (Hansen et al. 2000). In LULC products, anthropogenic land cover is typically classified as either cropland or urban extent. However, identification of these land uses is challenging and varies greatly across products (Potere and Schneider 2007; Fritz et al. 2011; Vancutsem et al. 2012). Traditional remote sensing classification approaches require grouping spectral signatures and subsequent accurate discrimination between groups i.e. land cover types (Pfeifer et al. 2012c), although emerging remote sensing techniques such as object-based classification reduce this reliance on unique spectral properties by allowing probabilistic class descriptions (Blaschke 2010). Classification may be relatively easy where vegetated landscapes are homogenous and extensive e.g. some croplands. However, in heterogeneous landscapes with small, patchy agricultural fields, discriminating cropland from natural land cover can be difficult (Tchuenté, Roujean, and De Jong 2011; Vancutsem et al. 2012).

One possible solution to classification problems is to use high-resolution image data (pixel resolution of 10 m or better, whereas moderate resolution data is between 10 and 250 m (Pfeifer et al. 2012c)) like that available through Google Earth. Google Earth is a free, easy-to-use program owned by Google Inc. that allows access to sub-meter pixel resolution data for over a quarter of the world's landmass and three-quarters of the global population (Google 2014). Google Earth's high-resolution data are useful as a platform for validating datasets (Fritz et al. 2011), used previously with urban extent (Schneider, Friedl, and Potere 2009) and land cover (Defourny et al. 2008). Another tool, Global Mapper, was developed for mapping global land cover with the aid of Google Earth (Gong et al. 2013), however, this tool is not widely available and could not be tested.

While Google Earth has the potential for wider use in scientific literature, particularly in LULC analyses (Potere 2008), one prominent challenge is that native analysis functions in Google Earth are minimal (Yu and Gong 2012), limited to drawing points, lines, and polygons. I previously used the polygon drawing feature to identify anthropogenic land use in West Africa (Riggio et al. 2013) and Mozambique (Jacobson et al. 2013). The time-consuming nature of these analyses spurred the creation of a new tool, "GE Grids", the first free, customizable creator of raster datasets for use with Google Earth. GE Grids creates a user-defined, interactive grid (raster) overlaid on Google Earth image data. This tool circumvents expensive, specialized programs and knowledge, and enables easy use of Google Earth's high-resolution data to create localized datasets. I use GE Grids to document anthropogenic land conversion in East Africa, a region of significant conservation importance (Myers et al. 2000; Ray, Hunter, and Zigouris 2005; Jenkins, Pimm, and Joppa 2013) experiencing rapid human population growth (UN 2013).

## 3.2 Methods

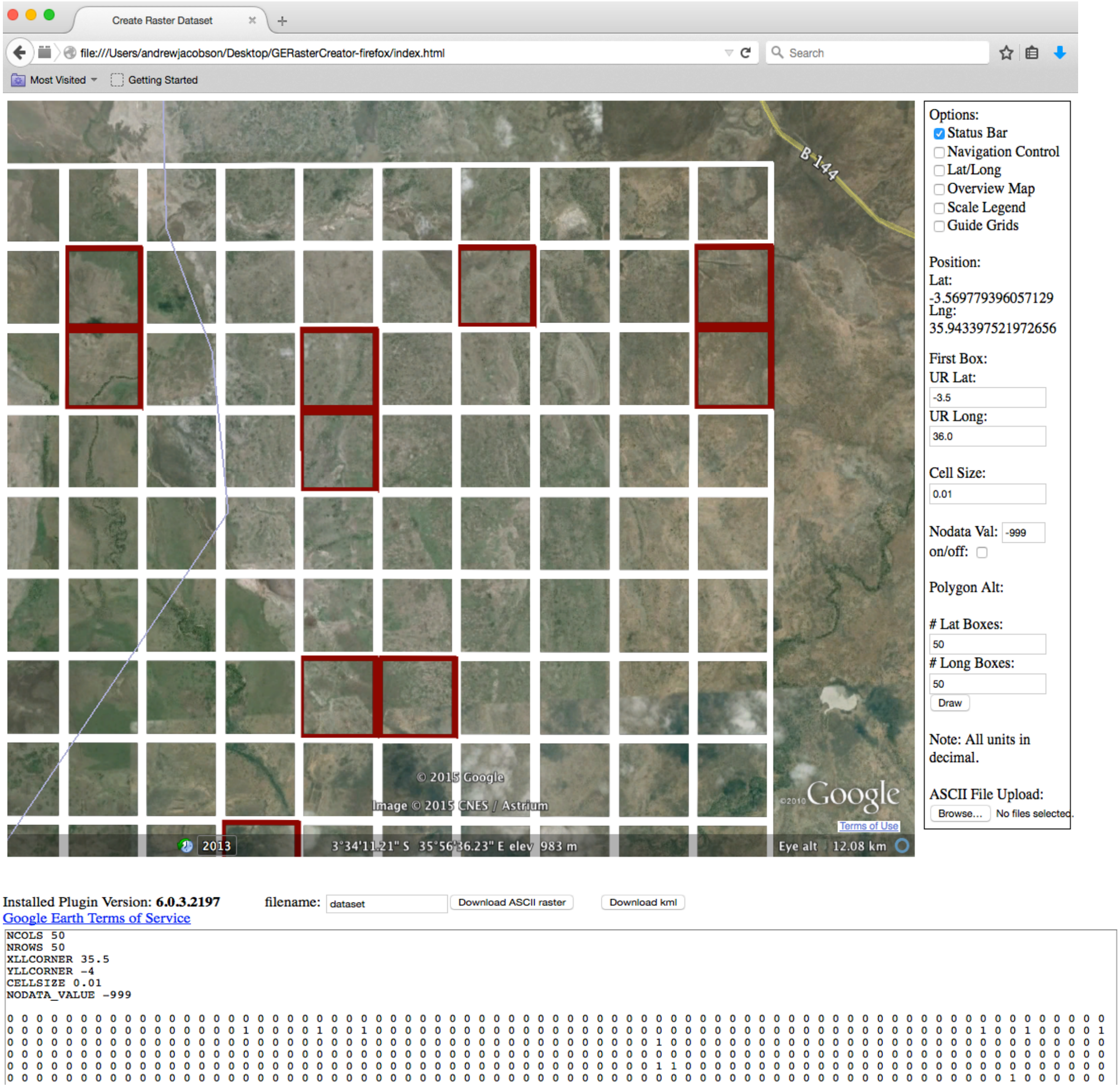
### 3.2.1 GE Grids program design and workflow

GE Grids is a browser-based application that provides a customized interface to map land cover using satellite and aerial data available in Google Earth. The application relies on the free Google Earth plug-in and Google's public application programming interface (API) as well as a plugin called "filesaver.js" written by Eli Grey and available on GitHub. The program is written in JavaScript and tested in the Google Chrome and Mozilla Firefox web browser.

The user interface of GE Grids is a combination of generic controls provided by Google Earth and input parameters for creating a grid. Controls allow the user to navigate around the Google Earth imagery and to enter the information necessary to specify or "draw" a grid (Figure 3.1). Options include: the latitude and longitude of the upper right-hand corner coordinates of the grid, the size of each cell (in Degrees – a function of Google Earth's use of the WGS 1984 coordinate system), and the number of cells on each axis. Each cell can be visually divided into 9 minor grids (3x3) using the "Grid Guides" function to ease the classification of heterogeneous cells. Although the study was done using square grids, the program supports any number of cells per side.

The overall work flow is summarized in Figure 3.2. Once the user creates a grid using the "Draw" feature, they can interact with the grid by clicking on the grid edges to change their color from white to red. This corresponds with a data value change from the default 0 to 1; or if the No Data function is clicked on, to -999 (or any other value chosen by the user). The result can be downloaded as a text file in ASCII raster format for import into GIS software or as a KML file to upload into Google Earth. The ASCII file can also be re-uploaded into GE Grids for editing and error checking.

This tool meets the legal requirements of the Google Earth API Terms of Service. GE Grids is free to all users, does not alter or blur imagery from Google Earth, and allows attribution of the image data to remain visible. In using this tool, users are also agreeing to abide by Google's Terms of Service. Importantly, the image data itself and the output from GE Grids should not be used for commercial purposes without further consent from Google. More information can be found in the ReadMe file, included in the zip file download of this tool.



Options:
☑ Status Bar
☐ Navigation Control
☐ Lat/Long
☐ Overview Map
☐ Scale Legend
☐ Guide Grids

Position:
Lat:
-3.569779396057129
Lng:
35.943397521972656

First Box:
UR Lat: -3.5
UR Long: 36.0

Cell Size: 0.01

Nodata Val: -999
on/off: ☐

Polygon Alt:

\# Lat Boxes: 50
\# Long Boxes: 50
Draw

Note: All units in decimal.

ASCII File Upload:
Browse… No files selected.

Installed Plugin Version: **6.0.3.2197**
Google Earth Terms of Service

filename: dataset
Download ASCII raster
Download kml

```
NCOLS 50
NROWS 50
XLLCORNER 35.5
YLLCORNER -4
CELLSIZE 0.01
NODATA_VALUE -999

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 1
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0
```

**Figure 3.1 Screenshot of the GE Grids program interface showing the control options, a portion of the interactive grid, and the layout of the ASCI text file.**

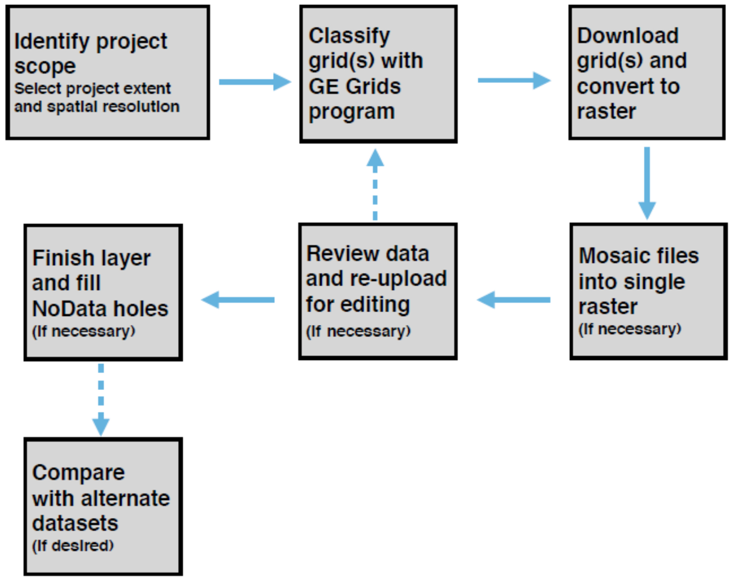


**Figure 3.2 User interaction diagram depicting the workflow using GE Grids.**

### 3.2.2 Application of GE Grids to East Africa

GE Grids was used to document anthropogenic land conversion throughout East Africa at a resolution of 0.01° (~1 km at the equator). Each run of GE Grids covered a 50 x 50 grid cell square (0.5° x 0.5°; ~2500 km$^2$). Each grid cell was visually evaluated for the presence of anthropogenic land cover (Figure 3). A cell was classified as 'converted' if 50% or more of the land was converted to human land cover (including agriculture, urban development, industry, mines, roads, and housing units such as bomas). Half of the grid cell was a visually easily identifiable threshold. These land covers are generally easy to visually distinguish from natural land cover although agriculture can be more challenging. Agriculture was distinguished from native vegetation by its geometric pattern either within the field itself or its border, and by the relative lack of "texture" in croplands compared to native vegetation. Tree plantations for either crops or timber were classified as converted if geometric patterns were detected; otherwise, they were generally indistinguishable from native vegetation. Deforested, degraded or grazed lands were not considered as converted. Grid cells partially covered by water were evaluated on the basis of the terrestrial land cover.

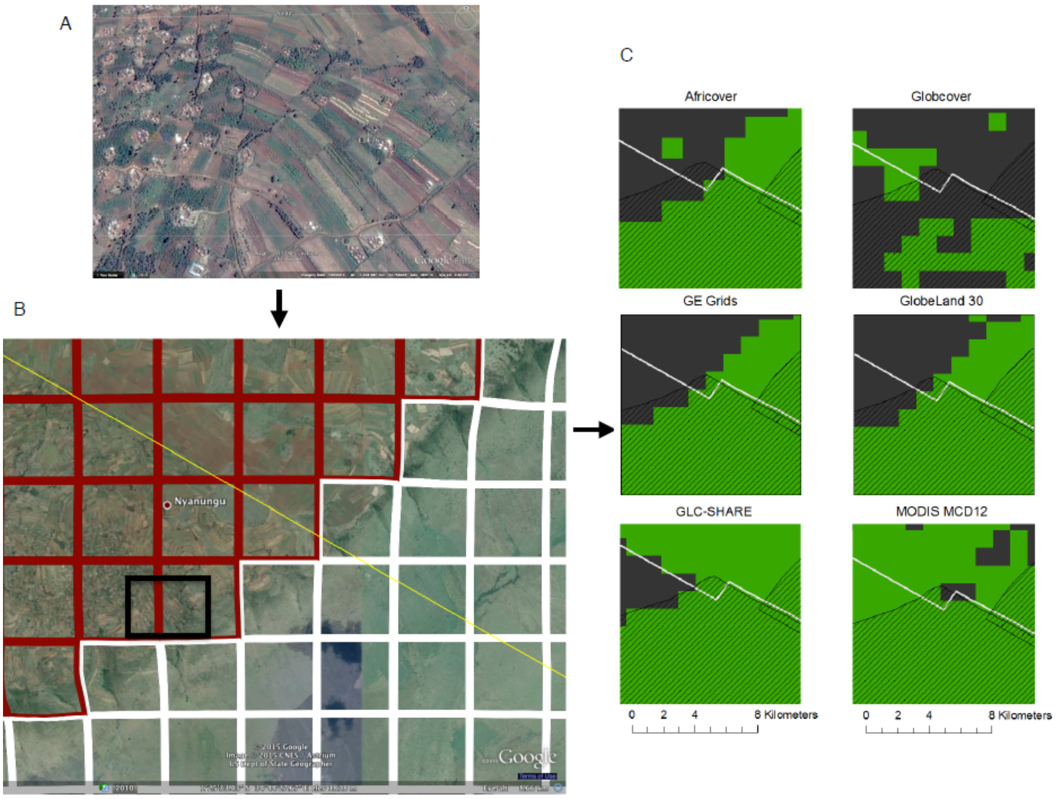


**Figure 3.3 The upper left image (A) is a screen capture from Google Earth of the western edge where Masai Mara NR (Kenya) and Serengeti NP (Tanzania) intersect. Fields and houses are clearly visible. Image B is the classification of this area in the GE Grids program, with a black box depicting the area of screen capture. Towns and fields occupy the land on top of the Great Rift Valley escarpment while natural vegetation lies below the escarpment. The clustered six images in C all show an identical area corresponding to the extent shown in B. These illustrate the differences between various dataset's depictions of anthropogenic land conversion. Dark gray is anthropogenic land conversion, green represents natural vegetation, the light grey line is the country border, and hashed regions are protected areas.**

If identification of land cover was impaired, primarily due to moderate resolution imagery (e.g. Landsat imagery) or cloud cover, or for any reason it could not be determined if >50% of a grid cell was converted to anthropogenic land cover; it was classified as No Data. These No Data areas were later filled outside of the GE Grids program.

After evaluating all cells in a 50 x 50 grid, the resulting file was downloaded in ASCII text format. Each text file was then imported into ArcGIS 10.2.1 (ESRI 2014) and converted to a raster. The individual files were mosaicked together on a per-country basis. No Data cells were filled using WorldPop, a human population density dataset with one kilometer resolution (Linard et al. 2012). To do so, I calculated the greatest percentage overlap between GE Grids and various densities of human population at five people per km$^2$ increments (WorldPop; Linard et al. 2012) on a per country basis. I counted areas of agreement, where a grid cell was either below a threshold human density and had no land conversion, or was above the threshold and had land conversion. The population density with the greatest percent overlap was then used as a threshold to fill in the "holes." The No Data grid cell was thus classified as anthropogenic land conversion or not based on whether it was above or below the human population density threshold value. Each country was then merged and clipped to remove islands in the Indian Ocean. Finally, the lakes class from the Global Lakes and Wetlands (GLWD v3; Lehner and Döll 2004) was overlaid to give context. The resulting dataset is a binary land classification layer of anthropogenic land conversion versus natural habitat.

All contributors, except the programmer (Andrew Stanish), evaluated a minimum of 50 grids each. Each grid was evaluated at least once and no grids overlap. I checked for consistency across grids in ArcGIS, and visually compared results with WorldPop and protected area extents (UNEP and IUCN 2014). Where concern arose due to discrepancies (e.g. stark differences between grids, or between GE Grids and WorldPop or protected area boundaries), grids were re-evaluated in GE Grids.

To illustrate the repeatability of a land cover classification using GE Grids I compared the results of the classification of a grid of 2500 cells (50 x 50). I chose a grid containing a mixture of conversion and natural lands, along with a combination of high and medium resolution imagery. For intra-user consistency, I classified this same grid a total of five times. For inter-user consistency, all contributors except the programmer classified this grid. Evaluated grids were compared in ArcGIS.

### 3.2.3 Product comparisons

Five datasets, one regional and four global, were spatially compared after the hole-filling process with the GE Grids classification: Africover (Alinovi, Di Gregorio, and Latham 2000), GlobeLand 30 (National Geomatics Center of China 2014), GLC-SHARE Beta Release 1.0 (Latham et al. 2014), Globcover v2.3 (Bontemps et al. 2011), and MODIS land cover MCD12Q1 (Friedl et al. 2010). Table 3.1 provides a summary of the datasets and they are briefly described below:

- Operated by the FAO, the Africover program created an East African regional land cover data set derived from Landsat imagery in 2000. It is a vector data set (all others are raster), so conversion to raster was also necessary after mosaicking countries together. The original data set was spatially re-aggregated to give priority to the agriculture class by reassigning polygons below a certain size threshold to agriculture. This resulted in a binary data set representing the extent of agricultural land use. This class was combined with the artificial surface class.
- China launched the Global Land Cover project in 2010 and released datasets for the year 2000 and 2010 in 2014. GlobeLand 30 is the world's first 30 m, global LULC dataset. Agriculture and urban classes were combined.
- The FAO released a new global LULC dataset, GLC-SHARE, in 2014 at ~1 km$^2$. A variety of datasets were "harmonized" to give a value per grid cell for each of 11 land cover classes such that all classes summed to 100% per cell. Agriculture and urban areas were summed and if they added up to be greater than 50%, the cell was classified as converted. Africover was a primary but not sole data contributor to GLC-Share in East Africa.
- The European Space Agency coordinates the GLOBCOVER project, a global land cover mapping initiative. GlobCover v2.3 is the most recent version. Four classes for Globcover were merged, post-flooding or irrigated croplands, rain-fed croplands, mosaic cropland (50-70%) and native vegetation (20-50%), artificial surfaces, and a fifth was evaluated, mosaic vegetation (50-70%) and cropland (20-50%).
- MODIS MCD12Q1 is an annually produced land cover product from the two NASA MODIS instruments aboard the Terra and Aqua platforms. The 2012 product was downloaded via the Land processes Distributed Active Archive Center server (NASA LP DAAC 2012) and mosaicked, clipped, and projected using GDAL (GDAL 2013). Two MODIS classes were merged, croplands, and urban/built-up, with a third evaluated, cropland/natural vegetation mosaic.

**Table 3.1 Summary of comparative land cover datasets.**

| Dataset (website) | Reference | Sensor | Year of data collection | Spatial resolution | Total # of classes (# related to anthro[a]) | Accuracy assessmen overall[b] (cropland) |
|---|---|---|---|---|---|---|
| Africover (www.glcn.org/activities/africover_en.jsp) | Alinovi, Di Gregorio, and Latham 2000 | Landsat TM | Burundi 1999,<br>Kenya 1999,<br>Tanzania 1997,<br>Rwanda 1999,<br>Uganda 2000-2001 | 30 m; spatially aggregated to polygon | Condensed to 6 (2) | NA |
| GlobeLand 30 2010 (www.globallandcover.com) | National Geomatics Center of China 2014 | 30 m multispectral images (e.g. Landsat TM, Landsat ETM+, HJ-1) | 2008 - 2011 | 30 m | 10 (2) | 83.5% (83.1%) |
| GLC-SHARE Beta Release 1.0 (www.glcn.org/databases/lc_glcshare_en.jsp) | Latham et al. 2014 | Varied | Burundi, Tanzania, Rwanda, and Uganda 2001;<br>Kenya 2010 | 30 arc seconds (~ 1 km) | 11 (2) | 80% |
| Globcover v2.3 (due.esrin.esa.int/globcover/) | Bontemps et al. 2011 | MERIS FR | 2009 | 300 m | 22 (5) | 58% |
| MODIS MCD12 Q1 collection 5; year 2012 (https://lpdaac.usgs.gov/products/modis_products_table/mcd12q1) | Friedl et al. 2010 | MODIS, bands 1-7 & EVI | 2012 | 500 m | 17 (3) | 75% (77%) |

[a] Number of classes related to anthropogenic land conversion
[b] Accuracy assessments for all land cover classes and cropland only as specified in their product descriptions.

For comparison purposes, all products were standardized to raster datasets at 0.01° resolution, then clipped and aligned to identical geographic extents. Each dataset was re-sampled according to the majority land cover within the 0.01° cell. An important component of prior land cover comparisons was the standardization of classes before comparison as products used different land cover categories and definitions (e.g. the International Geosphere Biosphere Project, IGBP, and the Land Cover Classification

System, LCCS; McCallum et al. 2006). This is trivial here as all and only anthropogenic classes were of interest. However, both GlobCover and MODIS MCD12 did have a class representing a mosaic of both natural and anthropogenic land covers. Therefore these datasets were compared with and without these mosaic layers.

## 3.3 Results

Using GE Grids, 1,479,121 cells for East Africa were classified as either predominately natural or converted to anthropogenic land cover (Figure 3.4). Nearly 30% (29.77%) of the region has been converted to anthropogenic land cover although this varies greatly by country (Table 3.2). Burundi and Rwanda have the highest proportions of anthropogenic land cover at 85.99% and 82.27% respectively. Kenya contains the greatest percentage of land still in a natural state (82.65%), mostly within the nation's arid north. Only 3.74% of cells were No Data and filled via human population density on a country-by-country basis (Table 3.3). The highest percent agreement varied by country and was 20 people/km$^2$ in Tanzania at 82%, 35 people in Burundi at 84%, 40 people in Rwanda at 91%, 60 people in Uganda at 81%, and 60 people in Kenya at 93%.

The consistency of a single classified grid (50 x 50 cells) from all contributors was 82.76%. Agreement between the five replicates of a single user was higher at 94.6%.

Africover has the highest overall agreement (87.63%) with the GE Grids product (Table 3.4). Africover classifies the least amount of land as natural that I found converted. GLC-Share classifies the least amount of land as converted that I found natural. Both GlobCover and MODIS MCD12 have higher agreement with GE Grids when mosaic cropland/natural vegetation land cover classes are combined with anthropogenic land cover classes.

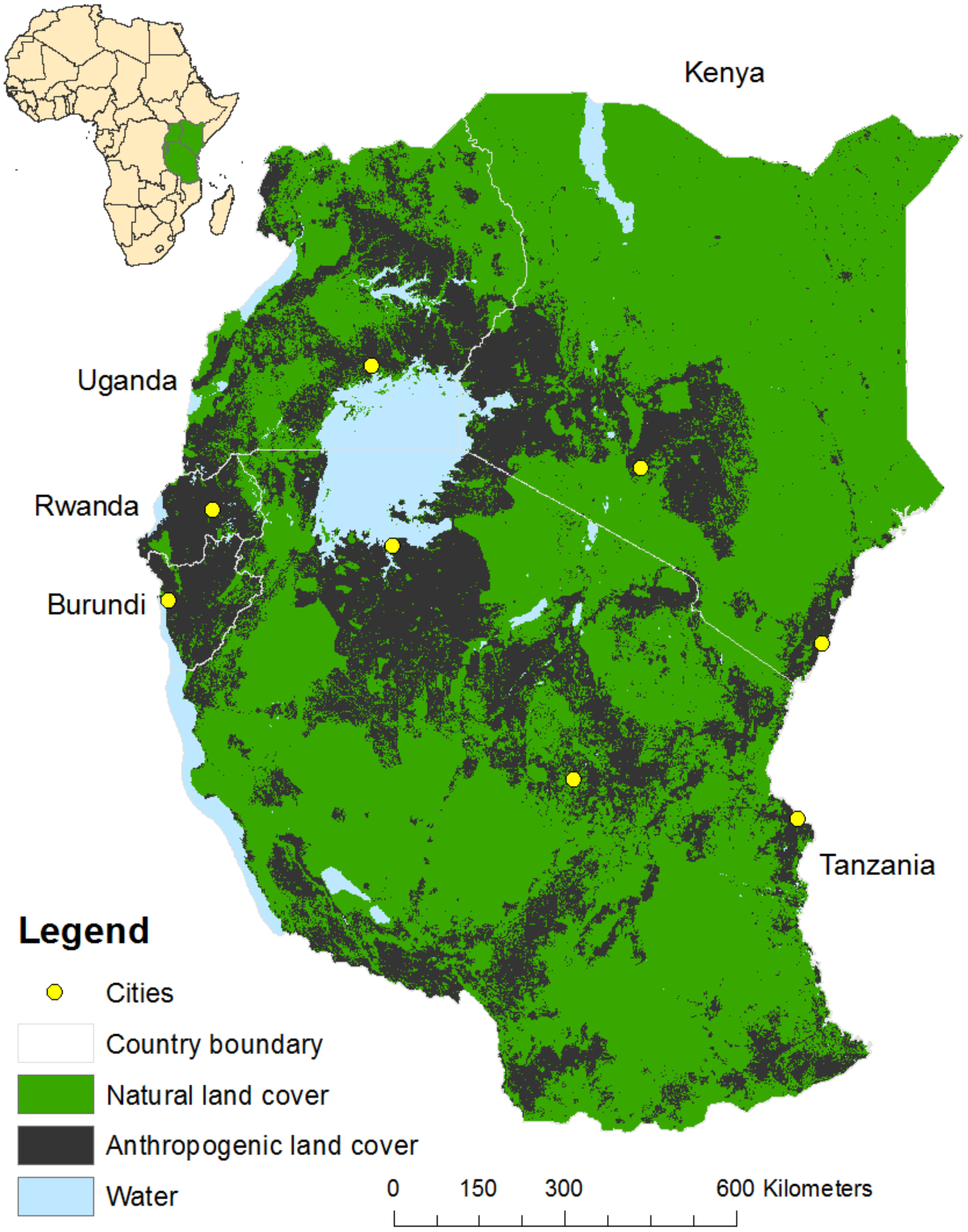


**Figure 3.4** GE Grids land cover classification map of East Africa. Cities displayed are either country capitals or have populations exceeding 500,000 people.

**Table 3.2 Percent terrestrial land cover classified as natural or anthropogenic by country after filling No Data holes.**

| Country | % Natural | % Anthropogenic | % No Data |
|---|---|---|---|
| Burundi | 14.01 | 85.99 | 1.92 |
| Kenya | 82.65 | 17.35 | 1.37 |
| Rwanda | 17.73 | 82.27 | 0.13 |
| Tanzania | 68.44 | 31.56 | 5.12 |
| Uganda | 56.54 | 43.46 | 3.74 |
| East Africa | 70.23 | 29.77 | 3.74 |

**Table 3.3 Agreement between GE Grids' classification of anthropogenic land cover and a range of thresholded human population density values for Tanzania. The greatest agreement for this country occurred at 20 people per km$^2$.**

| | Number of cells | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| People per km$^2$ | 10 | 15 | 20 | 25 | 30 | 35 | 40 | 45 | 50 |
| Both agree - not converted | 399,697 | 429,376 | 446,721 | 457,877 | 466,107 | 472,503 | 477,378 | 480,930 | 483,779 |
| Disagree - HPD converted | 103,504 | 73,825 | 56,480 | 45,324 | 37,094 | 30,698 | 25,823 | 22,271 | 19,422 |
| Disagree - GE Grid converted | 45,172 | 62,131 | 75,999 | 88,095 | 98,841 | 108,252 | 116,622 | 123,858 | 130,073 |
| Both agree - converted | 175,233 | 158,274 | 144,406 | 132,310 | 121,564 | 112,153 | 103,783 | 96,547 | 90,332 |
| % Agreement | 79.5% | 81.2% | **81.7%** | 81.6% | 81.2% | 80.8% | 80.3% | 79.8% | 79.3% |

**Table 3.4 The percent agreement between the GE Grids land cover classification of East Africa and those from comparative datasets. GlobCover and MODIS MCD12 both have an additional class of mosaic cropland/native vegetation that is added in the (+) comparison and absent in the (-). (NOTE. All Figures and Tables are only with the + version of these datasets due to their higher agreement.)**

| GE Grids v. Dataset | Africover | GlobeLand 30 | GLC-SHARE | GlobCover (+) | GlobCover (-) | MCD12 (+) | MCD12 (-) |
|---|---|---|---|---|---|---|---|
| Natural-Natural | 64.67 | 69.84 | 70.57 | 58.63 | 69.01 | 67.80 | 71.25 |
| Converted-Converted | 22.96 | 17.05 | 12.38 | 15.12 | 3.79 | 8.08 | 1.20 |
| Natural-Converted | 7.28 | 2.11 | 1.38 | 13.32 | 2.94 | 4.15 | 0.70 |
| Converted-Natural | 5.09 | 11.00 | 15.67 | 12.93 | 24.26 | 19.97 | 26.85 |
| Total % Agreement | 87.63 | 86.88 | 82.95 | 73.75 | 72.79 | 75.88 | 72.45 |
| Unweighted Kappa Statistic | 0.734 | 0.692 | 0.585 | 0.444 | 0.304 | 0.408 | 0.255 |

A spatial comparison between the GE Grids' classification and comparative products suggests that all global datasets had difficulty identifying development in southeastern Burundi and coastal regions of Kenya and Tanzania (Figure 3.5). On a per country basis, Burundi had the lowest rate of agreement from all comparative products with GE Grids, while Kenya had the highest (Table 3.5).

**Table 3.5 Percent agreement between GE Grids and each comparative land cover product analyzed per country.**

| | GlobeLand 30 | Africover | GLC Share | Globcover | MODIS | Average |
|---|---|---|---|---|---|---|
| Burundi | 44 | 86 | 62 | 62 | 62 | 63 |
| Rwanda | 67 | 89 | 83 | 68 | 79 | 77 |
| Uganda | 82 | 85 | 81 | 71 | 72 | 78 |
| Kenya | 95 | 85 | 90 | 77 | 85 | 86 |
| Tanzania | 84 | 83 | 79 | 72 | 70 | 77 |

The software design and approach satisfied the goals of this case study in East African land cover. GE Grids enabled evaluation of reasonably sized grids (50 x 50) using high-resolution satellite data. Since this is a manual process much larger grids would become burdensome. However, the study area was likely too large as it required ~600 individual runs of the program. However, the ASCII text files were readily converted into raster grids and, despite the large number of grids, were easily mosaicked without slivers or gaps.

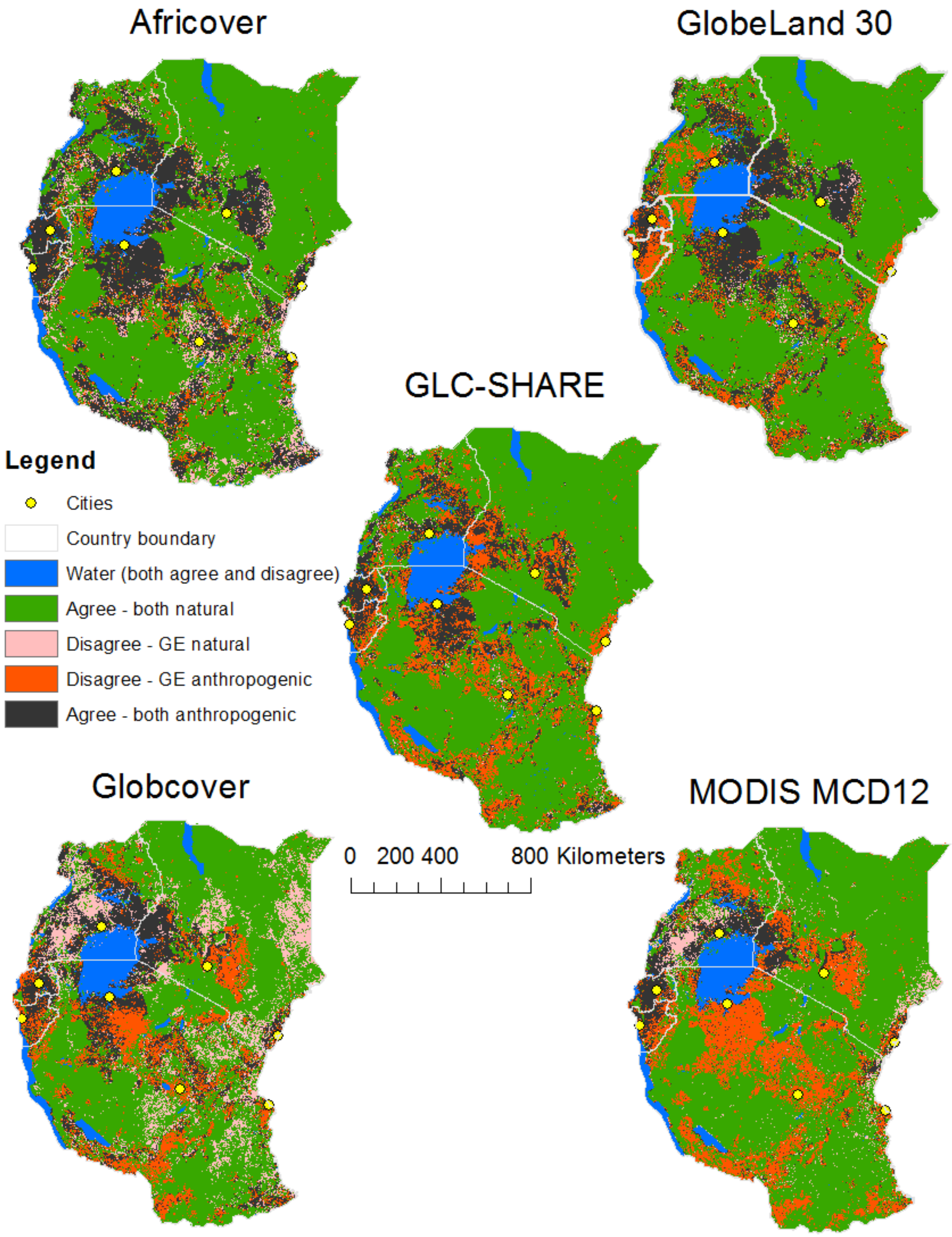


**Figure 3.5** **A spatial comparison of anthropogenic land cover in East Africa between GE Grids and comparative products.**

## 3.4 Discussion

### 3.4.1 Review of results

I introduce a new tool, GE Grids, and with it create a binary classification layer of anthropogenic land conversion versus natural habitat in East Africa. Although East Africa is a region of conservation significance (Jenkins, Pimm, and Joppa 2013), there is substantial disagreement over the extent of anthropogenic land cover among existing products (Fritz et al., 2011, 2010; Hannerz & Lotsch, 2008; Vancutsem et al., 2012). Accurately identifying this extent would provide a useful metric for previous or future change analyses. I found no existing peer-reviewed research measuring land cover change at the country or regional scale. Unsurprisingly, Burundi and Rwanda have the highest percentage of land conversion, as these countries also have the highest human population densities. Kenya and Tanzania have the lowest percentage of land conversion and population densities.

The reliability of this tool is important to consider as GE Grids relies on manual classification of image data. Using a test grid, I show that this process is highly repeatable (83% overlap between contributors) and even higher for a single user (95%). This suggests that multiple contributors following strict rules can produce output consistent enough to be merged together, although output by a single user will be more consistent.

Comparison of the resulting land cover product from GE Grids with existing datasets reveals several issues regarding the technical aspects of the different datasets. The differences may in part be an artifact of the varying resolution or age of the datasets, but more likely reflect the challenge of identifying anthropogenic land cover from moderate resolution imagery in savannah Africa. Although Africover is the oldest product, it has the highest percentage agreement with the GE Grids data set. This is likely due to the regional nature of the dataset and its comparatively high-resolution input data (30 m). GlobeLand 30, the only other comparative product with 30 m resolution, has the second greatest overlap with our layer. GlobeLand 30 has nearly the same overall agreement as Africover, yet on a country-by-country basis, its agreement is highly variable versus Africover's. GlobeLand 30 has both the lowest and the highest countrywide agreement (Burundi at 44% and Kenya at 95%) of any comparative product. This inconsistency strengthens the recommendation by Fritz et al. (2011) to review any data set for your area and application before use.

### 3.4.2 Classification challenges

The use of GE Grids to visually classify anthropogenic land conversion does present some new challenges. One issue is image data of moderate resolution or otherwise obscured land cover (commonly due to clouds). However, these regions can first be classified as No

Data and later modified using ancillary data layers where available. WorldPop, a human population density dataset (Linard et al. 2012), was used in this instance as the output was contrasted with land cover data sets. The high level of agreement between WorldPop and the GE Grids product (between 81 and 93% at the country level) validate their use.

Other issues with using GE Grids for identifying anthropogenic land conversion are inherent to Google Earth. These include positional error in data, variability in image date and resolution, and methodological variation among data providers and sensors. The positional accuracy of Google Earth data is debated, but errors are likely sufficiently small to allow for the evaluation of moderate-resolution remote sensing products across the globe (Yu and Gong 2012; Potere 2008). A significant drawback is the temporal variation of Google Earth data. Dates for high-resolution imagery from a random sample of 100 points throughout the study area range from August 10, 2001 to June 27, 2014. This variation makes it impossible to give a definitive reference date for this product. However, roughly 90% of sample points are from the 2010s. Unfortunately, the spatial coverage of various imagery dates cannot be estimated. Naturally, uses at smaller extents would have less temporal variation and represent a more precise period of time. Another challenge is that the imagery displayed in Google Earth is not easily integrated with GIS software, and is updated regularly, thus reducing the replicability of GE Grids' output over long durations.

Further challenges to the use of GE Grids for LULC classifications impact traditional remote sensing analyses as well. The potential misclassification of fallow or retired fields, especially in areas with shifting cultivation, can overestimate anthropogenic impact (Vancutsem et al. 2012). Another potential issue is the use of only one image date in classification (Sedano, Gong, and Ferrao 2005; Watson et al. 2015). The single image may be taken at a time when distinction between croplands and natural vegetation may be difficult; for example when following a fire, during dry seasons, or when lands are left fallow. Additionally, the single image precludes historical analysis (Watson et al. 2015)

However, previous research supports the idea that simple, rapid approaches to land cover mapping have benefits. See et al. (2013a) found that crowdsourced data from Google Earth highlighting the spatial distribution of cropland in Ethiopia had a higher overall accuracy than global land cover products. When analyzing the crowdsourced data itself, See et al. (2013b) found that users underestimate the degree of human impact and there was little difference between experts and non-experts in identifying human impacts (See et al. 2013b). These results suggest that the GE Grids process can produce accurate, conservative estimates of anthropogenic land conversion and can be effectively implemented by non-specialists. In addition, although classification challenges exist, the data set was produced through a manual process and does not use a model to identify impacts that can lead to inaccuracies within protected areas (e.g. Figure 3.3.c).

### 3.4.3 Software evaluation

The software performed well in this case study. Nearly 1,500,000 individual grid cells were evaluated over the course of 600 runs on the basis of image data provided via Google Earth. Despite multiple users, the program gave reliable results and the products easily interfaced with ArcGIS.

The software evolved during the case study as we made modifications to improve reliability and focus on essential program elements. A major addition was the ability to upload a previously evaluated grid cell, enabling us to edit individual runs of the program. Manual edits of raster grids are difficult in ArcGIS and much easier to complete in GE Grids. However, there is room for further improvements to evaluating and editing grid cells. We would like to add the ability to include KMZ files (Google Earth files) as an overlay while running the program. For instance, the user could then bring in protected area boundaries and specifically evaluate land cover on either side of the border. An additional modification should include the ability to access and toggle between different image data providers (like Google and Bing maps). Finally, moving beyond binary classification by allowing additional classes was discussed but was difficult to implement and increased the complexity for the user. This should be a priority in the future.

A substantial challenge to the software occurred during the publication process. Google is deprecating all plug-ins in its browser, Google Chrome, including the Google Earth plug-in used in this program (https://developers.google.com/earth/faq, accessed 10 December, 2016). This required prioritizing functionality on the Mozilla Firefox browser instead of Google Chrome. Indeed, future updates to the program will be required to maintain functionality. However, this method is generic in that the grids could be overlaid on other high-resolution imagery data providers.

## 3.5 Conclusions

Habitat loss via anthropogenic land conversion is a primary driver in biodiversity loss (Pimm et al. 2014). Therefore, identification of human-impacted areas is a critical first step in conservation planning and planning for ecological resilience (Baguette et al. 2013). Yet, existing global land cover products poorly and variably identify croplands and urban areas (Fritz et al. 2011; Fritz, See, and Rembold 2010; Vancutsem et al. 2012). Improvements in the identification of these important areas are necessary. GE Grids can aid conservation purposes by pinpointing anthropogenic land cover and providing complementary information for existing LULC data.

An important difference between traditional LULC mapping and GE Grids is that this tool only produces a binary output as opposed to assigning multiple land cover classes. But when identifying a particular land cover type is very important, such as anthropogenic land cover, GE Grids can be a valuable complement and validation to existing products. Traditional remote sensing techniques require potentially expensive remote sensing data (although this is changing), and specialized knowledge and software (Stensgaard et al. 2009; Pettorelli, Safi, and Turner 2014). Comparatively, using GE Grids is a free, simple, transparent process that can quickly confirm results from more complicated analyses.

Through this process I created a spatial data set of locations where natural land cover had been replaced with various human land covers, representing potential habitat loss for large carnivores. The extent of human land cover compared favorably with Africover, an older but localized and high-resolution data set. The GE Grids land cover data were amended (see next chapter) and used in the subsequent distribution modeling applications.

# Chapter 4

## Bringing ecology back: An ecological approach to selecting pseudo-absences and its impact on species distribution models

A cheetah lounging in the shade of a baobab stump in Ruaha NP. © Andrew Jacobson

# Chapter 4 Bringing ecology back: An ecological approach to selecting pseudo-absences and its impact on species distribution models

## Abstract

Species distribution modeling (SDM) is an increasingly popular tool in ecology and conservation biology with a wide range of uses, such as testing ecological hypotheses, or assessing a species' response to threats. Discriminatory approaches to distribution modeling require contrast data to compare with presence records. Although species occurrence records are readily available from a variety of sources, absence records are more difficult to reliably gather and the majority of distribution modeling studies use pseudo-absence (PsA) data instead. The selection of background extent and PsA data are an active area of research due to their importance in affecting model output, and a number of methods of PsA selection have been proposed. The choice of how to select PsA data is framed here in terms of the larger dynamic in ecology between inference and sampling design. The sampling design must be suitable from a biological and statistical standpoint as it constrains the ecological inferences resulting from the analysis. Importantly, the impact of this choice on the ecological inferences derived from the model is largely unexplored in the SDM literature. I address Research Aim II by demonstrating how model output is affected by the choice of PsA selection strategy, and the implications of this for identifying factors associated with cheetah range decline in East Africa. I do this by selecting PsA using five different methods and comparing their results and ecological inferences. Results show that not all previously proposed methods of selecting PsA data have good ecological justification. In addition, I find that selecting PsA data in two different ways, from the full background extent and from the extirpated range, and comparing the distribution model output can aid in identifying factors associated with cheetah range decline. Results suggest that human land cover, derived in Chapter 3, is a prominent factor in decline.

## 4.1 Introduction

Species distribution modeling (SDM) is an increasingly popular technique in ecology and conservation biology for a variety of uses such as testing ecological hypotheses, assessing species invasion potential, species recovery and reintroduction plans, reserve selection etc. (Guisan and Thuiller 2005; Franklin 2009; Merow et al. 2014; Guillera-Arroita et al. 2015). The statistical sophistication of species distribution modeling is rapidly increasing, with proliferating options for every step of the process. Part of the reason behind the accumulating number of approaches, is that there may be no single best approach to conducting SDM; the method should be selected according to the data, species, and ultimately, research question (Austin 2007).

Discriminatory approaches to distribution modeling require contrast data to compare with presence records. Although species occurrence records are readily available from a variety of sources, absence records are more difficult to reliably gather and pseudo-absence (PsA) data are frequently used instead. The selection of background extent and PsA data are an active area of research due to their importance in affecting model output, and a number of methods of PsA selection have been proposed (Chefaoui and Lobo 2008; VanDerWal et al. 2009; Barbet-Massin et al. 2012; Hertzog, Besnard, and Jay-Robert 2014). Yet, the proposed approaches, and discussion, have largely been from a statistical perspective. A longstanding critique of the SDM field is the need for stronger links between ecological theory and practice (Austin 2002; Elith and Leathwick 2009). Indeed, statistical "best practices" may not always correspond to "best practices" from an ecological perspective in terms of the inferences that can be made. The discussion of this issue has at least partially diverged from the ecology implied by the approaches, and the connection between sampling design and ecological inference has become hazy. This connection should be further explored and explicitly considered in order to draw appropriate inferences from these analyses.

Austin (2002; 2007) frames SDM in terms of three models: ecological, data, and statistical. The ecological model reflects the natural history of the species and represents our expectations about how the species responds to its environment. It also frames the species distribution in terms of how the species responds to environmental gradients, the shape of the response curve to these gradients, and what variables are most important. The data model reflects how and what environmental factors we measure, and at what time scale and grain. Fundamentally, it represents our sampling design. The statistical model provides a fit between the other two models, allows for tests of significance, and frames the interpretation of the results.

Correlative species distribution models can be broken down into two primary types, with some further subdivision. Generative (presence-only) approaches are typically based on detecting environmental conditions at known presences and identifying similar multivariate space. Hence, they identify what combination and range of values for environmental conditions are associated with habitat and search for similar conditions in multivariate space. There are few, true presence-only models such as envelope models, Ecological Niche Factor Analysis (ENFA), and Mahalanobis Distance (Barbet-Massin et al. 2012). The other primary type of SDM algorithm relies on a comparison of presence data with other point data (group discriminatory techniques). The contrast data used in discriminatory models is ideally absence records (Wisz and Guisan 2009; Elith et al. 2011) in order to distinguish habitat from not-habitat. An example of this type is occupancy modeling. However, since absence data can be difficult to obtain (Austin 2002), the contrast points are more commonly a sample of points drawn from the background extent, often called pseudo-absence (PsA) data. Discriminatory models can work with either absence or PsA data, although some like Maxent, are ideally suited for PsA data.

The selection of contrast points is a conceptually important step and is representative of the larger dynamic between inference and sampling design in ecology. The sampling design for a particular experiment (in this case, how to generate PsA points) informs the ecological conclusions that may be drawn from the results (Box 4.1). The ecological question of interest and whether the resulting conclusions are valid will therefore vary with sampling design (Gauch Jr. 2003). The analytical methods are similar for discriminatory approaches, regardless if presence data are contrasted with known absences or PsA, although the ecological inferences will differ. If absence data are available, the comparison is between presences and absences, and hence habitat and non-habitat. But if the distribution of the species is unknown and areas where the species is absent is unknown, then the contrast points can not be drawn specifically from areas where the species was observed or absent. The purpose of these points then is to characterize the available environment, and describe the range of conditions within the study area (Phillips et al. 2009). Thus, the contrast points drawn from the background extent contrast presence with locations that may or may not be habitat, in essence, available habitat.

### Box 4.1 Thought experiment on the dynamic between model construction and inference in species distribution modeling.

Imagine you want to construct a species distribution model for a terrestrial species and you have a few dozen observations to work with. Your landscape is full of forests, croplands, mountains, and lakes but let's say that this species is forest dependent. Let's also assume that the grain of the analysis is static regardless of the study extent. We could start by drawing PsA from a large region around the clustered presence data; that is generating PsA data from all the available habitats including the lakes. After running a distribution model, it's likely the model would achieve high evaluation scores because the presences and PsA are easy to discriminate from each other; they are quite dissimilar in environmental space. However, the results may be uninteresting, as you already know the species is terrestrial and forest-dependent. So then the analysis could be rerun but with the water features masked out. This may lower the evaluation criteria but the results would be more meaningful as it is already known the species is forest-dependent. Then continue masking out other areas in which the species is known to be absent, say croplands. Again, evaluation criteria will likely decrease but the results should become more meaningful and the ecological inferences more insightful. As you go through this process of refining the extent from which PsA are generated, the research question and model output will change. Initially the model may seek predictors that limit the species to the forest, but as known unsuitable habitat is masked out, the distribution model is increasingly attempting to compare species presence with locally available habitat, hence what part of the forest is the species found in. This is analogous to moving up Johnson's (1980) hierarchy of selection processes, from the 1st order (species geographical range) to the 2nd order (home range) or 3rd order (use of habitat components within home range). Naturally, however, the type and quality of the presence data will make some research questions more suitable than others. Thus, the available presence (and absence) data will precondition the research question, the research question will inform the appropriate study extent and PsA selection strategy, and those choices will affect the model results and ecological inferences.

This shift in implication between presence-absence and presence-PsA is important, and an appropriate statistical metaphor is the difference between one and two-sample t-tests. A one-sample t-test contrasts the sample mean to the mean from a random grab of the population whereas the two-sample t-test contrasts the means from two roughly equivalent samples. Using PsA data is equivalent to a one-sample t-test while using absence data is equivalent to a two-sample t-test. In a one-sample t-test, the null hypothesis is that there is no difference between the mean of the sample and of the population, in this case that habitat samples are equivalent to the available habitat. Hence, the null hypothesis is that the species exhibits no preferences for a single type of habitat within the available landscape, whereas the alternative is that the species does select particular habitats. In a two-sample t-test, the null hypothesis is that the two groups (habitat and non-habitat) are roughly equivalent samples from a larger pool of available habitat and the alternative is that the two samples (habitat and non-habitat) are different. Thus, using PsA data instead of absence data changes the interpretation of model results from probability of habitat (i.e. presence) to the preference of particular habitats against other available habitat.

Furthermore, the method of PsA selection alters what habitat is contrasted against, and motivates the ecological inference (Elith et al. 2011). When selecting PsA data in a particular fashion, the inference is that the species selects habitat compared to that set of PsA data. For instance, in order to investigate the realized niche of a species, the study extent must extend beyond the observed environmental tolerances of the species (Austin 2007). In practice this means contrasting presence data with PsA points drawn from outside the historical range of the species. Conceptually the model juxtaposes “where we saw the species” with “where the species could have occurred” and poses the research question: what environmental predictors govern the realized distribution of the species? Yet, researchers may also be interested in other scenarios such as “where we saw the species” vs. “where we used to see them.” This represents a different research question: what are the predictors responsible for an altered distribution? As the research question changes, the PsA generation strategy will change and interpretation of the output will to. Indeed, all aspects of the model output (e.g., response curves, variable importance etc.) should be interpreted in relation to the question.

The strategy for selecting PsA and the selection of the background extent are interrelated and are both critical steps in most SDM approaches (Merow, Smith, and Silander 2013; Phillips et al. 2009; Chefaoui and Lobo 2008; Anderson and Raza 2010). Typically, extent is either chosen for an ecological reason, i.e., some definition of species range, for convenience, i.e., a rectangle (Peterson 2001), to match a political designation (Peterson, Ball, and Cohoon 2002), or some other shape (Barve et al. 2011). Multiple researchers suggest there is a sort of sweet spot; if the background extent is too large then results may be less meaningful although evaluation metrics go up, or too small in which it is more difficult to disentangle habitat from available habitat (Thuiller 2004; Jiménez-Valverde, Lobo, and Hortal 2008; VanDerWal et al. 2009; Acevedo et al. 2012). Northrup et al. (2013) demonstrate a similar challenge with use-available designs in resource selection functions. Ultimately, the selection of the appropriate background extent will vary depending on the research question (Saupe et al. 2012; Merow, Smith, and Silander 2013) and data availability.

Researchers have proposed at least eight different methods of generating PsA (Table 4.1) as well as various more complicated combinations of them. Previous research has unsurprisingly found that different PsA generation strategies affect model performance (Jimenez-Valverde et al. 2008, Lobo et al. 2010, Acevedo et al. 2012), model predictions (Merow, Smith, and Silander 2013; Chefaoui and Lobo 2008), and impact other valuable aspects of SDM, such as variable importance (Stokland et al. 2011) and response curves (Merow, Smith, and Silander 2013; Thuiller et al. 2004; Lobo and Tognelli 2011). Indeed, this is unsurprising as the different selection strategies alters the collection of contrast points and their location in multivariate space relative to presence points (Elith et al. 2011; Senay, Worner, and Ikeda 2013).

I demonstrate how previously proposed methods of generating PsA points alter the ecological interpretation of model results with a case study exploring range contraction of the cheetah in East Africa. To illustrate how different PsA generation methods impact SDM output and interpretation of the output, I use real cheetah presence records and purposefully generate PsA data using five different methods (Figure 4.1). These methods are a subset of available PsA selection methods and were chosen because they have distinct and comparable ecological interpretations. I expand on the concepts behind delimitation of background extent and PsA generation in SDM. I use Random Forests (Breiman 2001), a frequently used and generally reliable approach (Cutler et al. 2007; Watling et al. 2012; Ahmed et al. 2015) to investigate the factors associated with cheetah range decline in East Africa. Cheetah are declining throughout Africa with habitat loss and human persecution believed to be primary threats (Durant et al. 2015).

**Table 4.1 Pseudo-absence selection strategies and their ecological meaning. Bold methods are used in the analysis.** Q – ecological research question posed by the method.

| Method (name) | Source(s) | Ecological meaning |
|---|---|---|
| **Draw PsA randomly from the entire study area (Full)**<br>Q - "Where we saw the species" with "where the species could have occurred" | Stockwell and Peters 1999; Wisz and Guisan 2009 | This approach depends on how the study area is delineated, and the extent of the area relative to the historical range of the species. If the study area is larger than the historical extent of the species, the algorithm should identify general or coarse features that govern the species' distribution, such as forest cover. In addition, the algorithm should produce full species response curves (as opposed to truncated response curves) and may aid in identifying the historical range limits of a species. A biogeographic perspective. |
| **Draw PsA randomly from the historical (and current) extent of the species** (**Historical**)<br>Q - "where we saw the species" vs. "where we might have seen the species" | Anderson and Raza 2010; Saupe et al. 2012; Merow et al. 2013 | This approach contrasts presence data against areas known to previously and currently provide habitat. This may suggest if, and how, current habitat is different from historical range if range has contracted significantly. A "macrohabitat" perspective. |
| **Draw PsA randomly from the resident range (Resident)**<br>Q - "where we saw the species" vs. "other areas they could likely be" | This study | This approach contrasts presence data with areas thought to currently provide habitat for the species. Are there parts of resident range that do not look like "habitat?" Could help identify differences in the area of occupancy vs extent of occurrence? This is a "micro-habitat" perspective that may be more interesting with telemetry or finer-scale movement data. However this approach may be particularly susceptible to error if using coarser observation data or a coarse analysis grain as presence and PsA data are likely to be relatively similar. Analyzing subsets of the presence data may allow for different questions to be modeled such as daily or seasonal differences in habitat use; although different analytical approaches may be better suited. |
| **Draw PsA randomly from** | This study | This approach contrasts presence data with specific areas that once but no longer provide habitat for the |

| | | |
|---|---|---|
| **the extirpated range (Extirpated)**<br>Q - "where we saw the species" vs. "where they used to occur." | | species. What predictors help distinguish between current locations and where the species used to be? This may be of greatest conservation interest. Since this area previously supported the species, if some predictors have changed little since historical times (e.g. climate), then they may be found relatively less important. |
| **Draw PsA randomly from "low suitability" regions (Low suitability)**<br>Q - "where we saw the species" vs. "environment in a different climatic region." | Zaniewski et al. 2002; Engler et al. 2004; Hengl et al. 2009; Barbet-Massin et al. 2012 | This approach uses environmentally-weighted exclusion and draws PsA from regions of "low suitability." The approach takes different forms regarding how the weighting is conducted but, in general, contrasts presence data with areas that are environmentally different from the habitat of the species of interest. Oftentimes, this is called a two-step modeling process because a first model is run to identify environmentally different (e.g. climate) regions. Then PsA are selected from these regions for use in a final model. Therefore, this approach ascertains how current presence is different from regions that are environmentally dissimilar to the species' habitat. Predictors used in the first step may feature prominently in the final model results. |
| Draw PsA randomly from outside a buffer | Hirzel, Helfer, and Metral 2001; Barbet-Massin et al. 2012 | This approach does not necessarily pose an ecological question. The purpose is to contrast presence data with points that are geographically separated and hence less spatially auto-correlated with the presence data. The PsA points may or may not be constrained by the historical extent of the species. |
| Draw PsA randomly from inside a buffer (or other geographic constraint) | VanDerWal et al. 2009 | This approach does not necessarily pose an ecological question although the meaning can vary widely depending on methods and how large or small the buffer is. With a small buffer around the presence data, the approach may be a "microhabitat" perspective. The question then becomes what habitat is selected over other areas that are locally available? This approach could be of more interest with finer resolution movement data and to identify 2nd or 3rd order habitat selection. Very large buffers have different meanings. For instance, buffers could be drawn to estimate the geographic extent that is potentially reachable via dispersal over geological time (Soberon and Peterson 2005; Barve et al. 2011), and the implications in this case are similar to Option A. This may be the preferred approach for modeling historical species' distribution. |

| | | |
|---|---|---|
| Draw PsA from other visited regions | Elith and Leathwick 2007; Phillips et al. 2009 | This approach contrasts presence data with particular locations where researchers looked for but could not find the species of interest. If field sampling was well designed, this is similar to a “true” species distribution model in that the question becomes, what are the differences in predictor variables between locations where the species was and was not found? If, however, the species is present but not seen at many of the ‘absence’ locations, then the question and implications are vague and results unlikely to be meaningful. |
| Draw PsA randomly from a biased background | Merow et al. 2013 | This approach is an attempt to counteract sampling error, by correcting from a model of "where humans go" to a model of where humans go and where the species of interest is seen. Variability in access to all areas of the study site, or variable species visibility across habitats are often causes of sampling error. Therefore, this approach restricts PsA selection to regions that are different from the sampled regions in some preselected way. The variables used to separate these regions may be emphasized in the final model. The ecological implications are vague. |
| Draw PsA randomly with a biased prior | Philips and Dudik 2008; Kramer-Schadt et al. 2013; Merow et al. 2013 | This approach uses a uniform background but attempts to generate PsA data whose bias matches the bias in the presence data (i.e., Gaussian kernel density methods, or bias grids in Maxent). The intent of the biased prior is to simulate sampling error such that both presence and PsA points have similar error structures. Thus, PsA points are drawn in areas that are ecologically similar or auto-correlated to the presence data. This forces the algorithm to attempt to distinguish between habitat where the species was seen and other likely suitable habitat. The ecological implications are vague. |

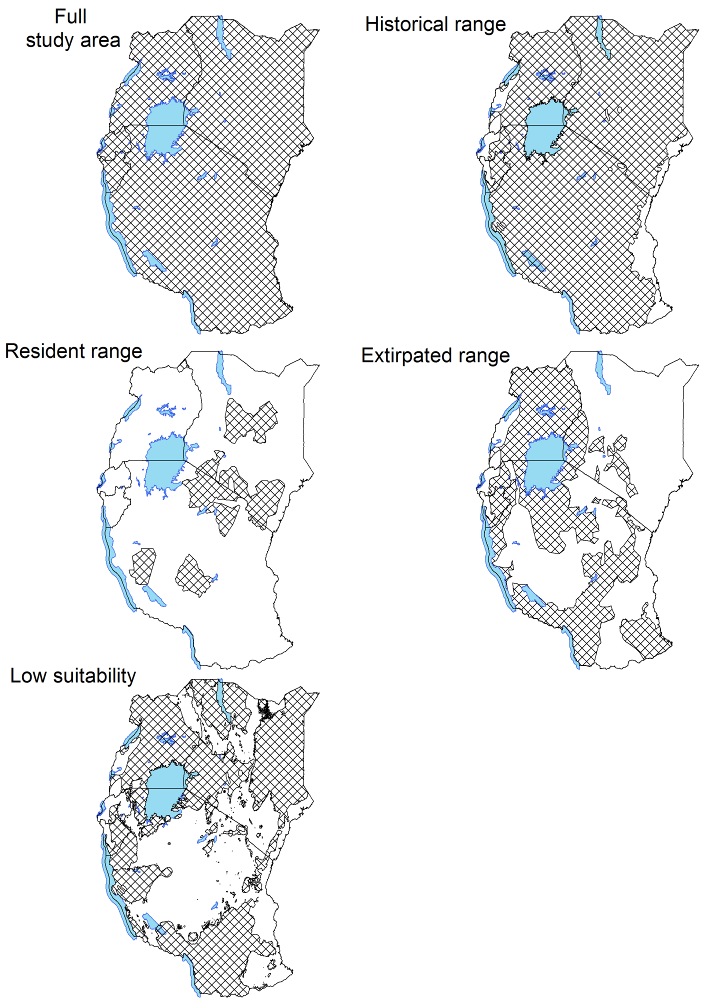


**Figure 4.1 Hash marks indicate areas from which the pseudo-absence data are drawn in the five selection strategies. Blue is water, and the outlines are country borders.**

## 4.2 Methods

### 4.2.1 Species data and study area

The study area is the East African community (~1,700,000 $km^2$ excluding inland water), a five country regional block including, from smallest to largest in area, Rwanda, Burundi, Uganda, Kenya, and Tanzania. I gathered cheetah presence data from 17 data contibutors from March 2013 to February 2015 representing a variety of sources including government authorities, individual researchers, and conservancy game scouts (Table 4.2). No data contributor was from Uganda, Rwanda or Burundi and no data points were gathered from those countries, although only Uganda has any resident range (IUCN/SSC 2007). The majority of data were from sightings and telemetry data. Data were standardized, duplicates and data older than the year 2000 removed, and some points withheld due to reliability concerns. I excluded all records that indicated more than five cheetahs in a single location as cheetahs rarely form groups this size (although litters can be as large as 5 or 6 individuals) (Laurenson 1995; Kelly et al. 1998). I excluded data where the notes and GPS location were mismatched (e.g., the GPS data were not in proximity to a particular geographic feature or political area identified in the notes). I also excluded points due to potential GPS inaccuracy (e.g., presence within a lake), or where the species was likely misidentified (e.g., a cheetah attack on livestock within a boma where there is likely confusion with leopard). Over 10,000 points were initially collected and subsequently reduced to 8,840. This represents the largest, most geographically comprehensive dataset assembled for the cheetah in East Africa.

From 8,840 locations, data were trimmed before model training. I checked for outliers using a scatterplot matrix and removed two data points. I spatially rarefied the occurrence data by removing all points within 10 km, as implemented in the SDM Toolbox (Brown 2014). This reduces the reliance of the training data on telemetry records and geographically focused research projects (e.g. Serengeti Cheetah Project) that are strongly autocorrelated. Those types of records can swamp a dataset and condition the model towards where that research occurred; after rarefying the data, isolated verified records (e.g. from the Tanzania Mammal Atlas Project) become relatively more important. Spatially autocorrelated occurrence points violate one of the standard assumptions of statistical analysis, that of independent and identically distributed data (Dormann et al. 2007). While the remaining presence locations were not necessarily independent, spatially rarefying occurrence data produces better results than some other standard techniques aimed at addressing spatial autocorrelation (Kramer-Schadt et al. 2013; Boria et al. 2014; Fourcade et al. 2014). The choice of this distance is ultimately subjective but seemed appropriate as it is slightly larger than the average daily movement distance of a cheetah; 6 - 8 km per day (Wilson et al. 2013; Scantlebury et al. 2014). In addition, smaller distance thresholds resulted in rapidly increasing number of data points since a number of data sources were highly localized (i.e. telemetry data).

Of the over 10,000 presence points collected, a total of 261 were used in model training. These data are subject to sampling bias because researchers sampled some regions of East Africa more intensively than others. Carnivore research projects, which provided much of the data, are typically established within carnivore range, thus potentially providing little evidence on the status of the carnivore outside known range. They are also typically established in more accessible regions, leaving more remote locations unstudied. Additionally, the training data suffer from imperfect detection as cheetah may inhabit open savannah as well as bushier habitats where they are harder to detect (Bissett and Bernard 2007). Finally, since the data have been collected opportunistically over time and from various sources, prevalence cannot be determined. The majority of SDM studies suffer from the same issues (Guillera-Arroita et al. 2015).

The cheetah was historically distributed from Africa to the Middle East and Asia but today occupy only 13% of their historical range in Africa (Durant et al. 2016), they still occur outside of the study area, i.e. the study area does not encompass their entire range. Within East Africa, the species historically ranged throughout nearly the entire study area except for mangroves and other coastal vegetation, and the Albertine rift mountains (IUCN/SSC 2007). To ensure consistency across vegetation types, I slightly modified the IUCN historical range. Using WWF ecoregions (Olson et al. 2001) and potential natural vegetation (van Breugel et al. 2012), I removed historical cheetah range from all Albertine rift forests, moorlands, mangroves and coastal vegetation. Thus defined cheetah historically ranged across ~1,575,000 km$^2$ in East Africa, or 93% of terrestrial East Africa.

Contemporary distribution of the cheetah was determined via an expert-based mapping approach at regional conservation workshops. The process of the Range Wide Conservation Planning workshop for the cheetah and wild dog, and results are detailed in Box 4.2. Resident range is distributed in seven patches, with a small patch in northern Uganda (Kidepo Valley NP), and otherwise split roughly equally in terms of area between Kenya and Tanzania (IUCN/SSC 2007). The species is extirpated from Rwanda and Burundi (IUCN/SSC 2007). Extirpated range is 680,000 km$^2$ (43% of historical range) whereas resident range is only 245,000 km$^2$ (16%).

**Box 4.2 Description of the mapping process used to assess the cheetah**

A mapping process established for the jaguar (Sanderson et al. 2002) was used during IUCN/Species Survival Commission conservation strategic planning workshops for the cheetah and wild dog (Figure 4.2). This process was replicated at three regional workshops in Africa. The Eastern Africa workshop occurred in 2007 (IUCN/SSC 2007). The distribution maps were drawn via expert opinion with the aid location records, and land cover and vegetation data. The maps were edited in subsequent national workshops.

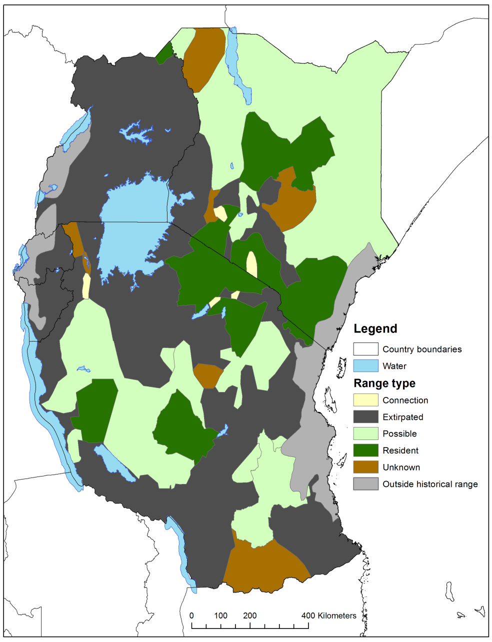


**Figure 4.2. Expert-derived cheetah range in East Africa**

Range categories include resident range, possible range, connecting range, unknown range, and extirpated range (itself divided into recoverable and unrecoverable). Resident range was defined as land where the species was known to be still resident based on (*i*) regular detection of the species in an area over a period of several years and/or (*ii*) evidence of breeding. Extirpated range was defined as where the species is extinct and the land is uninhabitable for the foreseeable future. The other range categories are not used here but their definitions can be found in (IUCN/SSC 2007).

**Table 4.2 Sources of cheetah occurrence records for East Africa.**

| Source | Region | Number of records | Data Collection Method(s) | Data Format |
|---|---|---|---|---|
| Alexandra Sutton | Vicinity of Masai Mara NR, Kenya | 3 | Sightings | GPS coordinates |
| Amy Dickman | Vicinity of Ruaha NP, Tanzania | 292 | Sightings | GPS coordinates |
| Asgar Pathan | Tsavo NP, Kenya | 5 | Sightings | GPS coordinates |
| Cherie Schroff | Tsavo NP, Kenya | 47 | Sightings | GPS coordinates |
| Elena Chelysheva | Masai Mara NR, Kenya | 138 | Sightings | KMZ file |
| Femke Broekhuis | Masai Mara NR, Kenya | 76 | Sightings | GPS coordinates |
| Ingela Jansson | Ngorongoro Conservation Area, Tanzania | 5 | Conflict records | GPS coordinates |
| Kenya Wildlife Authority | Kenya | 84 | NA | GPS coordinates |
| Laly Lichtenfeld | Vicinity of Tarangire NP, Tanzania | 22 | Sightings, spoor, camera traps, conflict records | GPS coordinates |

| | | | | |
|---|---|---|---|---|
| Mary Wykstra | Kenya | 2,739 | Telemetry, sightings | GPS coordinates |
| Michael Mbithi | Athi-Kapiti, Kenya | 81 | Sightings | KMZ file |
| Northern Rangelands Trust | Kenya | 2,267 | Sightings | GPS coordinates |
| Paul Schuette | southern Kenya | 10 | Sightings, camera traps | GPS coordinates |
| Phillip Henschel | Tsavo NP, Kenya | 56 | Transects | Presence grid cells |
| Sarah Durant | Serengeti ecosystem, Tanzania | 3,489 | Sightings | GPS coordinates |
| Tanzania Mammal Atlas Project | Tanzania | 818 | Sightings, camera traps | GPS coordinates |
| Various individuals | Tanzania, Kenya | 3 | Sightings | GPS coordinates |
| **Totals** | | **10,135** | | |

### 4.2.2 Predictor variables

Important predictor variables expected to influence cheetah distribution were selected based on ecological knowledge, a review of literature, and available data sets (Table 4.3). Predictors comprised nine general categories, normalized difference vegetation index (NDVI), soil nutrients, rainfall, temperature, elevation, water, human population density, human land cover, and human pressure (Supplemental Figures 4.1-3). All datasets were converted to raster format, and projected into the WGS 1984 Africa Albers Equal Area Conic coordinate system. Whenever possible, I gathered data from a 20 km buffer around the study area to account for the influence of environmental factors along the border of East Africa. All input datasets were processed using ArcGIS 10.2 (ESRI 2014) at 500 m resolution using a snap raster to align all grids. Besides 500 m, I set three additional scales. Wildlife may use the landscape at different scales and it is prudent to allow the model to select the most appropriate scale (Guisan and Thuiller 2005; Mashintonio et al. 2014). Average cheetah daily distance moved is around 6-8 km per day (Wilson et al. 2013; Scantlebury et al. 2014) therefore I selected scales of 3, 6, and 12 km. I used the Focal Statistics tool in ArcGIS to average values across squares with these diameters.

#### 4.2.2.1 NDVI

Vegetation quantity and structure affects large carnivore behavior and movement (East 1984; Loveridge & Canney 2009; Rabinowitz & Zeller 2010; Hayward et al. 2007). NDVI, a measure of vegetation abundance derived from remote sensing data, corresponds with a wide range of ecological processes, and is often a proxy for vegetation greenness and forage quality (Pettorelli et al., 2005). In Africa, large mammalian herbivore abundance correlates with above ground primary production and mean annual precipitation (Coe, Cumming, and Phillipson 1976). Prey abundance is a primary factor in determining carnivore carrying capacity (Hayward, Obrien, and Kerley 2007). In addition, cheetah select for areas with high prey densities (Vanak et al. 2013; Broekhuis et al. 2013; Swanson et al. 2014). However, prey abundance varies temporally, and is affected by human activities; hence prey abundance is poorly known across Africa. Therefore, NDVI is frequently used as a proxy for herbivore distribution and biomass, and is correlated with carnivore distributions (Pettorelli et al. 2009; Swanepoel et al. 2013). I did not assess the relationship of this, or other, proxies to actual prey biomass.

To construct the NDVI, I used NASA's MODIS MOD13Q1 16-day composite downloaded from the Land Processes Distributed Active Archive Center (LP DAAC). I acquired dry season images, sampled 12 August– 27 August, and rainy season, sampled 22 April – 7 May. For each season, I mosaicked seven scenes that covered the entire study area and repeated this for five years from 2008-2012. The MOD13Q1 dataset includes a pixel reliability layer. I masked out pixels classified as No Data, Snow/Ice, and Cloudy. After removing unsuitable pixels, I averaged NDVI values from 2008-2012 for each season producing two composite datasets.

I also created a standard deviation of NDVI values across different land cover types. Since species may be more or less related to particular habitats, and these habitats may have particular NDVI ranges, then a species may be naturally associated with a particular range of NDVI. Therefore, it may be more valuable to see if cheetahs utilize a specific range of NDVI values within their habitats. The study area was split into different regions based on their primary land cover type using GlobLand 30 (National Geomatics Center of China 2014). This data set had the second highest agreement, at 87% with results from GE Grids in East Africa (Jacobson et al. 2015). The only data set with better agreement did not have suitable land cover categories for this purpose and was discarded. The mean and standard deviation of NDVI values per land cover type (e.g. forest, cropland) was calculated. The individual standard deviation NDVI layer for each land cover class was then combined into a single data set to cover the entire study area.

#### 4.2.2.2 Climate

Climatic factors correlate with carnivore distribution via their effect on vegetation and hence herbivore biomass (Celesia et al. 2009; Loveridge and Canney 2009). For instance, Celesia et al. (2009) found that temperature and precipitation together explained 62% of the variation in the demography of a large carnivore, the lion *Panthera leo*, across its African range. Four climatic variables were downloaded from the WorldClim database (Hijmans et al. 2005), annual mean temperature, total annual precipitation, dry season precipitation, and precipitation seasonality. These are global interpolated values at a resolution of 30 arc-seconds, or roughly 1 $km^2$. Dry season precipitation is calculated as the amount of precipitation in the driest quarter of the year. Precipitation seasonality is the coefficient of variation in precipitation, and is a measure of the amount of variation in rainfall between the wettest and driest month. A higher value represents a larger difference in precipitation between the wettest and driest month. These values are interpolated and hence quality varies spatially and may degrade particularly in areas with sparse data or more locally variable climate (Hijmans et al. 2005).

**Table 4.3 Predictor variables used in the modeling approach. The grain of the analysis is 500 m. All variables except soil taxonomy and human impact are also calculated at three additional scales, smoothed over 3, 6 and 12 km roving windows.**

| General variable | Specific variable (*acronym used in analysis*) | Source | Mean; std dev | Range | Original resolution |
|---|---|---|---|---|---|
| NDVI | Wet season, April, NDVI averaged over 5 years ($NDVI_{wet}$) | MODIS NDVI | 5849; 1864 | -2000, 9857 | 250 m |
| | Dry season, August, NDVI averaged over 5 years ($NDVI_{dry}$) | MODIS NDVI | 3966; 1800 | -2000, 9802 | 250 m |
| | NDVI value in standard deviations away from the mean, assessed per land cover type ($NDVI\ Stdv_{dry}$) | MODIS NDVI | -0.04; 0.99 0; 1 | -7.45, 6.29<br>-3.39, 8.77 | 250 m |
| Soil nutrients | Organic carbon content (tons/ha) (*Soil Carbon*) | ISRIC-WISE | 136; 71 | 0, 1594 | 1 km |
| | Cation exchange capacity (cmolc/kg) (*Soil CEC*) | ISRIC-WISE | 19; 8 | 4, 96 | 1 km |
| | Soil taxonomy (USDA classification) | ISRIC-WISE | | 12 classes | 1 km |
| Rainfall | Annual precipitation (mm) (*Precip*) | WorldClim; Bio 12 | 872; 343 | 173, 2599 | 30 arc second |
| | Rainfall seasonality (higher values indicate greater seasonality) | WorldClim; Bio 15 | 81; 24 | 26, 141 | 30 arc second |
| | Dry season rainfall (mm) ($Precip_{dry}$) | WorldClim; Bio 17 | 40; 53 | 0, 389 | 30 arc second |
| Temperature | Mean annual temperature (°C) (*MAT*) | WorldClim; Bio 1 | 22.9; 3 | -4.0, 29.6 | 30 arc second |
| Elevation | Elevation (m) (*Elev*) | SRTM | 981; 546 | -1, 5842 | 250 m |
| | Slope (%) | SRTM | 1.8; 2.8 | 0, 44.5 | 250 m |
| | Terrain Ruggedness Index (*TRI*) | SRTM | 204; 175 | 0, 2527 | 250 m |
| Water | Distance to river and lake (m) | WWF hydrosheds & Africover (water) | 13722; 10257 | 0, 56732 | NA |
| Human population density | Human population density in 2015 (people/$km^2$) (*HPD*) | WorldPop | 71.6; 57.9 | 0, 2217 | 100 m |
| Human land cover | Percent human land cover (*% Human LC*) | GE Grids | 0.30; 0.40 | 0, 1 | 0.01 degree |
| Human pressure | Human pressure derived from human population density and roads using an inverse distance-weighted cost surface (*Human impact*) | WorldPop; DCW & OSM roads | 70.9; 13.3 | 0, 100 | WorldPop (100 m); roads (NA) |

#### 4.2.2.3 Landscape features

I collected elevation data from the USGS Shuttle Radar Topography Mission (Jarvis et al. 2008) at 250 m resolution. I calculated slope from the elevation data after resampling to 500 m resolution. Terrain ruggedness index (TRI) (Riley, DeGloria, and Elliott 1999) is a measure of the variation in topography of an area. Pettorelli et al. (2009) found cheetah avoid hills.

I also calculated distance to water, incorporating both rivers and lakes. Pettorelli et al. (2009) found a positive association with lakes and Hilborn et al. (2012) found that cheetahs commonly hunted near rivers. For data on river distribution, I used the WWF Hydroshed dataset (Lehner, Verdin, and Jarvis 2006). For data on lakes and other water bodies, I combined all the "water" classes from the regional land cover dataset Africover (Alinovi, Di Gregorio, and Latham 2000).

Soil is the medium for plant growth and regulator of water supplies (Hengl and Reuter 2009), hence indirectly affecting predator and prey distribution. Soil data sets are available from the ISRIC - World Data Center for Soils, downloaded in April 2014. These datasets are included in the SoilGrids1km (Hengl et al. 2014), a global 3D soil information system. These represent spatial predictions for a selection of soil properties at six standard depths. I used three layers in this analysis, USDA Soil Taxonomy orders, cation exchange capacity (cmol/kg), and soil organic carbon stock (tons/ha). Hengl et al (2014) caution that these data represent a first attempt at providing soil data information at a global scale and have low accuracy.

USDA Soil taxonomy is a classification of soil types that clusters soils according to various parameters. Soil formation is influenced by climate, relief, biological organisms, and parent material interacting over time (Brady and Weil 1996). Soils and vegetation are mutually associated and, at the larger extent, soil can correlate with habitat types (Jensen, Simonson, and Dosskey 1990; Brady and Weil 1996). I condensed soil types by order, the most basic classification, leaving 12 orders. Cation exchange capacity (CEC) is an important soil property influencing nutrient availability, soil pH, and soil structure stability (Hazelton and Murphy 2007). It is an inherent soil characteristic that influences the soil's ability to retain essential nutrients that are thereby accessible for plant growth. Soils with higher CEC have greater water holding capacity and organic matter has very high CEC as well. CEC, in combination with rainfall, was used as a proxy for ungulate biomass when predicting African lion distribution (Loveridge and Canney 2009). Soil organic carbon is also an important component of soil health. Like CEC, soil carbon increases water-holding capacity, contributes to structural stability, prevents nutrient leaching, is integral to making minerals available to plants, and linked with plant biomass (Brady and Weil 1996; Hazelton and Murphy 2007). I used the mean estimate of CEC and soil carbon for the topmost soil horizon.

#### 4.2.2.4 Human impacts

Previous studies suggest variables related to human densities and land cover are important in determining carnivore distribution (Basille et al. 2009; De Angelo, Paviolo, and Di Bitetti 2011; Swanepoel et al. 2013) and that cheetah avoid settlements (Pettorelli et al. 2009). I use three different human impact data sets, human population density, human land cover, and a derived metric of human pressure on natural resources.

I downloaded 2015 human population density data at 100 m resolution from the WorldPop project (Linard et al. 2012). WorldPop gridded population density estimates are matched with available census data and modeled using a Random Forest approach combined with various data sets such as land cover, roads, topography and nights at light (Stevens et al. 2015). I downloaded all five countries plus the surrounding seven countries in order to account for cross-boundary effects. I edited the data by masking out any human populations within national park or national reserve boundaries (downloaded from the World Database on Protected Areas; UNEP & IUCN 2014). Settlements are not allowed within these reserve types and projected densities here likely represent modeled impacts rather than permanent populations.

Human land cover is taken from (Jacobson et al. 2015) and detailed in Chapter 3. Land converted from natural habitats into primarily settlements or cropland, was distinguished manually using high-resolution imagery available on Google Earth (© Google Inc.). The data set was slightly amended before use here. As the purpose in Chapter 3 was to contrast the GE Grids process with existing land cover data sets, I filled no data holes using human population density data. For this (and subsequent) purposes, I filled no data holes with agriculture and settlement land cover types from the Africover data set (Alinovi, Di Gregorio, and Latham 2000). Africover had the highest and most consistent agreement with the GE Grids product and had greater agreement than the human population density product (see Chapter 3). The data was downscaled to 500 m and averaged across the different scales to become a percentage.

Other land cover data was not used in this analysis due to their reliance on NDVI in land cover classification (Townshend et al. 1991). Thus, using land cover data in addition to NDVI would be duplicative.

I created a cost surface of human pressure on natural resources. This variable is derived from human population density and road data. Urban centers can be hubs for charcoal, illegal logging and the wildlife trade; essentially acting as centers of demand for natural resources (Arnold, Köhlin, and Persson 2006; Brashares et al. 2011; Van Vliet, Nasi, and Taber 2011). These commodities are moved into cities along the transportation network (Arnold, Köhlin, and Persson 2006; Brashares et al. 2011; Van Vliet, Nasi, and Taber 2011). Thus, I created an inverse distance-weighting scheme from urban centers and roadways. Urban extent boundaries were set at human population density thresholds of 1,000 and 5,000 people per km$^2$. These thresholds were set to represent two different size categories of urban areas, with more populous cities having a greater pull of natural resources. In many places in Africa, natural resources like bushmeat and charcoal are normal goods and consumption remains high in urban areas, thus larger urbanized areas represent increased demand (Arnold, Köhlin, and Persson 2006; Brashares et al. 2011; Van Vliet, Nasi, and Taber 2011). The thresholds were selected by trial and error to identify only larger towns and cities. Then I calculated the cost distance away from these two thresholds of urban centers. I also calculated cost distance from two different road types, primary/trunk roads, and all other roads. Accurate road data sets in Africa are difficult to source, so I used two different data sets. I used the 1992 Digital Chart of the World road layer for minor roads (Danko 1992), which although it is old, is comprehensive. I used Open Street Map data for trunk and primary roads (© OpenStreetMap contributors). I selected these data based on visual comparisons with other data sets and with Google Earth. I calculated cost distance in ArcGIS using the Reclass by Function tool, scaled from 0 (low) to 100 (high). Cost distance for both roads and urban centers was set so that cost decayed at faster rates for smaller urban centers and along minor roads, than larger urban centers and primary roads, mimicking the constrained demand in natural resource for smaller urban centers and minor roads. Thus, the largest values represent urban centers and the smallest values represent those areas furthest from roads and urban centers.

### 4.2.3 Species distribution modeling

To demonstrate the effect of changing pseudo-absence points in a SDM framework, I chose a popular and effective algorithm, Random Forests (RF) (Cutler et al. 2007; Watling et al. 2012). RF is recommended for use at the regional scale, and is effective in consistently identifying important variables (Breiman 2001; Aguirre-Gutiérrez et al. 2013). I implemented Breiman and Cutler's random forest for classification and regression in the R statistical software (R Core Team 2015) using the randomForest package (Liaw and Wiener 2002).

#### 4.2.3.1 Random Forests

Random Forests (RF) were born out of the machine-learning field and ecologists started using them in the mid-2000's (Cutler et al. 2007). RF are extensions of classification and regression trees, or CART, models (Breiman et al. 1984). In a standard classification situation, classification trees recursively partition data into increasingly homogenous classes. At each step, or node, in fitting a classification tree, the algorithm selects an optimal predictor variable and cut-off value that results in the most homogenous subgroup as measured by an impurity metric, such as the Gini index (Breiman et al. 1984). Splitting continues until further subdivision no longer reduces the Gini index. At this point, the terminal nodes are as pure as possible and the "tree" is fully grown. The tree can be "pruned" by removing some of the lower branches, or nodes. This results in new terminal nodes that are more heterogeneous and thereby less fit to the training data.

CART models have several benefits and drawbacks. On the plus side, CARTs are non-parametric, are not subject to distributional assumptions, do not require variable transformation, can use categorical, ordinal and continuous data simultaneously, can incorporate complex variable interaction, and are capable of handling high-dimensional data (Evans et al. 2011). On the other hand, CARTs can succumb to over-fitting, the final tree may not be the optimal solution, and results can exhibit high variance (Evans et al. 2011). These drawbacks spurred the creation of RF such that the algorithm can build thousands of trees and average the results to mitigate challenges associated with single CART models.

RF are structured as a collection of individual CART trees with predictions combined across all the trees. The algorithm selects bootstrap samples from the data. Approximately 63% of the data points are selected in the bootstrap; those that do not occur in the bootstrap are called out-of-bag (OOB) observations. For each bootstrap, the algorithm fits a tree and, importantly, only a few randomly selected variables are available for use at each node in tree building. This is repeated a large number of times. Cutler et al. (2007) recommend 500 and I used 5,000. The number of available variables is termed the mtry, and is typically set at the square root of the number of variables (Liaw and Wiener 2002; Strobl et al. 2008; Evans et al. 2011) although Cutler et al. (2007) found that the algorithm is insensitive to the chosen value. With each tree fully grown, the OOB samples are used to assess accuracy, and OOB error rates are averaged across all trees. Thus, OOB error rates are essentially cross-validated accuracy estimates (Cutler et al. 2007). This is a strength of RF and hence independent observations to assess error are not required although still recommended (Evans et al. 2011).

Variable importance in RF is derived rather uniquely (Cutler et al. 2007). As stated earlier, each tree has a misclassification rate based on OOB data. RF randomly permutes the values of the variable for the OOB observations and fits a new tree. Variable importance is thus the difference in misclassification rates for the modified and original OOB data divided by the standard error (Cutler et al. 2007). As implemented in the randomForest package, RF outputs both the Gini index and a mean decrease in accuracy value. Strobl et al. (2008) recommended using the mean decrease in accuracy value.

#### 4.2.3.2 Pseudo-absence point creation

I selected PsA using five different strategies (Figure 4.3). Three options were previously recommended and have been used in multiple studies, while two were developed in this study (see Table 4.1). The two new strategies were devised because they have potentially useful ecological implications and use prior knowledge on cheetah distribution. Most importantly, these five methods were selected because they have clear, defined and distinct ecological implications. While some strategies may be more appropriate with other research questions or address statistical issues with presence data, the ecological inferences from these approaches may be less identifiable or easily contrasted. Hence, the five selected options present a range of describable ecological implications that can be readily compared.

Each strategy used a different selection process in generating PsA data but all excluded water. In Method A, “Full,” I selected PsA points randomly from the entire study extent (Stockwell and Peters 1999; Wisz and Guisan 2009). In Method B, “Historical,” I selected

PsA points randomly from the historical range of the species (Anderson and Raza 2010; Saupe et al. 2012; Merow, Smith, and Silander 2013). In Method C and D, "Resident" and "Extirpated," I selected PsA points randomly from within regions defined as resident or extirpated range respectively. Note that the PsA drawn from extirpated range should not be considered real absence data, as there may be unused habitat within the range. In Method E, "Low Suitability," I selected PsA points from a region that is climatically different, as proposed by Barbet-Massin et al. (2012) and similar to approaches advocated by (Zaniewski, Lehmann, and Overton 2002; Engler, Guisan, and Rechsteiner 2004). To do this I created a surface range envelope, equivalent to running BIOCLIM, using the BIOMOD2 package in R (Thuiller, Georges, and Engler 2014). I used four climatic variables and via the surface range envelope algorithm, identified regions that were climatically different from the presence data.

In order to examine the factors associated with cheetah range decline, the Full and Extirpated options are most appropriate. The Full method contrasts current presence records with PsA drawn from outside the historical extent of cheetah thus identifying what predictors are important in delimiting current range. The Extirpated method contrasts presence records with PsA drawn from extirpated range and thus investigates what predictors best identify differences between current and extirpated range. Thus, contrasting these two options enables a comparison of the factors that are responsible for current habitat, and emphasizing what factors are associated with former habitat.

The appropriate number of PsA points is also a concern as it affects prevalence, the ratio of presence to contrast points (McPherson, Jetz, and Rogers 2004; Phillips and Dudik 2008; Barbet-Massin et al. 2012). It is not extensively studied, but prevalence can impact model accuracy (McPherson, Jetz, and Rogers 2004; Barbet-Massin et al. 2012). Barbet-Massin et al. (2012) used virtual species to determine the importance of varying the number and selection strategy of PsA point creation across different modeling algorithms. Specifically for classification techniques like RF, the number of PsA points had a greater effect on model accuracy than the selection strategy (Barbet-Massin et al. 2012) (opposite of Stokland et al. 2011 who used boosted regression trees). They advocated that with the RF algorithm, prevalence should be one, hence an equal number of presence and PsA points. Barbet-Massin et al. (2012) also recommended, particularly in instances where the study area is large, to draw a large number of PsA and then randomly subsample and average the SDM results across multiple pulls. I selected 5,000 PsA points for each methodology and subsampled 261 for each of 10 model runs. PsA points were created in R using the raster (Hijmans 2016) and dismo packages (Hijmans et al. 2016).

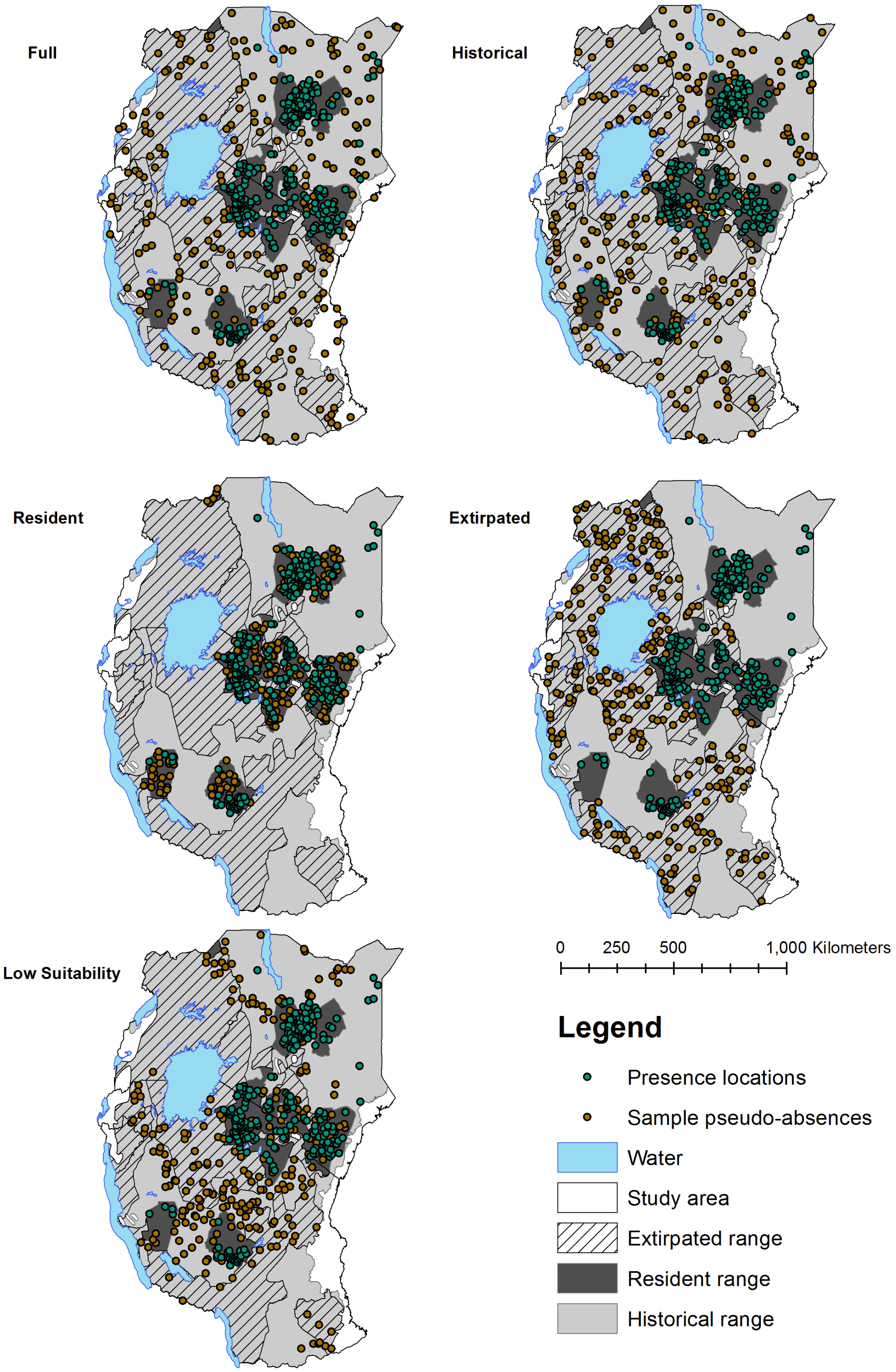


**Figure 4.3** **Location of presence and PsA data. Only 500 of the 5,000 PsA are shown for illustrative purposes.**

#### 4.2.3.3 Variable selection

I started with 63 variables and selected the most important uncorrelated predictors. Following the recommendations of Strobl et al. (2008) and similar to (Forester, Dechaine, and Bunn 2013), I used a large number of trees (5,000) and averaged variable importance across several mtry values (i.e., the number of predictors that can be used at each split in the tree). I ran 25 different RFs at 5,000 trees each at four different mtry values to create 100 estimates of variable importance. Mtry values were set at roughly half the default value, the default value, and then roughly two times and five times the default value (similar to recommendations in Liaw and Wiener 2002). I selected eight predictors for use in the final models. There was no clear guidance on the appropriate number of predictors (Franklin 2009), this number was similar to other studies, and this accounted for most uncorrelated variables whose importance was within roughly a factor of 10 of the most important variable.

I reviewed the Pearson correlation coefficients of the predictors and retained only a single predictor if more than one was correlated above 0.7 (Dormann et al. 2013). I did not discard highly correlated variables outright; rather, I selected only the most important of the correlated variables during the variable selection process. I also reviewed variable collinearity using the rfUtilities package (Evans and Murphy 2015), although none were collinear.

Finally, I reviewed the similarity between the range of predictor values within and outside protected area boundaries. Protected areas may not contain a full range of predictor values and presence data may be biased towards protected areas. Therefore, I visually compared boxplots of the predictors used in the final models inside and outside protected areas. The protected area network of East Africa is quite extensive (14% of terrestrial area) and covers varied habitats, (Jason Riggio et al. unpublished data). No predictor variable had substantially different ranges of values between these two regions except for the three human impact variables. In addition, 48% of presence points (after spatially rarefying) were within protected areas and a similar percentage, 55%, of resident range is also within protected areas. This suggests that presence points and resident range are roughly similarly represented in protected areas.

#### 4.2.3.4 Model implementation

After variable selection, I implemented five final RFs, one per PsA selection strategy. I ran each model ten times, subsampling the PsA points. I chose 10 runs as a compromise between processing time and obtaining an understanding of the variation resulting from different PsA pulls. I averaged the ten runs to create an ensemble for each PsA selection strategy. I reviewed model consistency by checking the Pearson correlation coefficient between all ten runs.

Using the rfUtilities package (Evans and Murphy 2015), I compared variable response curves within each PsA selection strategy (i.e. between the 10 runs subsamples) and between the five different PsA selection strategies.

The exact meaning of the output from SDM is typically vague or rarely explicitly considered (Aarts, Fieberg, and Matthiopoulos 2012; Phillips and Elith 2013; Yackulic et al. 2013). Generally speaking, it is commonly interpreted as a probability of occurrence somewhere between the potential and realized distribution of a species (Jiménez-Valverde, Lobo, and Hortal 2008). Recently, Guillera-Arroita et al. (2015) argued that both the data and statistical models impact interpretation of results. They argue that SDM output can, from least to most informative, vary in meaning from relative likelihood of observation, relative ranking of probability of occurrence at various sites, relative likelihood of occurrence, and actual probability of occurrence. As the cheetah presence data, suffers from sampling bias, imperfect detection and unknown prevalence, similar to SDM of most large mammals, the output reveals only the relative likelihood of observation. Henceforth, I will call the output “likelihood of observation.”

#### 4.2.3.5 Model evaluation

Proper evaluation of a distribution model consists of multiple measures, and while focus is often on a model’s predictive performance, that is only one part of model evaluation (Austin 2007; Franklin 2009). Reliance solely on predictive accuracy is unwise because a model may have good metrics but relate poorly to the underlying biology (Warren and Seifert 2011). A model should be reviewed to determine if it is ecologically realistic, robust, and has ‘good’ predictive ability within clearly defined criteria. Hence, to determine ecological realism, I reviewed the response curves and compared the spatial output to the cheetah’s current distribution. I evaluated model robustness by reviewing the consistency within and between model runs. Finally, in terms of predictive performance, I selected several threshold independent evaluation metrics.

To verify model predictions, I chose three threshold independent evaluation measures: OOB error rates, the Boyce Index (Boyce et al. 2002) calculated in the ecospat package (Broennimann et al. 2015), and AUC. I selected threshold independent measures to relieve from selecting a threshold as it represents a loss of information (Hirzel et al. 2006; Freeman and Moisen 2008; Guillera-Arroita et al. 2015). I also calculated the predicted average probability of observation for each ensemble over the different expert-derived distribution zones: resident, extirpated, and outside the historical range. A higher probability in resident range, and probabilities close to zero in the extirpated and outside the historical range are expected.

The AUC is derived from the Receiver Operating Characteristic (ROC) plot (Pearce and Ferrier 2000). The ROC plot is a graph of Sensitivity (true positive rate) on the y-axis and 1-Specificity (or the false-positive rate) on the x-axis. The Sensitivity and Specificity values are plotted by evaluating the model prediction at a large number of thresholds and together compose the ROC curve. Each threshold creates a binary response in the model prediction and when compared to the evaluation data gives a particular sensitivity and specificity value. Therefore the ROC curve is built up across a number of thresholds, identifying their associated sensitivity and specificity values, and plotting these points on the graph. This means that the AUC is evaluated across all possible thresholds rather than a particular set threshold chosen by the user (such as threshold-dependent evaluation statistics like Cohen’s Kappa). After creation of the ROC curve, the AUC is calculated by summing the area under the ROC curve (AUC). A value of 0.5 is considered random and values closer to 1.0 indicate performance better than random. A number of serious concerns have arisen about AUC (Lobo, Jiménez-Valverde, and Real 2008). An important concern is that as the area of extent over which the prediction is applied increases, AUC scores may also increase (Lobo, Jiménez-Valverde, and Real 2008). When used in presence-only situations like this, the AUC should also be interpreted differently than if real absence points are used because misclassification of PsA should be expected, roughly at a proportion that corresponds with the amount of habitat in the study area (Franklin 2009). AUC from presence/background data describes the probability that the model scores a random presence site higher than a random background site (Phillips et al. 2009). In addition, AUC cannot be compared across species or study regions since differing amounts of the study area are inherently suitable (Anderson and Gonzalez Jr. 2011) (but also see Lobo et al. 2008 and Peterson et al. 2008).

The Boyce Index (Boyce et al. 2002), recommended by Hirzel et al. (2006), is a threshold independent evaluation measure. It is calculated via a Spearman's rank correlation coefficient between the frequency of evaluation points and bin ranks of predicted suitability. The frequency of presence locations within ranked bins of predicted suitability is expected to increase monotonically with bin number (Franklin 2009). The coefficient varies from -1 to 1 with 1 representing a model in which higher predicted probabilities are consistent with presence locations, values close to zero indicating a model similar to random, and negative values in which low probabilities are predicted in areas with more presences.

The three measures of predictive performance, the OOB, AUC and Boyce Index are a good combination of threshold independent measures. The AUC is the most commonly reported evaluation score and balances both omission and commission errors (Franklin 2009; Hijmans 2012). However, the OOB is a resampling approach that increases the robustness of evaluation measures whereas the AUC and Boyce Index are just a single measure. However, both the AUC and OOB incorporate absence (in this case PsA) data and since some PsA are expected to be in suitable habitat, this is expected to alter the interpretation and lower the values (Hijmans 2012). Boyce does not incorporate PsA data, only comparing the density of presences to the binned suitability classes (Hirzel et al. 2006). A few presences may occur in areas of lower suitability but more presences should occur in areas of higher suitability and the Boyce Index recognizes this relationship.

## 4.3 Results

I ran five RF models of cheetah occurrence in East Africa to demonstrate how the ecological inferences of SDM results vary by PsA strategy (Figure 4.4). The five PsA selection strategies posed different modeling questions. The Full model contrasted "where we saw the species" with "where the species could have occurred". The Historical option contrasted "where we saw the species" with "where we might have seen the species". In the third and fourth options, the modeling questions here posed "where we saw the species" vs. "other areas they could likely be" (Resident), and "where the species used to occur" (Extirpated). The fifth option, Low Suitability, posed "where we saw the species" vs. "environment in a different climatic region."

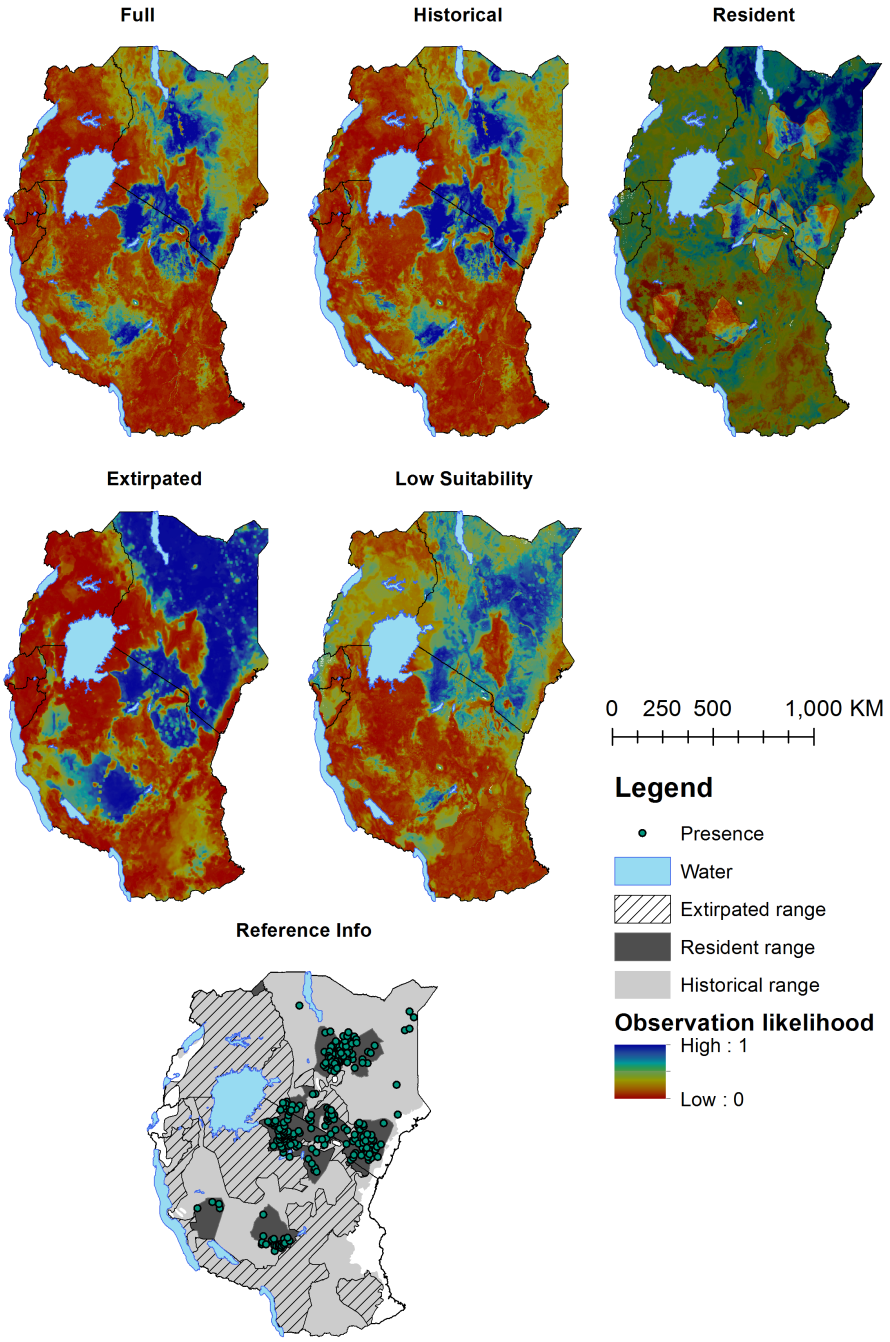


**Figure 4.4 Likelihood of observation for cheetah based on different PsA selection methods. The resident range is highlighted in the Resident option as the predictions are more relevant within that area.**

The eight most important uncorrelated variables were selected for the final model of each option (Table 4.4). Variables averaged at the largest scale (12 km) were more frequently but not universally selected over predictors averaged at smaller scales. Human population density was selected in every method, although it was less important than percent human land cover.

**Table 4.4 Variables ranked by importance for each of the five modeling options. Variable names/acronyms listed in Table 4.3.**

| **Full** | | **Historical** | | **Resident** | |
|---|---|---|---|---|---|
| Variable (scale) | Rank | Variable (scale) | Rank | Variable (scale) | Rank |
| MAT (6km) | 1.00 | MAT (12km) | 1.00 | NDVI $_{wet}$ | 1.00 |
| NDVI $_{dry}$ (3km) | 0.99 | % Human LC (12km) | 0.95 | Soil Carbon (12km) | 0.89 |
| % Human LC (12km) | 0.98 | NDVI $_{dry}$ (3km) | 0.81 | Bio 17 (12km) | 0.86 |
| NDVI $_{wet}$ (3km) | 0.71 | NDVI $_{wet}$ (6km) | 0.72 | Elevation | 0.84 |
| HPD (6k) | 0.59 | HPD (6km) | 0.53 | NDVI $_{dry}$ (12km) | 0.76 |
| Soil Carbon (12km) | 0.50 | Soil Carbon (12km) | 0.53 | Soil CEC (6km) | 0.65 |
| Soil CEC (12km) | 0.49 | Soil CEC (12km) | 0.49 | HPD (12km) | 0.49 |
| Precip $_{dry}$ (12km) | 0.46 | Precip $_{dry}$ (12km) | 0.48 | Human Impact | 0.39 |

| **Extirpated** | | **Low Suitability** | |
|---|---|---|---|
| Variable (scale) | Rank | Variable (scale) | Rank |
| Precip (12km) | 1.00 | % Human LC (12km) | 1.00 |
| % Human LC (12km) | 0.93 | MAT (12km) | 0.88 |
| NDVI StDv $_{dry}$ (12km) | 0.57 | NDVI $_{dry}$ (3km) | 0.75 |
| MAT (12km) | 0.50 | HPD (6km) | 0.66 |
| Precip $_{dry}$ (12km) | 0.48 | NDVI $_{wet}$ | 0.65 |
| HPD (12km) | 0.47 | Precip $_{dry}$ (12km) | 0.64 |
| Soil Carbon (12km) | 0.46 | Soil Carbon (12km) | 0.60 |
| Human Impact | 0.27 | Soil CEC (12km) | 0.52 |

Response curves showed a range of consistency between the 10 different PsA draws (Figure 4.5). Some variables, primarily percent human land cover and human population density, showed little variation between the random draws. Based on a visual comparison, the more important variables typically had greater consistency between draws than less important variables. Visual comparison of response curves of eight of the most commonly used variables between all PsA selection strategies indicated relatively high consistency in the response curves (Figure 4.6). For some variables, such as percent human land cover or soil carbon, the response was fairly consistent although the slope for each option varied. For other variables, such as the soil CEC, there was limited consistency in response across the selection strategies.

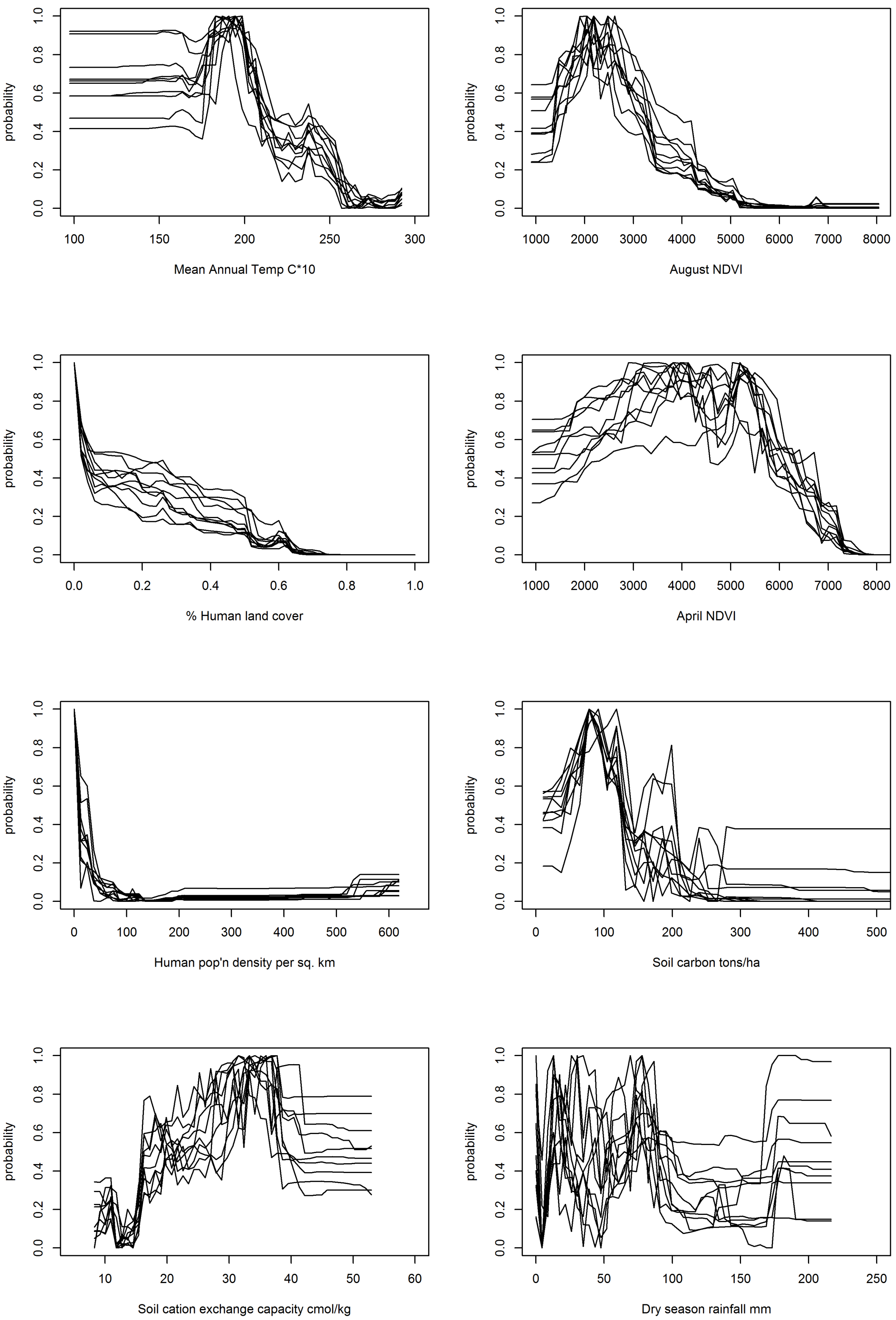


**Figure 4.5 Response curve variation across the 10 PsA pulls in the Full option. Variables are ranked in order of importance; starting at the upper left and working across and down. The y-axis represents probability of observation.**

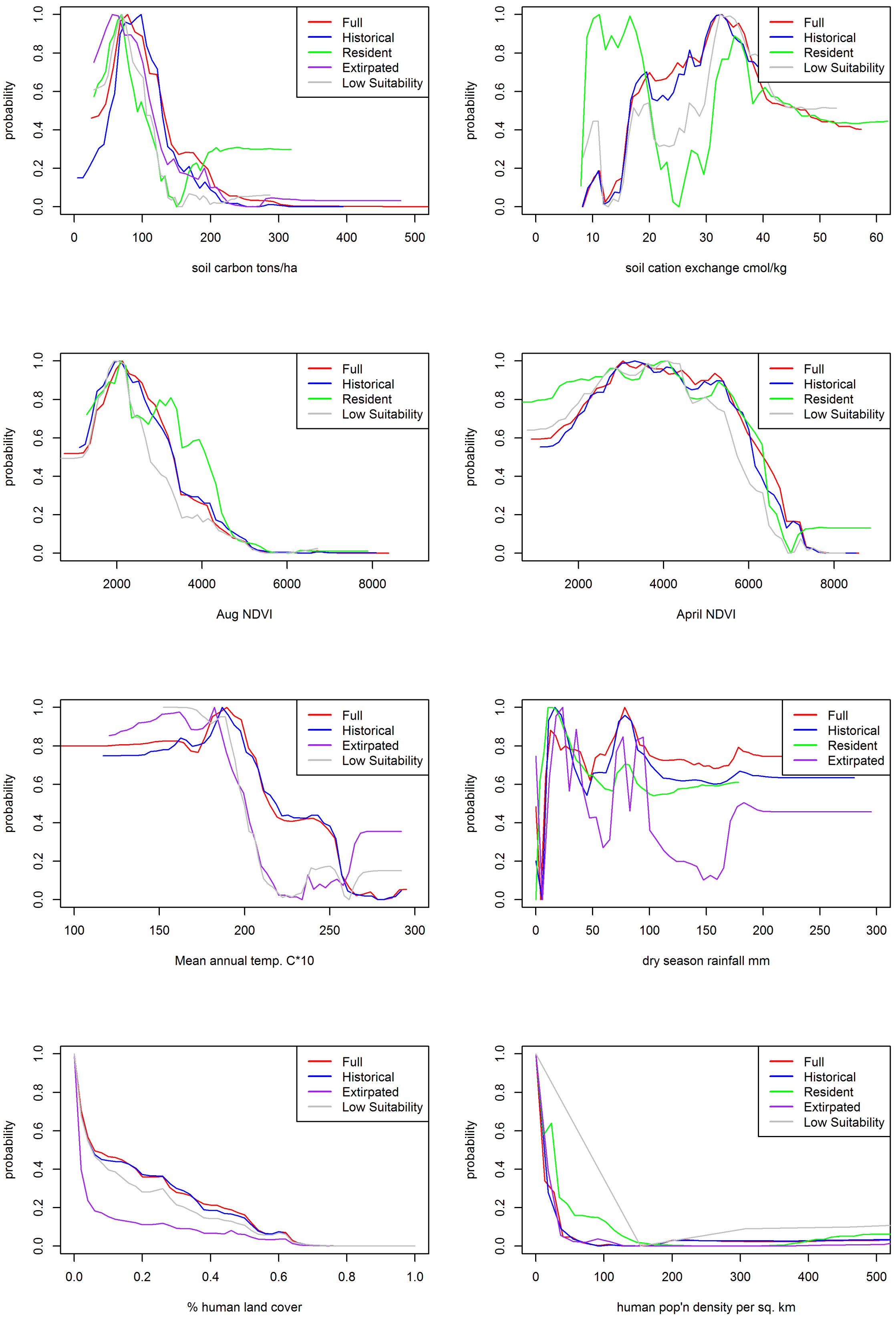


**Figure 4.6** **Variation in response curves between PsA generation strategies for the eight most common variables. The y-axis represents probability of observation.**

Within each selection strategy, correlation between the 10 different PsA draws varied (Table 4.5). Four of the five strategies had correlations 0.89 and greater across all 10 draws. The Extirpated method had the highest overall correlation between PsA draws. Correlation between the ensembles of the five different PsA strategies varied widely (Table 4.5). The Full and Historical methods produced highly similar output, while the Resident and Extirpated methods produced the least similar output. The spatial prediction from the Resident option was the greatest outlier of the group, although the Low Suitability option was also poorly correlated to the remaining options.

All models had high and relatively consistent AUC and Boyce Index scores although OOB error varied more substantially between PsA options (Table 4.6). All evaluation metrics had small standard deviation between different PsA pulls. The Full option had the highest Boyce Index although the Extirpated model had the lowest OOB error rate and highest AUC value.

Likelihood of observation in expert-derived zones varied widely (Table 4.7). In resident range where the highest probabilities are expected, mean likelihood of observation roughly doubled between the lowest and highest values, Resident and Extirpated options respectively. In extirpated range where low probabilities are expected, the Extirpated model produced the lowest mean values although the Full and Historical options were very similar. In the region outside the historical range of the cheetah, where probabilities close to zero are expected, the mean likelihood of observation was lowest in the Full option and highest in the Resident method.

**Table 4.5 Correlations of the spatial predictions among the different PsA selection strategies. Values on the diagonal are the 'internal consistency' of the model, aka the range of correlations between the 10 random PsA draws.**

| | Full | Historical | Resident | Extirpated | Low Suitability |
|---|---|---|---|---|---|
| Full | 0.89 – 0.95 | | | | |
| Historical | 0.987 | 0.92 – 0.95 | | | |
| Resident | 0.422 | 0.439 | 0.75 – 0.86 | | |
| Extirpated | 0.755 | 0.73 | 0.394 | 0.95 – 0.99 | |
| Low Suitability | 0.67 | 0.693 | 0.627 | 0.637 | 0.9 – 0.95 |

**Table 4.6 Mean and standard deviation of evaluation statistics for five different methods of selecting PsA points. The highest evaluation metrics are in bold.**

| | Boyce ± SD | AUC ± SD | OOB ± SD |
|---|---|---|---|
| Full | **0.89 ± 0.01** | 0.98 ± 0.00 | 0.15 ± 0.02 |
| Historical | 0.89 ± 0.01 | 0.98 ± 0.00 | 0.15 ± 0.01 |
| Resident | 0.89 ± 0.01 | 0.96 ± 0.01 | 0.35 ± 0.02 |
| Extirpated | 0.88 ± 0.01 | **0.99 ± 0.00** | **0.05 ± 0.01** |
| Low Suitability | 0.89 ± 0.01 | 0.97 ± 0.00 | 0.19 ± 0.01 |

**Table 4.7 Mean and standard deviation of the likelihood of observation in expert-derived distribution zones. The highest probability in resident range, and lowest probabilities in extirpated and outside historical range are in bold.**

| | Resident range | Extirpated range | Outside historical range |
|---|---|---|---|
| Full | 0.62 ± 0.26 | 0.11 ± 0.13 | **0.091 ± 0.11** |
| Historical | 0.61 ± 0.26 | 0.11 ± 0.13 | 0.102 ± 0.11 |
| Resident | 0.42 ± 0.20 | 0.36 ± 0.10 | 0.338 ± 0.09 |
| Extirpated | **0.81 ± 0.26** | **0.11 ± 0.17** | 0.116 ± 0.16 |
| Low Suitability | 0.57 ± 0.26 | 0.21 ± 0.17 | 0.263 ± 0.18 |

## 4.4 Discussion

Using presence data from the cheetah in East Africa, I demonstrate how the choice of PsA selection strategies alters modeling results and the ecological inferences of the output. I compared three widely used PsA selection strategies and two strategies developed here, each with distinct ecological inferences. Of these PsA strategies, two, where PsA data are drawn from the whole study extent (Full) and extirpated range of the cheetah (Extirpated), were found to be most useful in predicting cheetah range decline in East Africa. These two strategies identify the predictors associated with current cheetah distribution and those related to formerly suitable habitat. The other PsA generating strategies are likely valuable given alternative modeling questions although the Low Suitability option where PsA are drawn from different climatic regions makes little ecological sense. The various PsA generating strategies resulted in different sets of contrast points, some of which were more geographically similar (and hence similar in multivariate space) than others. This resulted in varied sets of important predictors, response curves, and spatial predictions. The Extirpated model had the highest evaluation criteria and the spatial prediction that best matched expert-derived distributional zones. Annual precipitation and percent human land cover were the most important predictors in distinguishing current from extirpated range.

This study chose to focus on the impact of varying PsA data rather than on the impact of bias in the presence data. Substantial prior efforts have been made to understand these impacts (Dormann et al. 2007; Phillips et al. 2009). The impact of biased presence data is probably best observed through simulation experiments where there is a known distribution and the training data can be purposefully biased in multiple ways and results compared to the known distribution. This study acknowledges that the training data are biased. A similar argument could be made that the impact of different PsA selection strategies should be seen via simulation; indeed this is what has been done to this point - understanding the statistical impacts of different PsA selection strategies. Yet, it's important to understand the ecological implications of how various PsA selection strategies impact the modeling question, and this is relatively poorly covered in the existing literature and is best done with a case study.

### 4.4.1 Modeling outcomes

I followed best practices in generating PsA points for Random Forests (Barbet-Massin et al. 2012). I maintained an equal weighting between PsA and presence data by generating 5,000 points, subsampling an equal number of PsA as presence data and averaging the result across the multiple runs.

Generating PsA data from different strategies fundamentally alters the geographic locations environmental data are sampled from and hence their location in multivariate space. Thus, variation in variable importance, response curves, model prediction and model performance should be expected across different PsA generation strategies.

Variable selection and importance was influenced by the PsA selection strategy. The scale of the predictor also varied although averaging at the largest scale (12 km) was by far the most common. Only one time was a variable selected at the 500 m scale, and this was for the Resident option. Only human population density and soil carbon were selected in all five options. Cheetah habitat in East Africa is highly fragmented and, as expected, anthropogenic factors were important in all options, although they were least important in the Resident option. This is probably because areas of resident cheetah range have relatively few people or little habitat conversion.

In general, resource gradients are expected to be a unimodal curve, going from "too little" to "too much," with "just right" in the middle (Austin 2002; Merow et al. 2014a). Previous literature has shown that to get a full response curve, the study extent must be large enough to capture a range of environmental conditions that encompass the species' environmental tolerances (Thuiller et al. 2004). The response curves for both wet and dry season NDVI, for soil carbon and CEC showed mostly unimodal response curves although with some truncation of the curves at the extremes. Both mean annual temperature and dry season rainfall had even greater truncation of values at the extremes. This was expected, as although East Africa is a heterogeneous region, it is only a portion of cheetah range. As expected with anthropogenic variables, there was a strong negative relationship.

Response curves varied within model options, sometimes substantially. Within an option (i.e. across the 10 PsA pulls), some level of variation is expected. Indeed, greater variation is expected at the extremes as there are fewer data at these resource levels and hence a response curve is more likely to be influenced by the chance inclusion of a handful of points. Similarly, this is unlikely to have significant impact on model output as this occurs only at more extreme values. In addition, greater variation may be expected for less important predictors. Overall, I did not quantify how much variation there was within the response curves nor is it known if there is a threshold at which there becomes too much variation and the output treated skeptically. Further research here is needed.

The Full and Historical options were highly correlated while the output of the Resident and Extirpated options were poorly correlated. The highly correlated output was expected in the Full and Historical options as PsA were drawn from largely similar regions since only the Albertine rift and coastal vegetation along the Indian Ocean were outside of historical range. They also were very similar in terms of variable importance and response curves. Different model parameters could lead to greater differentiation of results (i.e., a larger study extent that included more non-historical range). The Resident and Extirpated options result in PsA drawn from non-overlapping areas of expert-defined resident or extirpated range. Between these two options, evaluation statistics, variable importance, and response curves are markedly different, and resulted in a poorly correlated spatial prediction.

The spatial predictions from the 10 random PsA draws were highly consistent for all but the Resident option. Interestingly, the worst internal correlation was from the method with the strongest geographic constraint placed on the PsA points, i.e. they were drawn from the smallest geographic region. Yet this option forced the PsA points to be drawn in close proximity to the presence data, which likely affected discrimination between presence and PsA. The Resident option also had the worst OOB error, was likely overfit, and the spatial prediction is unrealistic.

The Extirpated model had the highest mean likelihood of observation in resident range and lowest likelihood of observation in extirpated range. The Extirpated model achieved the best discrimination between resident and extirpated range, and provided the closest match to the expert-derived range maps. The Full option, the only option to select PsA points outside of historical range achieved the lowest mean likelihood of observation in regions outside of historical range. This suggests that PsA points selected from areas outside of historical range may have contributed to properly identifying what its historical range was (similar to results from Elith et al. 2011).

The mean likelihood of observation was highest within resident range as expected, but was essentially the same for both extirpated range and outside historical range. I expected, but did not find, the mean value outside historical range to be lower than the extirpated range and close to zero. This may be because I did not select the most suitable predictors or because the predictors I did select did not find clear thresholds separating historical range from unsuitable habitat. However, only the Full option may be expected to pertain to this, as only the Full option could contrast presence with PsA selected from outside historical range. Indeed, the Full option had the lowest mean likelihood of observation in this category. Overall the Full and Extirpated options most closely matched the expert-drawn range map. The Resident and Low Suitability options gave the highest values in extirpated range, and lowest in resident range suggesting that these options did a poor job matching the expert-derived distribution. Yet, I did not expect a good match to the current distribution for the Resident option as this strategy contrasts where cheetah were actually found relative to other areas they could likely be. Hence, this option is identifying predictors that distinguish places cheetah were seen within known cheetah range (similar to 3rd order habitat selection (Johnson 1980)) and hence it may be appropriate to restrict the output of this option only to resident areas. Thus, this option may help distinguish areas of increased probability of presence within known range and have less relevance to non-habitat. However, the Low Suitability option has little ecological relevance and also matched poorly with the known distribution.

The predictive performance of the different model options varied but in general, all the options had good evaluation statistics. All the AUC values were >0.96. The Full and Extirpated options had the highest evaluation criteria overall. The Resident option had the worst AUC and a substantially higher OOB error than the other options. Curiously, the

Resident option had one of the highest Boyce Index values although all those values were quite similar.

Although it should be noted, that the Full option worked well in this situation since the cheetah was historically present across ~95% of the study area. Indeed, the historical range of the species extends far beyond these boundaries. If, on the other hand, the species of interest had a very small historical range within East Africa, then a more restrictive extent from which to select PsA may be more appropriate. In addition, if the historical range of the species is unknown, then it may be more challenging to estimate the appropriate size of the study extent. In this case, some reasonably and biologically appropriate buffer could be determined, i.e. some distance potentially reachable by dispersal over time (Soberón and Peterson 2005; Barve et al. 2011). Running a SDM and ensuring there is a full response curve may be a useful to know when a large enough buffer has been selected.

### 4.4.2 Theory and concepts

Most discriminative SDM algorithms rely on, or at least conceptually can be thought to use, three resource gradient curves. The three curves are the background environmental conditions in the study extent (i.e., a univariate histogram), what the algorithm believes are the background environmental conditions via the presence and PsA data (the sampled background), and the species response to each predictor variable. The true environmental background is likely unknown. The algorithm is given a limited amount of data from which to replicate the true environmental background; the fewer the number of points or the more biased they are, the less likely the algorithm will properly represent the true background. The algorithm uses the limited presence data to estimate the species response to different resource gradients. The algorithm compares the observed presences against the sampled background to generate the species' response curves (see Merow, Smith and Silander 2013). The more presence data the better the algorithm can generate the species response curve; and the more PsA data (along with presence data), the better the algorithm can match the sampled background to the true environmental background. Errors in either curve produce worse model output. Two of these curves, the sampled background and observed presences, are explicitly considered in the Maxent program, and termed probability density distributions (Phillips, Anderson, and Schapire 2006; Elith et al. 2011; Merow, Smith, and Silander 2013). With decision tree algorithms like Random Forests, the curves are more metaphorical and can be useful as an aid to understanding how the algorithm works. I suspect that the Maxent algorithm typically has high evaluation metrics (Elith et al. 2006) because the program default is to draw 10,000 background points, thus making the sample background curve very similar to the true background.

However, these three curves do not assist in understanding how to generate PsA or how they be constrained. More research is needed to determine the appropriate balance between ecological inference and statistical best practice but I identify two key concepts in generating PsA. The most important aspect is to match the strategy with your research question and obtain a good characterization of the background. Manly et al. (2002) argued that PsA data should be generated to minimize sampling bias and to truthfully represent the background of the study area. In only the simplest sense, is the background the full study extent; at a minimum, for terrestrial species, water bodies are masked out. In other cases, additional habitats will be removed from the study extent because of their known unsuitability for the species. Indeed, the research question will affect the choice of study extent and what areas are masked (see Elith et al. 2011).

Generally speaking, its known how to achieve higher evaluation scores – greater ecological distance between presence and PsA data (Elith et al. 2011; Senay, Worner, and Ikeda 2013; Hertzog, Besnard, and Jay-Robert 2014). This makes the prediction task easier, possibly why narrow-ranged endemics tend to have higher evaluation scores (Franklin 2009). Thus, the goal for PsA generation is not to optimize model performance by finding a way to achieve the greatest difference in environmental space from your presence data (Franklin 2009; Elith et al. 2011; Saupe et al. 2012), but to select an appropriate method(s) that will address the research question.

The second aspect to consider is how many PsA should be drawn. Barbet-Massin et al. (2012) recommend different approaches to this question based on the choice of algorithm; for Random Forests and other decision tree models they recommend balancing the number of PsA points with the number of presence points. A larger number of points may be needed to characterize a larger study area, or one with greater environmental variability (Barbet-Massin et al. 2012). Therefore, the number of PsA points to generate is related to the modeling algorithm, the extent, the grain of the analysis, and the composition of the landscape.

### 4.4.3 Ecological interpretation of pseudo-absence generation strategies

The modeling questions posed by the different PsA generation strategies varied substantially across the five options. The Full and Historical options both posed broad scale, biogeographic or "macrohabitat" questions, comparing cheetah presence to where they may have occurred or could have been seen. The Resident option posed a more local scale question, where cheetah were actually found relative to other areas they could likely be. This option, and selecting PsA from within a small buffer around presence data which would contrast the distribution of presence locations with locally available sites, may be more useful with finer-scale telemetry data. However, alternative analysis methods are likely preferable to distribution modeling for this type of question (Pearce and Boyce 2006), such as resource selection functions (Manly et al. 2002). The implications of the Resident and Extirpated models were perhaps most different from each other, asking questions about where cheetah were found in relation to known resident or extirpated range. The inference from the fifth option, Low Suitability, had little ecological meaning as it contrasted presence data with areas that were in a different climatic region, which were not necessarily bound by historical presence. It's likely that this option achieved high evaluation statistics (Zaniewski, Lehmann, and Overton 2002; Engler, Guisan, and Rechsteiner 2004), because it was enhancing the differences between presence and PsA data in environmental space. Thus, it is important that PsA generation strategies are explicitly linked to the research question, such that the appropriate ecological inferences can be made.

As PsA generation techniques change, so must interpretation of their output. For instance, with variable importance, April (wet season) NDVI at 500m was the most important variable for the Resident method while the mean annual temperature averaged at 6 km scale was the most important for the Full option. This suggests that fine scale differences in April NDVI were important when comparing presences to other locations within resident range but that at broader scales when looking at cheetah presence across all of East Africa, mean annual temperature was a more important variable in determining presence.

Interpretation of response curves between the different PsA options in particular requires a careful approach. For each response curve, what is the presence data contrasting with/what is it relative to? For instance, when comparing response curves for percent human land cover, all options show declining trends as human land cover increases towards 100%. However, the initial slope was most negative in the Extirpated model, and least negative in the Full and Historical options. This difference may result from the high overlap of human land cover with extirpated areas, while across the entire study area or historical range, percent human land cover was less of a driving variable in determining habitat. The response curves for soil CEC between the different options were especially dissimilar. Take the response curve for the Resident method that exhibited an implausible bimodal response. This suggested that presences were highly localized in both low and high levels of CEC, while they were not commonly found within resident range at very low or medium levels of CEC. This is most likely a result of sampling bias rather than a behavioral response to the environment. Thus, reading a response curve from presence-PsA data is not simple and must be interpreted according to what the contrast data represent.

### 4.4.4 Why are cheetahs declining across the landscape?

The Full and Extirpated options were most relevant to the study aim of investigating predictors related to cheetah range decline in East Africa. Comparing the important variables, spatial predictions, and response curves between these options builds inference into the factors associated with cheetah distribution and their decline.

No previous global or regional scale analysis of cheetah distribution exists, only local scale habitat use studies within protected areas (see Pettorelli et al. 2009; Durant et al. 2010a; Rostro-García, Kamler, and Hunter 2015). Hence, the existing publications focus on habitat selection at fine scales, in the 3rd order, rather than at the global or regional scale (1st order; Johnson 1980). Mean annual temperature, August (dry season) NDVI, and percent human land cover were clustered as the most important variables in determining cheetah presence relative to the entire study area. Previous research has suggested that temperature affects cheetah activity (Cozzi et al. 2012), and they avoid human settlements (Pettorelli et al. 2009). Annual precipitation and percent human land cover were the two dominant variables in determining cheetah presence relative to where they used to be found. Surprisingly, percent human land cover and human population density did not become substantially more important variables in the Extirpated option compared to the Full option, although the third anthropogenic variable, human pressure, was included. But it should be noted that the Extirpated model may accentuate the importance of anthropogenic variables as the area is climatologically suitable since it was historically occupied.

The spatial predictions of these two options were similar although the Extirpated option predicted more cheetah habitat in southwestern Tanzania, and northern and eastern Kenya. Both options produced plausible spatial predictions, with the Extirpated model representing more of a best-case scenario for cheetah persistence. Indeed, the model may have overemphasized regions that are largely free of human impact while including potentially less biologically suitable habitats (such as the Mathew's Mountain Range in Kenya).

Possibly the most compelling reason for comparing these two options in order to understand cheetah range decline, is to contrast the two sets of response curves. This allows the direct comparison of what the response curve is for determining cheetah presence at a regional level vs. a response curve for distinguishing current from former habitat. A few variables are worth investigating. With mean annual temperature, the response curve for the Extirpated option is higher at both extremes while depressed in the middle (Figure 4.6). With dry season rainfall, the response curve for the Extirpated option is much lower at all values above ~100 mm (Figure 4.6). This is likely due to the fact that higher levels of dry season rainfall (i.e., more consistent year-round rain) are associated with human presence. This suggests that cheetahs may fail to persist in once suitable habitats to live in more extreme environments (mostly hotter and drier). Indeed, this difference is particularly obvious when comparing the response curves for annual precipitation (Figure 4.7). This was the most important variable for the Extirpated option but since it is highly correlated with NDVI was not selected in any other option. To investigate the cheetah failing to persist in previously suitable habitats, I ran the Full model a second time substituting annual precipitation for the NDVI predictor. Then, I compared the response curves between the Extirpated and modified Full option and indeed the Extirpated response curve was shifted towards the dry end with probability of presence depressed at medium rainfall values. A shift towards more use of hot and dry climates matches the spatial prediction of the Extirpated option where NE Kenya has high likelihood of observation.

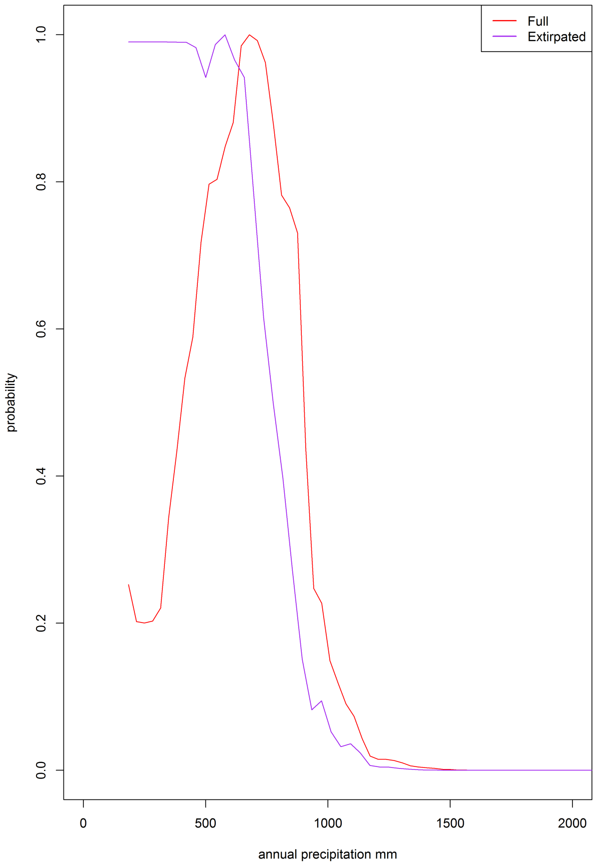
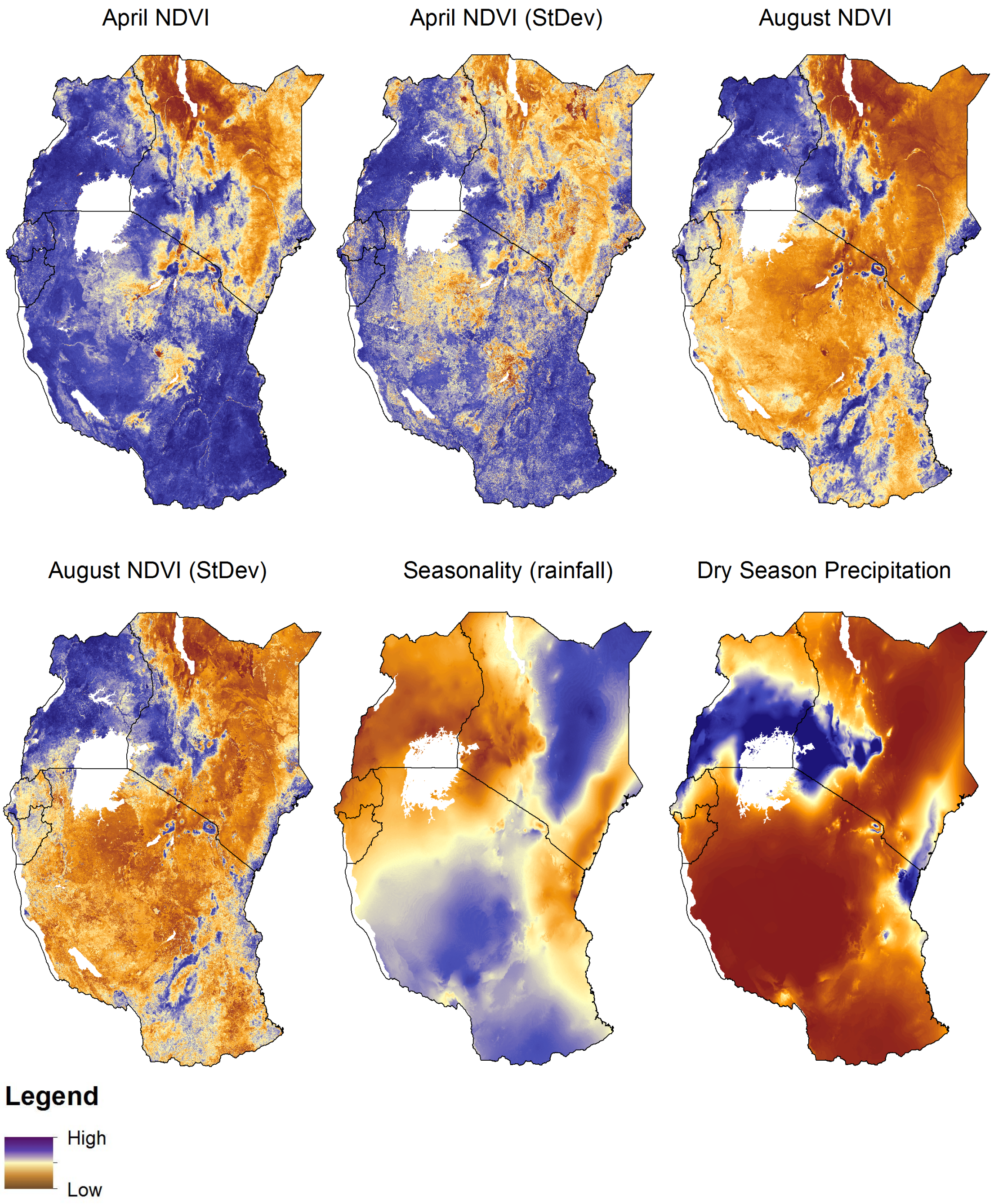


**Figure 4.7 Variation in response curves between the Full and Extirpated options for annual precipitation. I modified the Full option in this case to replace NDVI with annual precipitation for demonstration purposes.**

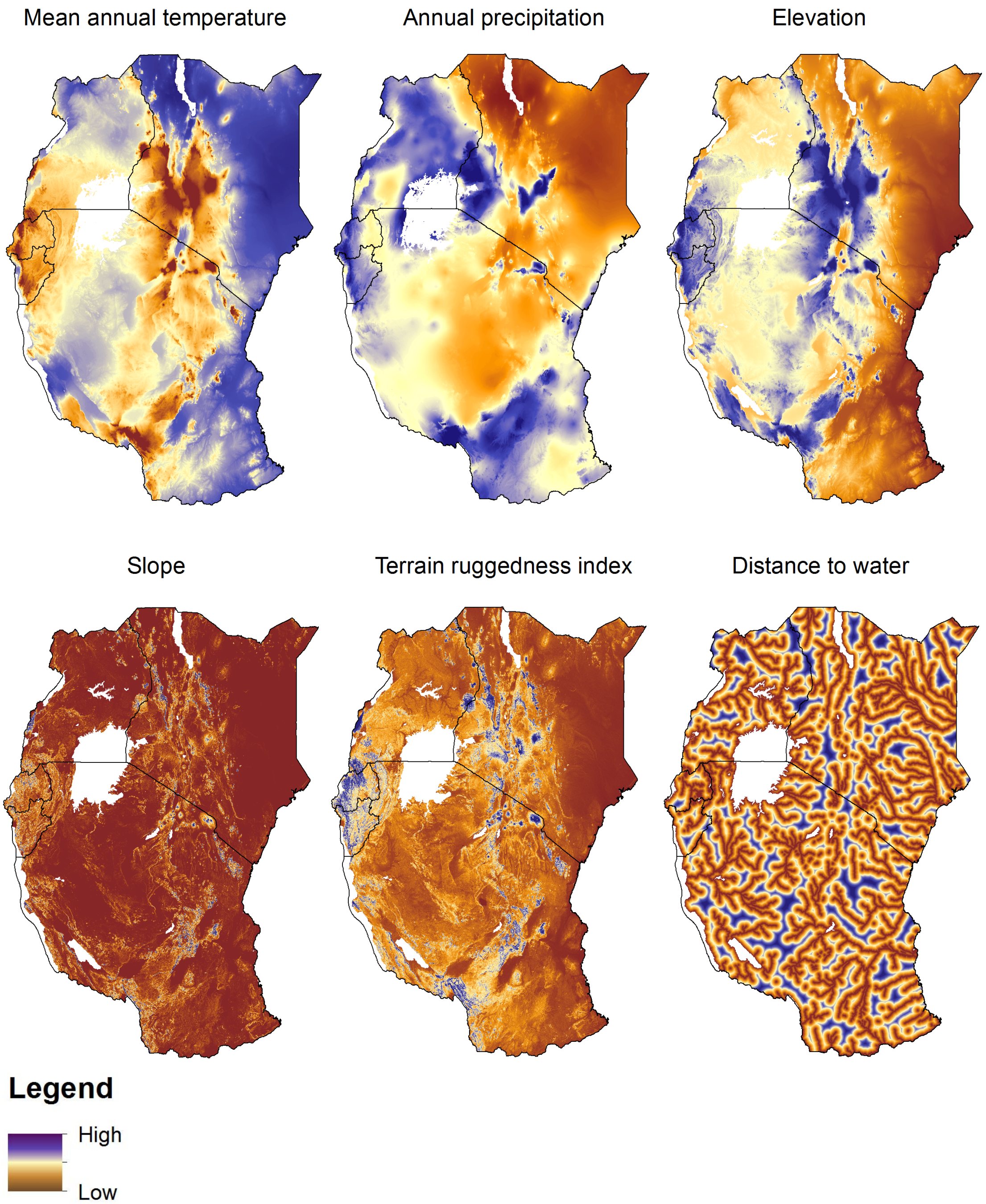
## 4.5 Conclusions

Recently, Guillera-Arroita et al. (2015) argued that the type of presence data, along with the issues of sampling bias, and imperfect detection jointly affect the meaning of SDM output. They argued the meaning of SDM output varies from the relative likelihood of observation (lowest information content) to the actual probability of occurrence (greatest information content). This may help disabuse modelers from thinking that species distribution model output actually represents the real or potential distribution of a species. Extending the conclusions of Guillera-Arroita et al. (2015), if PsA data are used, then the method of selecting PsA data will also influence the modeling question and interpretation of results. While the quality of presence data may influence what the output means (Guillera-Arroita et al. 2015), the ecological inference one can draw will vary depending on how one generates the contrast data.

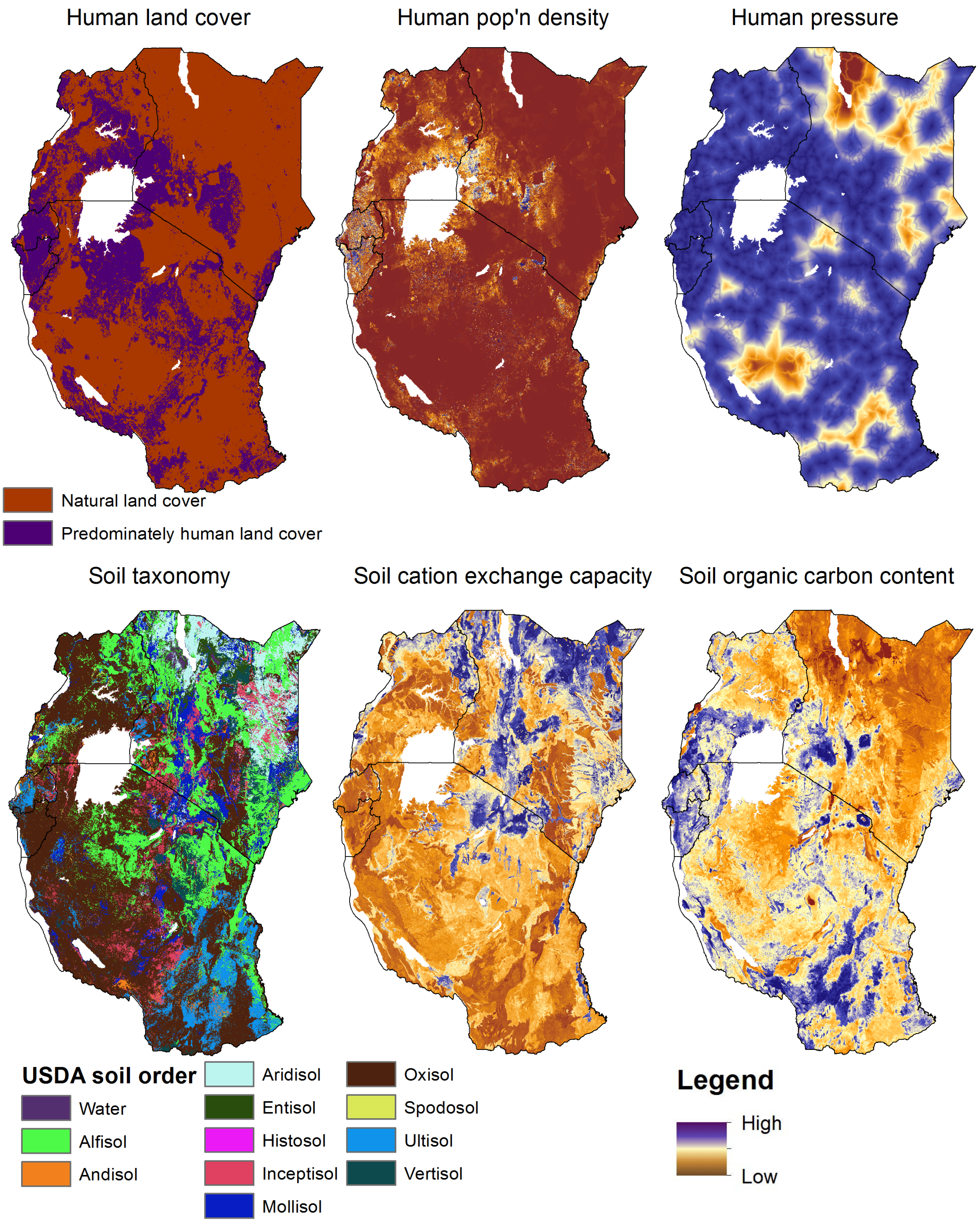
As opposed to chasing the PsA selection strategy that produces the highest evaluation metric, the modeling question should be matched to the proper method of selecting PsA data. Some of the existing PsA generation strategies in fact have little ecological meaning. The PsA generation strategy along with the study extent should be identified at the outset in order to match the research question to the inferences one can draw from the SDM output. Indeed, one can build ecological inference by combining different PsA generation strategies and contrasting predictions, the influence of predictors or shapes of response curves. Multiple comparisons, like using multiple SDM algorithms, may help build support for the outcome and strengthen ecological conclusions.



**Supplemental Figure 4.1** **Predictors used in the analysis.**



**Supplemental Figure 4.2** **Predictors used in the analysis.**



**Supplemental Figure 4.3** **Predictors used in the analysis.**

# Chapter 5

## People or the plow? Using distribution models to assess predictors associated with large carnivore range decline in East Africa

Lion skulls in various stages of decomposition at the Ruaha Carnivore Project camp in southern Tanzania. © Andrew Jacobson

# Chapter 5 People or the plow? Using distribution models to assess predictors associated with large carnivore range decline in East Africa

## Abstract

The ability of large carnivores to coexist in human-dominated landscapes has become a pre-requisite for their survival in many parts of the world. This recognition has driven greater research into understanding what anthropogenic factors are associated with continued carnivore presence, or extinction. As shown in Chapter 4, comparing species distribution models with pseudo-absences (PsA) generated separately from the full study area and extirpated areas can provide an effective approach to assess species' response to threats. Added to this, I develop a third modeling approach that predicts extinction probability. Here, these three modeling approaches are used to build ecological inference to answer Research Aim III - to identify the most important predictors of presence for four large carnivores (cheetah, African wild dog, leopard and lion), and the environmental factors correlated with their decline in East Africa. Using this approach I test three hypotheses: increasing anthropogenic impacts threaten large carnivore persistence; carnivores with larger home ranges are more sensitive to habitat loss; and habitat loss is a greater factor impacting carnivore distributions than human population density. I find evidence to support each hypothesis in that 1) increasing human population densities and human land cover negatively impact carnivore presence; 2) cheetah and leopard are the most and least vulnerable species to increasing human land cover respectively; and 3) that human land cover is a greater threat to carnivore distribution than human populations per se. In addition, I compare spatial predictions of the models, combined with thresholds of human population density and human land cover, with expert-derived species range maps. This enables identification of current range that may be gravely threatened by human impacts, as well as areas of potentially suitable habitat outside known range. This approach makes use of widely available data for many species, and thus has general applicability to modeling distributions and assessing threat.

## 5.1 Introduction

Threats to biodiversity continue to increase (Butchart et al. 2010) yet certain mammals are more predisposed to extinction via a combination of intrinsic biological traits and external factors such as anthropogenic and environmental influences (Purvis et al. 2000;

Cardillo et al. 2005; Di Marco et al. 2014). Many risk factors scale with body mass and large-bodied carnivores are particularly susceptible to extinction (Cardillo et al. 2005; Davidson et al. 2009). Problematically, carnivore loss challenges ecological systems as carnivores regulate ecosystems and perform valuable ecological services (Terborgh et al. 2001; Ray et al. 2005; Ripple et al. 2014).

Human activities directly and indirectly imperil carnivores. Humans intentionally kill carnivores for a variety of reasons and with a variety of methods, such as shooting, poisoning, trapping etc. (Kruuk 2002). Carnivores are also incidentally killed such as via snaring by-catch (Becker et al. 2013a; Lindsey et al. 2013b) and on roads (Woodroffe and Ginsberg 1999; Swanepoel et al. 2015). Furthermore, human activities can indirectly impact carnivores, such as via prey depletion (through snaring etc.) (Lindsey et al. 2011; Lindsey et al. 2013b; Wolf and Ripple 2016), competition for space/resources by humans and domesticated animals (Oriol-Cotterill et al. 2015a), and through the degradation and conversion of natural habitats to human dominated land cover (i.e. agriculture and urban land) (Ehrlich and Holdren 1971; Vitousek et al. 1997; Foley et al. 2005; Ripple et al. 2014). However, many factors such as land use, policy, culture, and dependence on local natural resources can affect the relationship between human populations and activities that harm wildlife (Linnell, Swenson, and Andersen 2001; Dickman 2010). Thus growing human populations have been described as an “insidious threat” to wildlife (Ripple et al. 2014), but their impact can be mediated.

Habitat loss is one of the key consequences of increasing human populations, and has long been recognized as a preeminent risk to biodiversity (Fahrig 1997; Fahrig 2003). Habitat loss has a multitude of negative impacts including effects on species richness, genetic diversity, population growth rates, species interactions, and breeding, dispersal and foraging success (Fahrig 2003). Habitat loss is primarily driven by agricultural expansion and urbanization (Foley et al. 2005). Home range size is one of several, often linked, biological traits that are associated with increased susceptibility to habitat loss and extinction (Cardillo et al. 2005; Davidson et al. 2009). Species with larger home ranges are thought to be more sensitive to habitat loss as they are more prone to edge effects (Woodroffe and Ginsberg 1998; Cardillo et al. 2008). Ranging behavior affects the intensity of interactions with humans, and those species that traverse greater distances are more vulnerable due to more frequent contact with humans and human-related disturbances (Woodroffe and Ginsberg 1998). Human population density and the amount of habitat loss are two commonly used metrics to assess pressures to wildlife (Woodroffe 2000; Crooks 2002; Cardillo et al. 2004; Cardillo et al. 2008).

Protected areas can serve as important reservoirs of large carnivore populations, by separating them from human impacts and habitat loss, but most are too small to contain viable populations themselves (Linnell et al. 2005; Durant et al. 2010b; Lopez-Bao et al. 2015). As the ability of large carnivores to coexist in human-dominated landscapes has become a prerequisite for their survival in many parts of the world (Lopez-Bao et al. 2015), understanding what anthropogenic factors, and their thresholds, are associated with presence or extinction has become increasingly important (Fahrig 2001; Burdett et al. 2010; Bouyer et al. 2014).

East Africa is one of the fastest growing regions in the world (Bongaarts 2009; UN 2013). Several countries' populations in the region could quintuple by 2100 (UN 2013). Agriculture and rangelands are expanding rapidly, as are East African economies (Ahlers et al. 2014). Yet, East Africa is of particular importance for carnivores (Ray, Hunter, and Zigouris 2005), such as the lion, and cheetah. As East Africa continues to transform, identifying the factors associated with carnivore presence are critical in order to assist in their conservation.

A variety of studies have examined the environmental and anthropogenic factors affecting large carnivore distribution and habitat use in Africa. These studies, primarily at site or country-level scales, have documented the avoidance of large carnivores to human settlements and croplands including: cheetah (Pettorelli et al. 2009; Pettorelli et al. 2010), leopard (Toni and Lodé 2013; Swanepoel et al. 2013), lion (Schuette, Creel, and Christianson 2013; Elliot et al. 2014b) and wild dogs (Whittington-Jones et al. 2011; Jackson et al. 2016). Abade, Macdonald, and Dickman (2014) found a slightly positive relationship between lions and leopards with settlements near Ruaha National Park (NP), but this may have been due to greater search effort near the park boundary. Overall, carnivore species richness is significantly lower in cultivated areas in contrast to protected or pastoral areas (Msuha et al. 2012; Kiffner et al. 2014). In sum, prior research has shown large carnivores avoid human settlements, areas of high population density, and croplands generally, but the relative sensitivities among carnivores and between these threats is less known.

Understanding the relative importance of human populations and human land cover on wildlife as well as the differences in susceptibility of species to these threats is important. Species respond differently to the same threat (Harcourt, Parks, and Woodroffe 2001), even those within a guild (Janecka et al. 2016). For instance, the observed difference in tolerance of cheetah and leopard (cheetah is less tolerant) to humans has been noticed for a long time (Myers 1975). Furthermore, there may be an important distinction in the implication of using human land cover and human population density. While the two factors are obviously related, pastoralists live at relatively low human densities and rarely cultivate. On the other hand, farming regions can have extensive areas of cropland also with low human densities. Both situations have low densities of people, but different amounts of habitat loss. I try to tease apart if the presence of people is enough to deter carnivores, or if carnivore absence is driven by habitat loss. If the differences are substantial, the result could influence management objectives, affecting pastoralism practices and land use planning.

Also, it is important to understand species responses to human impacts at different spatial scales for conservation purposes and for management (Sunquist and Sunquist 2001; De Angelo, Paviolo, and Di Bitetti 2011). Yet, there have been no published distribution models conducted at regional scales (larger than a country) for large African carnivores although important biological processes, such as metapopulation dynamics, occur at this scale (Hanski 1998; Hanski and Ovaskainen 2000).

This study aims to identify the most important predictors of presence for four members of the large carnivore guild in East Africa, cheetah, wild dog, leopard, and lion, and the environmental factors correlated with their decline. This is accomplished by using the Random Forests algorithm (Breiman 2001), a widely used distribution model (Cutler et al. 2007; Watling et al. 2012; Ahmed et al. 2015). Drawing pseudo-absences (PsA) from different regions as demonstrated in Chapter 4 is used to assist and strengthen ecological conclusions. Here I use this process to identify areas of conservation importance and test several ecological hypotheses aimed at distinguishing between the impacts of human density and land use change on the persistence of large carnivores (Table 5.1). These hypotheses are: 1) increasing anthropogenic impacts threaten large carnivore persistence; 2) carnivores with larger home ranges are more sensitive to habitat loss; and 3) habitat loss is a greater factor impacting carnivore distributions than human population density.

**Table 5.1 Hypothesis and prediction framework.**

| General hypothesis and justification | Prediction(s) |
|---|---|
| **H1: Increasing anthropogenic impacts threaten large carnivore persistence.**<br>H1A: Human populations are associated with increased resource demands (Ehrlich and Holdren 1971), increased demand for meat (either through domestic or wild sources) (Ripple et al. 2014), biodiversity loss and prey depletion (Vitousek et al. 1997), direct persecution of carnivores (Woodroffe and Ginsberg 1998; Woodroffe 2000; Harcourt, Parks, and Woodroffe 2001), and habitat loss via land conversion (Ehrlich and Holdren 1971; Vitousek et al. 1997; Foley et al. 2005).<br>H1B: Conversion of natural habitat to anthropogenic land cover reduces biodiversity (Pimm and Raven 2000; Fahrig 2003; Haddad et al. 2015), prey biomass (Worden, Reid, and Gichohi 2003; Reid, Thornton, and Kruska 2004; Ripple et al. 2015), impacts foraging opportunities (Benton, Vickery, and Wilson 2003; Stewart et al. 2016), reduces denning or refuge quality (Oriol-Cotterill et al. 2015b), and increases chance of sighting (Oriol-Cotterill et al. 2015b). | Prediction 1A: Increasing human population densities negatively impact large carnivore presence.<br>Prediction 1B: Increasing human land cover negatively impacts large carnivore presence. |
| **H2: Carnivores with larger home ranges are more sensitive to habitat loss.**<br>Larger home ranges increase opportunity for threats due to human impacts on protected area boundaries i.e. edge effects (Woodroffe and Ginsberg 1998; Cardillo et al. 2008). | Prediction 2: Cheetah have the largest home ranges and are most vulnerable to increases in human land cover, while wild dogs, lion and leopard are less vulnerable in that order. |
| **H3: Habitat loss is a greater factor impacting carnivore distributions than human population density.**<br>Habitat loss is a more proximate threat than human population density. Although both can be mediated to an extent, the impact of human populations on carnivore presence is more variable due to: cultural/religious tolerance and social issues (Dickman 2010; Li et al. 2013); differential reliance on natural resources and hunting (Homewood and Rodgers 1991; Nelson 2009); governmental laws and wildlife management structures (Linnell, Swenson, and Andersen 2001; Balme, Slotow, and Hunter 2009; Dickman, Macdonald, and Macdonald | Prediction 3: Human land cover is a more important variable restricting carnivore distributions than human population density. |

2011); national security/war (Kanyamibwa 1998; Douglas and Alie 2014); and ecosystem productivity (e.g. remaining cover and prey densities) (Athreya et al. 2015).

## 5.2 Methods

This chapter uses distribution models to explore large carnivore range decline. I use the same study area and predictors as Chapter 4; see that chapter for details. However, this chapter investigates a larger set of carnivore species and uses a subset of the PsA selection techniques.

### 5.2.1 Study area, species data and predictor variables

In this chapter, four carnivores are investigated, the cheetah, wild dog, leopard, and lion. I used the same cheetah presence data as described in Chapter 4 and used the same data collection methods for the wild dog and lion, but focused more heavily on published literature for leopard presences. Data collection lasted approximately two years, from March 2013 to February 2015. Presence records came from a variety of sources including government authorities, research projects, conservancy game scouts, and individual observations (Table 5.2). There were 35 total data contributors, although only one was from Uganda and no data points were gathered from Burundi or Rwanda. Data types included telemetry data, playback surveys, conflict records, scat identification, camera trap records, and sightings, although the majority of data were telemetry records and sightings. Lion, leopard, and a single wild dog record were the only data collected in Uganda. I collated ~70,000 lion, ~17,500 wild dog, ~10,000 cheetah, and 545 leopard records. I digitized records from two published maps pinpointing leopard presences (UWA 2010; Andanje et al. 2016). Data were standardized, duplicates and data older than the year 2000 removed, and some points withheld due to reliability or accuracy concerns. The year 2000 was chosen as an easily communicated cutoff when requesting data. I also visually checked for outliers using a scatterplot matrix and removed two cheetah points, one wild dog, and one lion point. These locations may have been mistakenly attributed or recorded in a nearby city rather than where the individual was actually spotted. I spatially rarefied the occurrence data to reduce spatial autocorrelation by removing all duplicate points within 10 km, as implemented in the SDM Toolbox (Brown 2014). This distance was about the average of the daily distance moved for the wild dog (Pomilia, Mcnutt, and

Jordan 2015; Hubel et al. 2016a) the species with the largest average daily distance moved. This also served to reduce the discrepancy between heavily surveyed locations and other regions with fewer observers (e.g. western Tanzania).

After I spatially rarefied the occurrence data, a total of 261 presence records were used in model training for the cheetah, 256 for the wild dog, 223 for the leopard, and 399 for the lion. These data suffer from sampling bias because researchers sampled some regions of East Africa more intensively than others. Additionally, the species suffer from imperfect detection, as species can be more difficult to see in densely vegetated habitats. Finally, since the data have been collected opportunistically over time and from various sources, prevalence cannot be determined (Guillera-Arroita et al. 2015). The majority of SDM studies suffer from the same issues (Guillera-Arroita et al. 2015).

All four large carnivores are declining but continue to range outside of East Africa (Woodroffe and Sillero-Zubiri 2012; Bauer et al. 2015; Durant et al. 2015; Jacobson et al. 2016). As habitat generalists, each species historically ranged widely inside East Africa (Table 5.3). Cheetah historical range was detailed in the previous chapter. Wild dogs were excluded from Albertine Rift montane forest (Olson et al. 2001; IUCN/SSC 2007). Leopards ranged completely across the region (Jacobson et al. 2016). Lions were only excluded from areas above 3,500 m in elevation as this was approximately the maximum recorded elevation in the study area (Nowell and Jackson 1996).

Each species had expert-derived distribution maps that included areas of resident (or extant) and extirpated range along with areas of unknown or temporary occupation. Range maps for cheetah and wild dog were created through a regional workshop organized by the Range Wide Conservation Program for cheetah and wild dog and detailed in Box 4.2 (IUCN/SSC 2007). The current lion range map is detailed in Riggio et al. (2013) and has been slightly updated since then (Amy Dickman et al. unpublished data). The provenance for the leopard range map is Jacobson et al. (2016).

Predictor variables were the same as those in Chapter 4. I also reviewed additional variables for potential inclusion, including distance to settlement, livestock densities, and protection status. I chose not to include distance to settlement as some were inaccurately located (as assessed in Google Earth), and even small settlements impact carnivore movement (Schuette, Creel, and Christianson 2013) and only larger settlements were included in the data set. Livestock density did not appear accurate at local scales. Distance to protected area was also not included as a predictor and its inclusion was assessed separately (Appendix 5A).

**Table 5.2 Presence records and sources.**

| Source | Region | Cheetah records | Wild dog records | Leopard records | Lion records | Data Collection Method(s) | Data Format |
|---|---|---|---|---|---|---|---|
| Alayne Cotterill | northern Kenya | | | | 31,627 | Telemetry | Presence grid |
| Alex Piel & Fiona Stewart | Mahale Mtns, Tanzania | | 2 | | 4 | Spoor | GPS coordinates |
| Alexandra Sutton | Vicinity of Masai Mara NR, Kenya | 3 | 2 | 2 | 7 | Sightings | GPS coordinates |
| Amy Dickman | Vicinity of Ruaha NP, Tanzania | 292 | 36 | | 2,211 | Sightings | GPS coordinates |
| Andimile Martin | western Tanzania | | 2 | 4 | 4 | Spoor and scat | GPS coordinates |
| Asgar Pathan | Tsavo NP, Kenya | 5 | | | | Sightings | GPS coordinates |
| Bernard Kissui | northern Tanzania | | | | 1,396 | Telemetry, sightings | GPS coordinates |
| Cherie Schroff | Tsavo NP, Kenya | 47 | | | | Sightings | GPS coordinates |
| Christian Kiffner | Katavi and Lake Manyara NP, Tanzania | | 1 | 5 | 85 | Playbacks, sightings, spoor | GPS coordinates |
| Dennis Ikanda | Vicinity of Selous GR, Tanzania | | | | 62 | Sightings, kills | GPS coordinates |
| Elena Chelysheva | Masai Mara NR, Kenya | 138 | | | | Sightings | KMZ file |
| Emmanuel Masenga | northern Tanzania | | 315 | | | Conflict records, sightings | GPS coordinates |
| Femke Broekhuis | Vicinity of Masai Mara NR, Kenya | 76 | | | | Sightings | GPS coordinates |
| Helen O'Neill | Vicinity of Serengeti NP, Tanzania | | 1 | | 264 | Sightings | GPS coordinates |
| Henry Brink | Selous GR, Tanzania | | 32 | | 309 | Sightings | GPS coordinates |
| Individual sightings | Tanzania, Kenya | 3 | 8 | 3 | 4 | Sightings | GPS coordinates |
| Ingela Jansson | Ngorongoro Conservation Area, Tanzania | 5 | 2 | 106 | 62 | Conflict records | GPS coordinates |
| IUCN records (Hans Bauer) | Tanzania, Kenya | | | | 44 | Sightings, aerial survey | GPS coordinates |
| Kenya Wildlife Authority | Kenya | 84 | | | | NA | GPS coordinates |
| Laly Lichtenfeld | Vicinity of Tarangire NP, Tanzania | 22 | 31 | 41 | 138 | Sightings, spoor, camera traps, conflict records | GPS coordinates |
| Literature | Various | | | 309 | | Sightings, camera traps | GPS coordinates |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Mary Wykstra | Kenya | 2,739 | | | | Telemetry, sightings | GPS coordinates |
| Maurus Msuha | Tarangire NP, Tanzania | | | | 18 | Camera traps | GPS coordinates |
| Michael Mbithi | Athi-Kapiti, Kenya | 81 | | | | Sightings | KMZ file |
| Mwangi Githiru | eastern Kenya | | 34 | | | Sightings, road kill | GPS coordinates |
| Northern Rangelands Trust | Kenya | 2,267 | 420 | | | Sightings | GPS coordinates |
| Paul Schuette | southern Kenya | 10 | 1 | 4 | 200 | Sightings, camera traps | GPS coordinates |
| Paulo Wilfred | Ugalla GR, Tanzania | | 2 | | | Sightings | GPS coordinates |
| Phillip Henschel | Tsavo NP, Kenya | 56 | 10 | | 280 | Transects | Presence grid |
| Roland Kays & Burce Patterson | Vicinity of Tsavo NP, Kenya | | | | 2,193 | Telemetry | GPS coordinates |
| Rosie Woodroffe | Kenya | | 16,383 | | | Sightings, telemetry, spoor, vocalization | GPS coordinates |
| Sarah Durant | Vicinity of Serengeti NP, Tanzania | 3,489 | | | 45 | Sightings | GPS coordinates |
| Shivani Bhalla | northern Kenya | | | | 106 | Sightings | GPS coordinates |
| Stephanie Dolrenry | southern Kenya, northern Tanzania | | | | 29,756 | Telemetry | Presence grid |
| Tanzania Mammal Atlas Project | Tanzania | 818 | 131 | 71 | 7 | Sightings, camera traps | GPS coordinates |
| The Uganda Wildlife Authority | Uganda | | | | 608 | Sightings | GPS coordinates |
| Totals | | 10,135 | 17,413 | 545 | 69,430 | | |

**Table 5.3 Species range in the study area (km$^2$).**

| | Historical range (% of study area) | Resident range (% of historical range) | Extirpated range (% of historical range) |
|---|---|---|---|
| Cheetah | 1,576,891 (93%) | 245,516 (16%) | 680,947 (43%) |
| Wild dog | 1,675,029 (98%) | 341,922 (20%) | 795,175 (47%) |
| Leopard | 1,702,850 (100%) | 1,042,312 (61%) | 369,434 (22%) |
| Lion | 1,701,479 (99.9%) | 504,523 (30%) | 835,869 (49%) |

### 5.2.2 Species distribution modeling

I used Random Forests (RF) to investigate carnivore distributions in East Africa. I implemented Breiman and Cutler's RF for classification and regression in the R statistical software (R Core Team 2015) using the randomForest package (Liaw and Wiener 2002). RF is an extension to the more familiar classification and regression trees. The algorithm constructs many trees (I used 5,000), and includes an additional parameter that limits the number of variables for use at each node in a tree. Predictions are averaged across trees. I began with 63 variables and selected the eight most important uncorrelated predictors using the mean decrease in accuracy values for the final model, the same as in Chapter 2.

I set up three modeling approaches for each species (Figure 5.1). The first two are detailed in Chapter 2 (Full and Extirpated). The Full option uses presence data contrasted with PsA data pulled from the entire study area and poses the modeling question of "where we saw the species" vs. "where the species could have occurred." The Extirpated option uses presence data contrasted with PsA from the species' extirpated range and poses the modeling question of "where we saw the species" vs. "where the species no longer occurs." The third approach is novel to this chapter and is the opposite of typical distribution models; instead of modeling probability of presence it modeled probability of absence. This approach, termed Extinct Suitability, drew random points from within extirpated areas as presence data, and drew random points from the historical range of the species, as PsA data. In effect, the PsA data from the Extirpated option was reversed to become "presence" data in this option. The modeling question here is "where the species used to occur" vs. "where we might have seen the species." Thus, it attempts to describe what are the predictors associated with absence, or extinction. Higher values therefore indicate a greater probability of absence, and conversely, a smaller probability of occurrence. Contrary to the other two approaches then, lower values in this approach indicate greater habitat suitability. I drew 5,000 PsA points randomly in R using the 'raster' (Hijmans 2016) and 'dismo' packages (Hijmans et al. 2016) for each modeling option.

After variable selection (see Chapter 2), I ran one final model for each of the four species and three approaches for a total of 12 models. As outlined in Chapter 2, I ran each model ten times with subsampling of PsA points. I averaged the ten runs to create an ensemble for each PsA selection strategy. Due to the limitations of the presence data, the spatial predictions should be interpreted as the relative likelihood of observation (Guillera-

Arroita et al. 2015). I reviewed model consistency by checking the Pearson correlation coefficient between all ten runs. I determined variable importance with two methods, first using the mean decrease in accuracy as calculated in the randomForest package. The second approach, called the PIMP test, is a permutation variable importance measure (Altmann et al. 2010) and implemented in the 'vita' package in R (Celik 2015). The PIMP algorithm permutes the response vector (whether or not the data point was presence or PsA), grows a new random forest for every permutation, and averages variable importance across all permutations. To compare, randomForest calculates variable importance after permuting the values of the variable itself using the out of bag observations. Finally, to rank the variables, I divided the value of each variable by the highest value in each modeling approach. I used these two metrics to test H3.

Using the rfUtilities package (Evans and Murphy 2015), I compared variable response curves between the four different species and three modeling approaches. I reviewed the shape of the response curves for each species against HPD and percent human land cover to test H1.

To determine if species had significantly different responses to human land cover (H2), response curves were created for all 10 PsA draws for each species. A single-parameter regression replicated each curve going from 1,0 to 0,1 with the parameter representing the concavity of the curve. The concavity parameter thus represents the change in the probability of observation over the change in percent human land cover. The parameter value was adjusted until the sum of the squared difference from the original response curve to the predicted curve was minimized. After identifying the concavity parameter for all forty curves, a one-way analysis of variance was used to determine if the parameter values were different between species. Following that, the Tukey HSD test was used to identify differences between species. These analyses were run separately for the Full and Extirpated options.

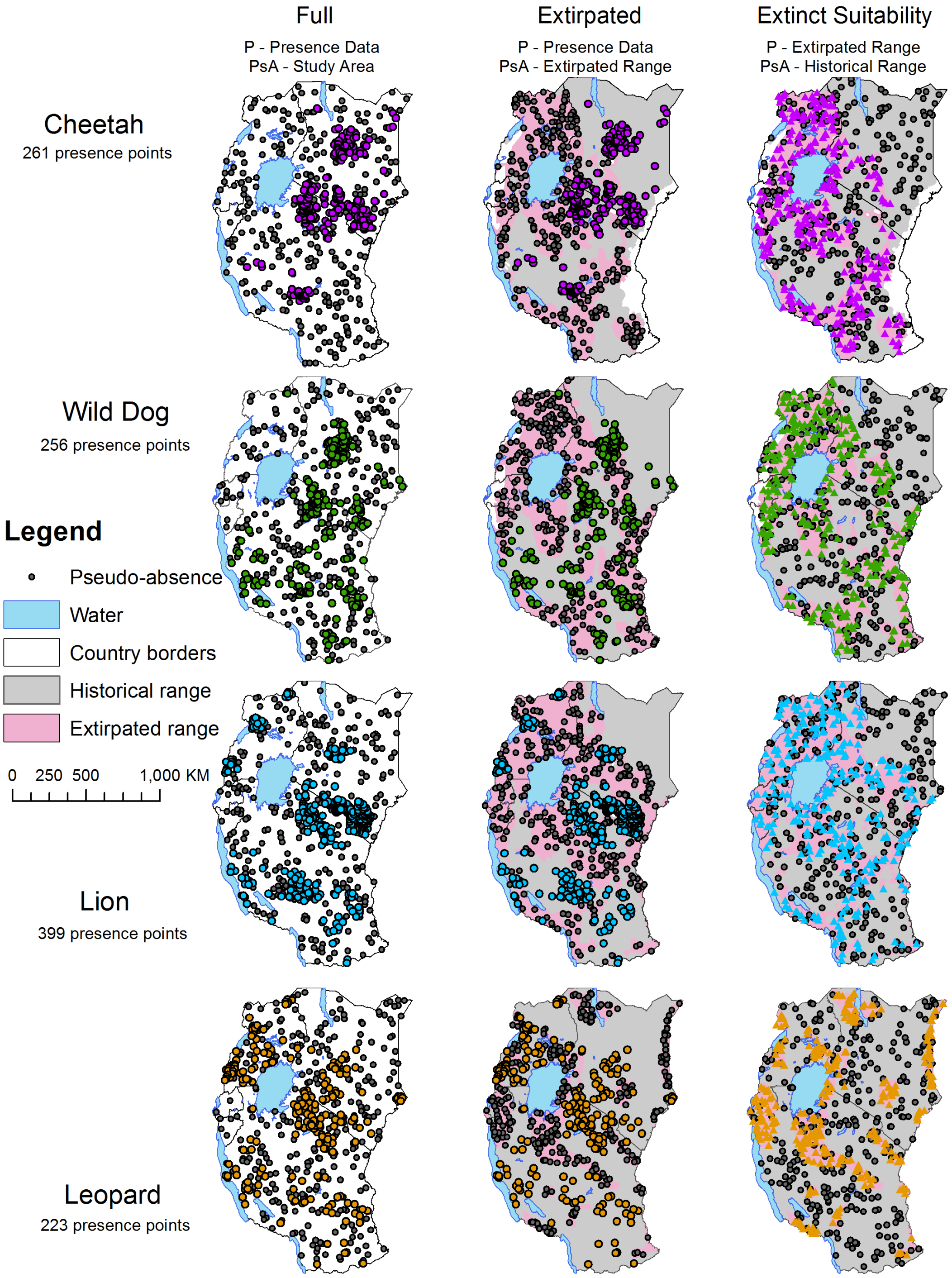


**Figure 5.1 Training data (colored dots) for each species and modeling approach. Only 300 random PsA points out of 5,000 are shown for illustration purposes. The colored triangles in the Extinct Suitability option are randomly selected points within extirpated range used as training data; they are not actual presence records. Note that the color used for each species will be carried through subsequent Figures where relevant.**

### 5.2.3 Model evaluation

Proper evaluation of a distribution model consists of multiple measures depending on the aim and application of the model (Austin 2007; Franklin 2009). A model should be ecologically realistic and robust, with 'good' predictive ability within some clearly defined criteria. To assess ecological realism, I reviewed the response curves and compared the spatial output to each species' current distribution. I noted model robustness by reviewing the consistency within and between model runs. Finally, in terms of predictive performance, I selected several threshold independent evaluation metrics. A threshold for spatial predictions is not recommended if not necessary for the application (Guillera-Arroita et al. 2015), and is noted as particularly difficult for generalist species (Barbosa et al. 2013). To verify model predictions, I used three metrics: the so-called 'out-of-bag' (OOB) error rates, Boyce Index (Boyce et al. 2002) calculated in the ecospat package in R (Broennimann et al. 2015), and the area under the receiver operating characteristic curve (AUC). I also calculated the predicted average probability of observation for each ensemble over the different expert-derived distribution zones: resident, extirpated, and outside the historical range. Finally, I calculated the correlation between the expert-derived resident and extirpated ranges of the four species as well as the correlation between all four species and three modeling approaches to compare the interspecific patterns of modeled outputs and reality.

## 5.3 Results

### 5.3.1 Model predictions

I created an ensemble Random Forest model for cheetah, wild dog, leopard and the lion in East Africa from 10 random PsA pulls, and for each of three model options, Full, Extirpated, and Extinct Suitability. The Full, Extirpated, and Extinct Suitability spatial predictions were compared with the species' range maps (Figures 5.2-5). The mean likelihood of observation and standard deviation for each of the distribution range categories were calculated for each species (Table 5.4). In each case, the mean likelihood in resident range was substantially higher, while the suitability in the extirpated region was lower between the Full and Extirpated options. The leopard had the highest mean likelihood averaged across the entire study area.

Within each selection strategy and across all four species, the internal correlation from the 10 different PsA draws had a similar trend, with the greatest correlation among the Extirpated model and the worst among the Extinct Suitability option (Table 5.5). The pattern in correlation of the spatial predictions largely matched that of the expert-derived range maps (Table 5.6). In both the range maps and the model output, cheetah and leopard had the lowest correlation with each other. On the other hand, the wild dog was strongly correlated with both the lion and cheetah in the range maps and model output.

The Full and Extirpated approaches for all four species had AUC scores near one and all Boyce Index values were above 0.86 (Table 5.7). For nearly all species and evaluation metrics, there was a clear progression that the Extirpated models had the best evaluation metrics while the Extinct Suitability had the worst metrics. The leopard had the lowest evaluation metrics overall although the best Extinct Suitability metrics.

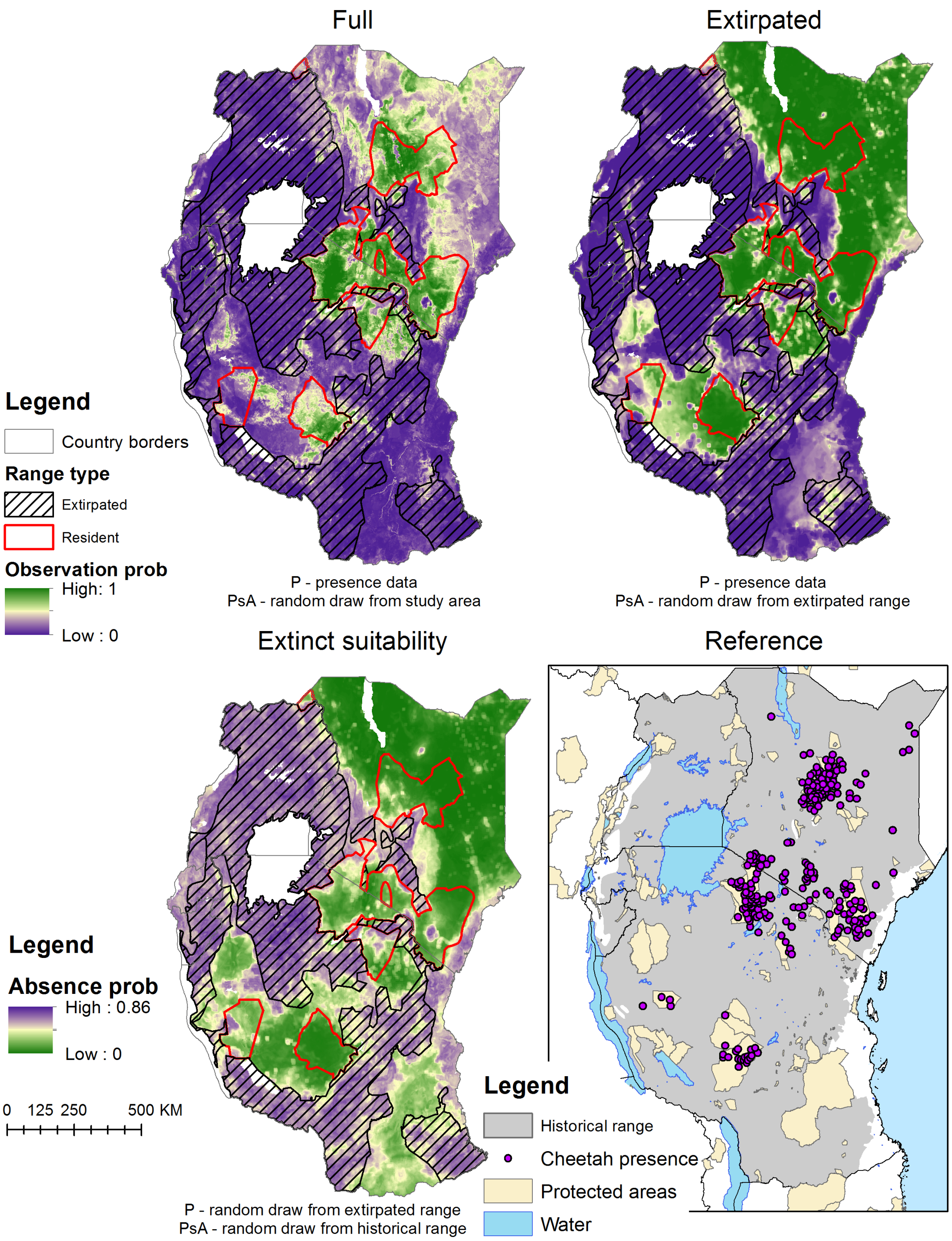
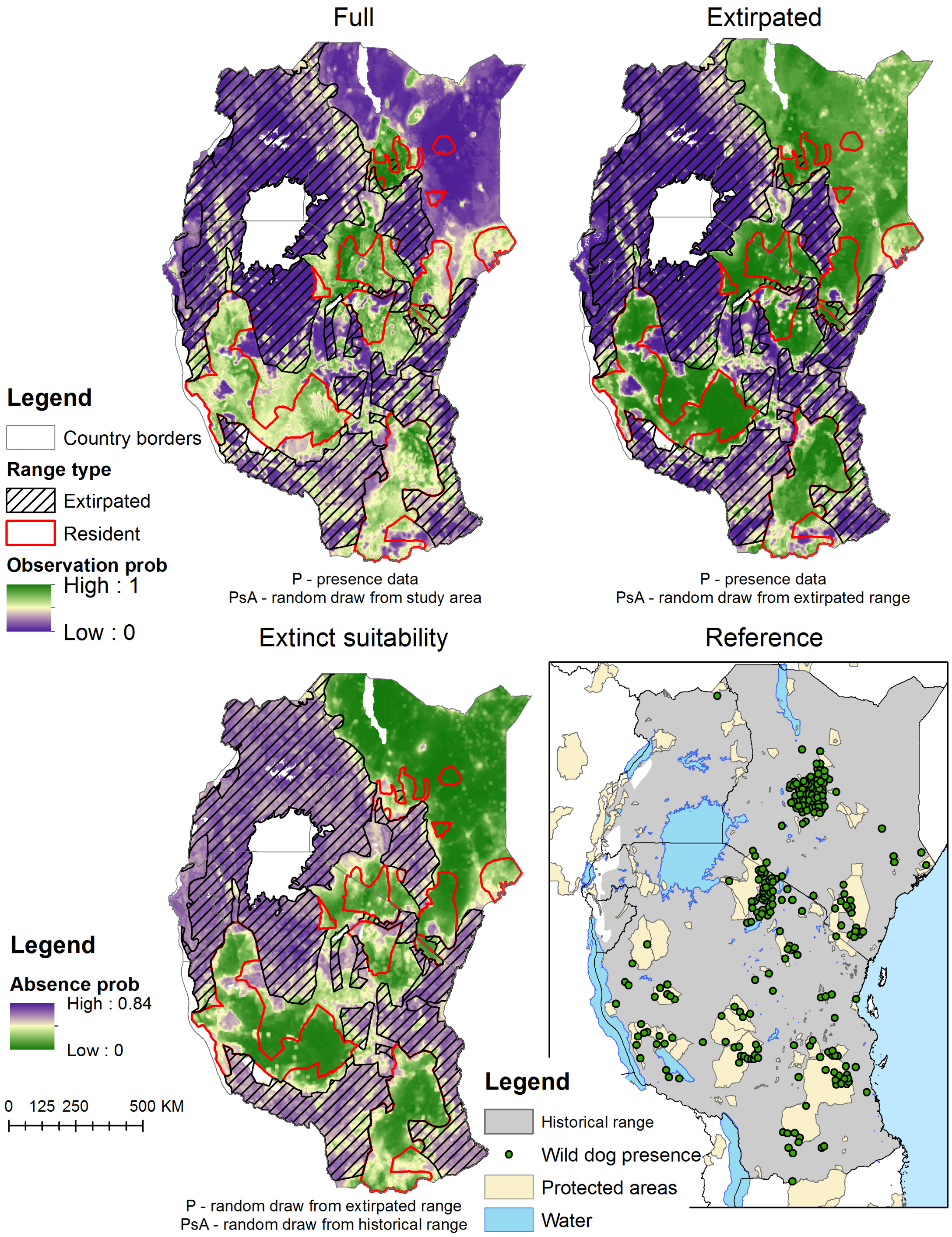


**Figure 5.2 Cheetah spatial predictions from each modeling approach. Observation probability represents likelihood of observation in the Full and Extirpated options. Extinct Suitability estimates the probability of absence; hence higher values indicate higher probability of absence. Thus, the color gradient is flipped as lower values here are similar to higher likelihood of observation in the Full and Extirpated options. The reference figure gives information on the historical range of the species, the training data, and location of protected areas.**



P - presence data
PsA - random draw from study area

P - presence data
PsA - random draw from extirpated range

P - random draw from extirpated range
PsA - random draw from historical range

**Figure 5.3 Wild dog spatial predictions from each modeling approach. Observation probability represents likelihood of observation in the Full and Extirpated options. Extinct Suitability estimates the probability of absence; hence higher values indicate higher probability of absence. Thus, the color gradient is flipped as lower values here are similar to higher likelihood of observation in the Full and Extirpated options. The reference figure gives information on the historical range of the species, the training data, and location of protected areas.**

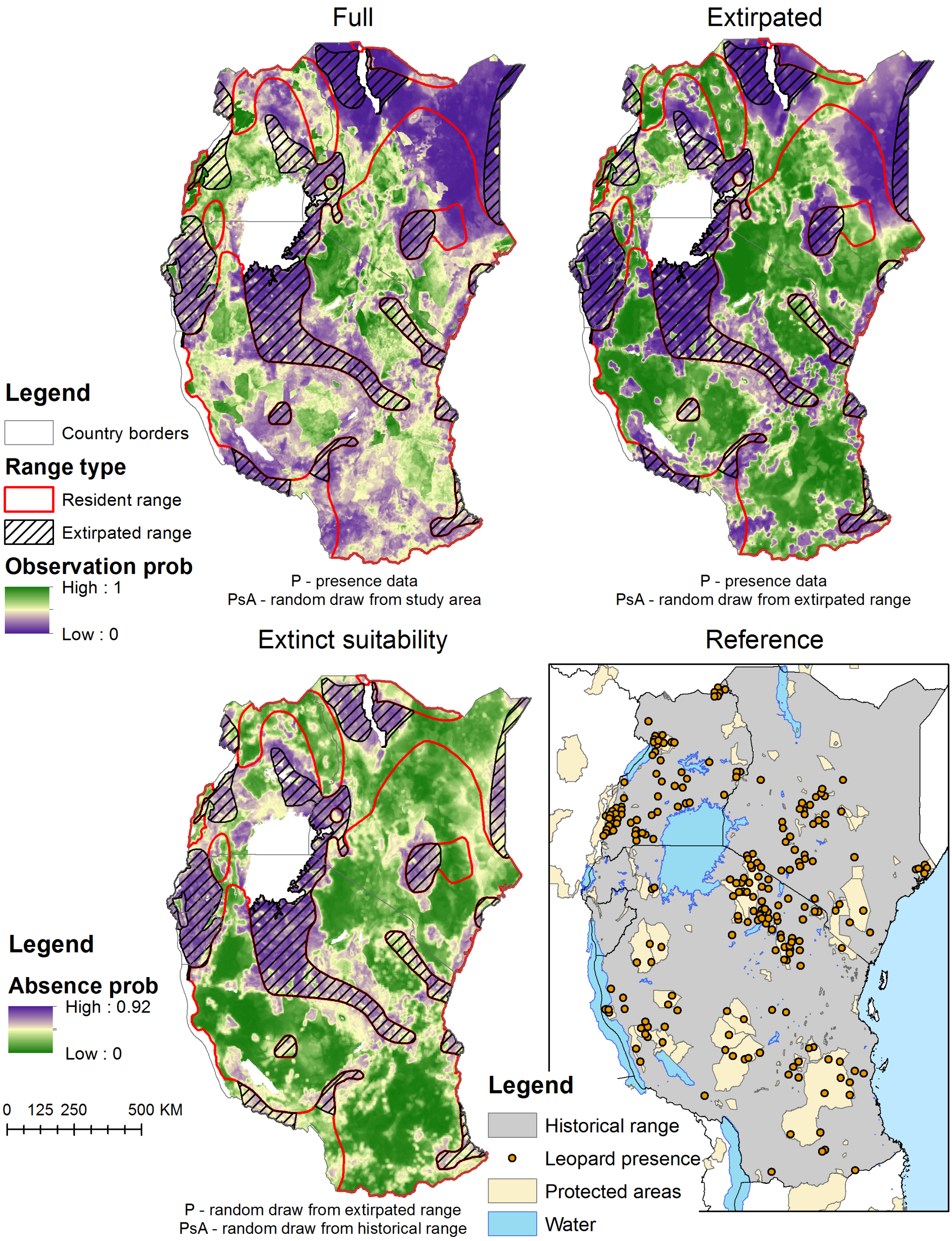


**Figure 5.4** **Leopard spatial predictions from each modeling approach. Observation probability represents likelihood of observation in the Full and Extirpated options. Extinct Suitability estimates the probability of absence; hence higher values indicate higher probability of absence. Thus, the color gradient is flipped as lower values here are similar to higher likelihood of observation in the Full and Extirpated options. The reference figure gives information on the historical range of the species, the training data, and location of protected areas.**

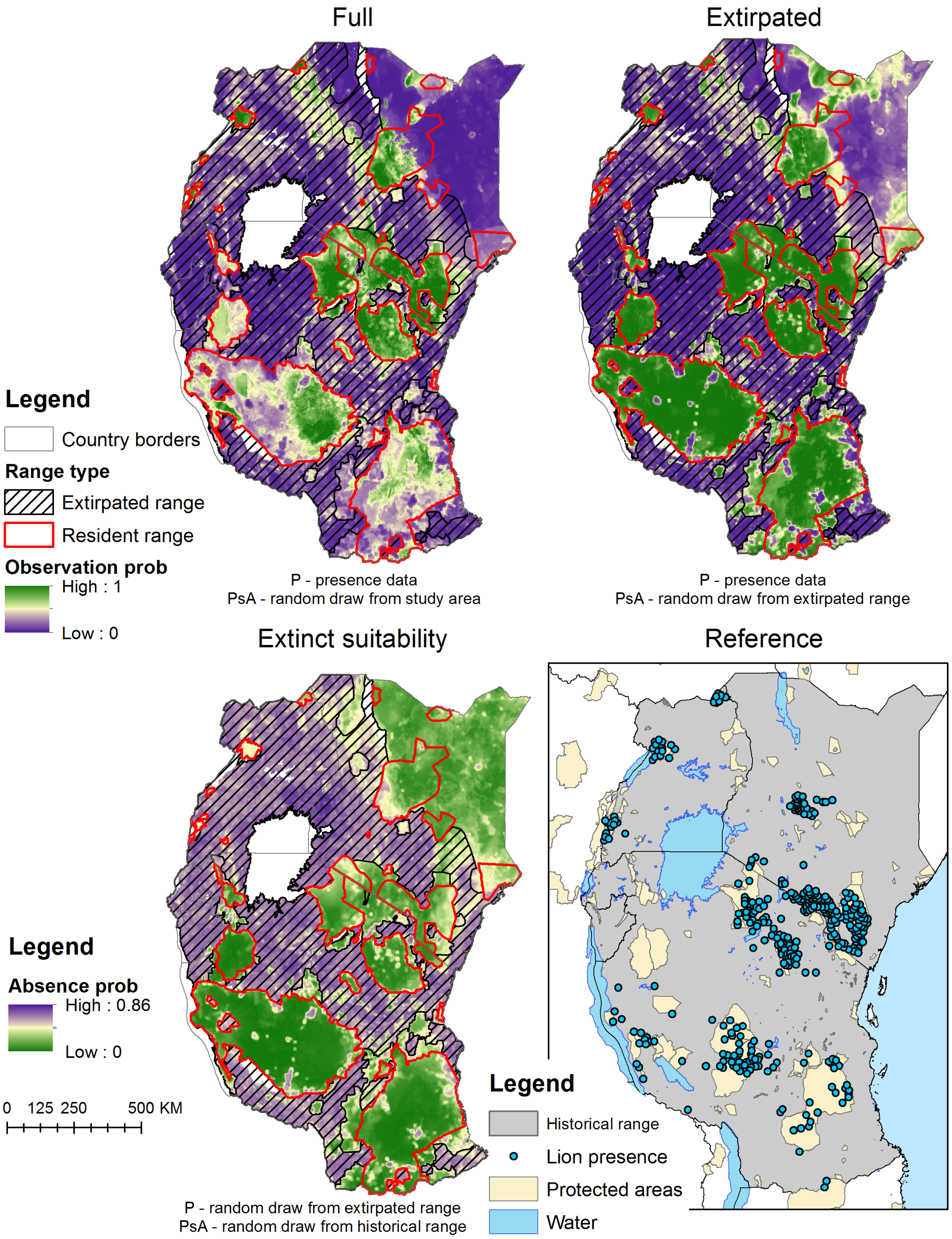


**Figure 5.5** Lion spatial predictions from each modeling approach. Observation probability represents likelihood of observation in the Full and Extirpated options. Extinct Suitability estimates the probability of absence; hence higher values indicate higher probability of absence. Thus, the color gradient is flipped as lower values here are similar to higher likelihood of observation in the Full and Extirpated options. The reference figure gives information on the historical range of the species, the training data, and location of protected areas.

**Table 5.4 Mean and standard deviation of the probability of observation for the ensemble prediction of each species, and modeling approach. The probabilities are averaged within each expert-derived range category, and across the full study area. Higher probabilities within Full and Extirpated options indicate higher likelihood of observation and are expected to be high in resident range and low in extirpated and outside historical range. Higher probabilities of the Extinct Suitability option indicate a higher likelihood of absence, and are expected to be lower in resident range and higher in extirpated and outside historical range. The best values for each modeling option are in bold.**

| | Modeling option | Resident | Extirpated | Outside historical range | Overall suitability |
|---|---|---|---|---|---|
| Cheetah | Full | **0.62 ± 0.26** | **0.11 ± 0.13** | **0.09 ± 0.11** | 0.24 ± 0.25 |
| | Extirpated | **0.81 ± 0.26** | **0.11 ± 0.17** | 0.12 ± 0.16 | 0.4 ± 0.39 |
| | Extinct Suitability | 0.21 ± 0.18 | 0.61 ± 0.11 | 0.54 ± 0.14 | NA |
| Dog | Full | 0.56 ± 0.20 | 0.22 ± 0.19 | 0.16 ± 0.12 | 0.31 ± 0.25 |
| | Extirpated | 0.78 ± 0.24 | 0.2 ± 0.23 | **0.1 ± 0.14** | 0.48 ± 0.36 |
| | Extinct Suitability | 0.22 ± 0.17 | 0.6 ± 0.11 | 0.61 ± 0.85 | NA |
| Leopard | Full | 0.42 ± 0.22 | 0.21 ± 0.16 | NA | 0.36 ± 0.22 |
| | Extirpated | 0.63 ± 0.31 | 0.19 ± 0.2 | NA | 0.5 ± 0.34 |
| | Extinct Suitability | 0.26 ± 0.19 | **0.64 ± 0.14** | NA | NA |
| Lion | Full | 0.49 ± 0.25 | 0.17 ± 0.16 | 0.47 ± 0.07 | 0.27 ± 0.25 |
| | Extirpated | 0.81 ± 0.26 | 0.15 ± 0.18 | 0.72 ± 0.14 | 0.4 ± 0.37 |
| | Extinct Suitability | **0.16 ± 0.16** | 0.6 ± 0.1 | 0.44 ± 0.06 | NA |

**Table 5.5 Correlations between spatial predictions for each species and modeling approach. The values on the 1:1 line are the range of internal correlations for all 10 PsA pulls.**

| | | **Cheetah** | | | **Dog** | | | **Leopard** | | | **Lion** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Full | Extirpated | Ext Suit | Full | Extirpated | Ext Suit | Full | Extirpated | Ext Suit | Full | Extirpated | Ext Suit |
| Cheetah | Full | 0.89-0.95 | | | | | | | | | | | |
| | Extirpated | 0.77 | 0.95-0.99 | | | | | | | | | | |
| | Ext Suit | -0.60 | -0.92 | 0.81-0.87 | | | | | | | | | |
| Dog | Full | 0.46 | 0.32 | -0.30 | 0.86-0.94 | | | | | | | | |
| | Extirpated | 0.67 | 0.85 | -0.88 | 0.59 | 0.96-0.98 | | | | | | | |
| | Ext Suit | -0.55 | -0.86 | 0.74 | -0.33 | -0.90 | 0.79-0.87 | | | | | | |
| Leopard | Full | 0.23 | -0.05 | 0.12 | 0.60 | 0.14 | 0.10 | 0.81-0.9 | | | | | |
| | Extirpated | 0.34 | 0.19 | -0.18 | 0.77 | 0.48 | -0.21 | 0.77 | 0.93-0.97 | | | | |
| | Ext Suit | -0.45 | -0.46 | 0.49 | -0.65 | -0.69 | 0.53 | -0.44 | -0.81 | 0.79-0.86 | | | |
| Lion | Full | 0.58 | 0.35 | -0.26 | 0.79 | 0.51 | -0.28 | 0.69 | 0.78 | -0.60 | 0.91-0.94 | | |
| | Extirpated | 0.51 | 0.47 | -0.49 | 0.80 | 0.73 | -0.55 | 0.52 | 0.79 | -0.77 | 0.80 | 0.96-0.98 | |
| | Ext Suit | -0.45 | -0.66 | 0.73 | -0.55 | -0.86 | 0.81 | -0.17 | -0.49 | 0.71 | -0.45 | -0.78 | 0.76-0.86 |

**Table 5.6 Correlation of species range maps below 1:1 line, and between modeled outputs above 1:1 line. Correlation between species range maps is a comparison of resident and extirpated range. Correlation between modeled outputs represents the correlation between each modeling option, averaged across all three options.**

| | Cheetah | Wild dog | Leopard | Lion |
|---|---|---|---|---|
| Cheetah | | 0.65 | 0.24 | 0.5 |
| Wild dog | 0.61 | | 0.43 | 0.65 |
| Leopard | 0.33 | 0.4 | | 0.6 |
| Lion | 0.49 | 0.63 | 0.48 | |

**Table 5.7 Evaluation metrics for each species and modeling approach. The best values for each modeling approach are in bold.**

| | | Boyce Index | AUC | Out of bag error |
|---|---|---|---|---|
| Cheetah | Full | **0.89 ± 0.01** | 0.98 ± 0.00 | **0.15 ± 0.02** |
| | Extirpated | 0.88 ± 0.01 | **0.99 ± 0.00** | **0.05 ± 0.01** |
| | Extinct Suitability | 0.86 ± 0.01 | 0.82 ± 0.01 | 0.32 ± 0.01 |
| Wild dog | Full | 0.87 ± 0.01 | **0.99 ± 0.00** | 0.23 ± 0.01 |
| | Extirpated | **0.89 ± 0.01** | 0.99 ± 0.00 | 0.11 ± 0.01 |
| | Extinct Suitability | 0.85 ± 0.01 | 0.81 ± 0.01 | 0.34 ± 0.03 |
| Lion | Full | 0.88 ± 0.00 | 0.99 ± 0.00 | 0.17 ± 0.01 |
| | Extirpated | **0.89 ± 0.01** | 0.99 ± 0.00 | 0.08 ± 0.01 |
| | Extinct Suitability | 0.86 ± 0.01 | 0.81 ± 0.01 | 0.34 ± 0.02 |
| Leopard | Full | 0.86 ± 0.01 | 0.98 ± 0.01 | 0.28 ± 0.02 |
| | Extirpated | 0.87 ± 0.01 | 0.99 ± 0.00 | 0.12 ± 0.01 |
| | Extinct Suitability | **0.86 ± 0.02** | **0.87 ± 0.01** | **0.28 ± 0.01** |

### 5.3.2 Variable importance and response curves

The eight most important and uncorrelated variables were selected for the final model for each species and modeling option (Table 5.8). Variables averaged at the largest scale (12 km) were more frequently but not universally selected over predictors averaged at smaller scales. Three predictors, mean annual temperature, human population density, and percent human land cover, were selected in every method. The two different variable importance methods, mean decrease in accuracy and PIMP, produced different ranks of important variables, although the top variable frequently stayed the same. The PIMP test suggested in the Extinct Suitability option of three species, that a handful of variables had negative contributions to model predictions i.e., their removal would have produced a better model.

Response curves varied between species for both Full and Extirpated options (Figures 5.6 & 5.7; Supplemental Figure 5.1). Some variables (e.g. soil CEC and dry season NDVI) had relatively consistent responses across all species. In two variables (soil carbon and human impact), all species but the leopard had relatively consistent responses. A few other variables (e.g. rainfall seasonality) showed more widely varying response between all species. Supporting Predictions 1a and 1b, the response curves for both HPD and percent human land cover were strongly negative. The response curves for human impact were negative as well except for the leopard, which showed increasing suitability as human impact increased.

**Table 5.8 Variable importance of the final models relativized by the most important variable, using the mean decrease in accuracy (Rank) and PIMP test (P rank), for each species and modeling option.** See Table 4.3 for variable acronyms.

**Cheetah**

| Full | Rank | P Rank | Extirpated | Rank | P Rank | Extinct Suitability | Rank | P Rank |
|---|---|---|---|---|---|---|---|---|
| MAT (6km) | 1 | 1 | Precip (12km) | 1 | 0.88 | Precip (12km) | 1 | 0.90 |
| % Human LC (12km) | 0.91 | 0.84 | % Human LC (12km) | 0.91 | 1 | % Human LC (12km) | 0.78 | 1 |
| NDVI $_{dry}$ (3km) | 0.90 | 0.96 | NDVI StDv $_{dry}$ (12km) | 0.56 | 0.30 | HPD (12km) | 0.65 | 0.53 |
| NDVI $_{wet}$ (3km) | 0.72 | 0.81 | MAT (12km) | 0.48 | 0.16 | Human Impact | 0.58 | 0.37 |
| HPD (6km) | 0.57 | 0.53 | Precip $_{dry\ season}$ (12km) | 0.46 | 0.19 | Soil Carbon (12km) | 0.53 | 0.37 |
| Soil Carbon (12km) | 0.52 | 0.38 | HPD (12km) | 0.46 | 0.44 | Precip $_{seasonality}$ (12km) | 0.43 | -0.09 |
| Soil CEC (12km) | 0.49 | 0.30 | Soil Carbon (12km) | 0.42 | 0.14 | NDVI StDv $_{dry}$ (12km) | 0.41 | 0.15 |
| Precip $_{dry\ season}$ (12km) | 0.48 | 0.16 | Human Impact | 0.27 | 0.12 | MAT (3km) | 0.41 | 0.14 |

**Wild dog**

| Full | Rank | P Rank | Extirpated | Rank | P Rank | Extinct Suitability | Rank | P Rank |
|---|---|---|---|---|---|---|---|---|
| % Human LC (12km) | 1 | 1 | % Human LC (12km) | 1 | 1 | HPD (12km) | 1 | 0.91 |
| MAT (12km) | 0.97 | 0.59 | NDVI $_{dry}$ (12km) | 0.73 | 0.41 | Precip (12km) | 0.92 | 0.71 |
| Precip (12km) | 0.86 | 0.76 | MAT (12km) | 0.71 | 0.24 | % Human LC (12km) | 0.88 | 1 |
| Precip $_{seasonality}$ (6km) | 0.64 | 0.23 | Soil CEC (12km) | 0.64 | 0.32 | NDVI StDv $_{dry}$ (12km) | 0.75 | 0.55 |
| HPD (12km) | 0.63 | 0.46 | HPD (12km) | 0.61 | 0.65 | Precip $_{seasonality}$ (12km) | 0.68 | 0.44 |
| NDVI StDv $_{dry}$ (12km) | 0.56 | 0.31 | Precip $_{seasonality}$ (12km) | 0.53 | 0.19 | Human Impact | 0.63 | 0.34 |
| TRI (500m) | 0.41 | 0.10 | Human Impact | 0.51 | 0.49 | MAT (12km) | 0.46 | 0.07 |
| Soil Carbon (12km) | 0.37 | 0.32 | NDVI $_{wet}$ (500m) | 0.34 | 0.22 | Soil Carbon (12km) | 0.37 | 0.15 |

**Leopard**

| Full | Rank | P Rank | Extirpated | Rank | P Rank | Extinct Suitability | Rank | P Rank |
|---|---|---|---|---|---|---|---|---|
| HPD (6km) | 1 | 1 | % Human LC (12km) | 1 | 1 | % Human LC (12km) | 1 | 1 |
| % Human LC (12km) | 0.87 | 0.76 | Precip (6km) | 0.58 | 0.48 | Precip (12km) | 0.73 | -0.01 |
| Precip $_{seasonality}$ (12km) | 0.84 | 0.39 | Soil CEC (12km) | 0.51 | 0.29 | HPD (12km) | 0.70 | 0.62 |
| Precip (6km) | 0.71 | 0.61 | MAT (12km) | 0.51 | 0.31 | MAT (3km) | 0.63 | -0.04 |
| MAT (12km) | 0.65 | 0.43 | Precip $_{seasonality}$ (6km) | 0.46 | 0.18 | Precip $_{seasonality}$ (12km) | 0.57 | 0.13 |
| Soil Carbon (6km) | 0.42 | 0.53 | HPD (12km) | 0.45 | 0.38 | Human Impact | 0.52 | 0.43 |
| NDVI StDv $_{wet}$ (12km) | 0.42 | 0.25 | Soil Carbon (12km) | 0.26 | 0.15 | Soil CEC (12km) | 0.47 | -0.02 |
| Human Impact | 0.22 | 0.14 | Human Impact | 0.25 | 0.08 | NDVI StDv $_{wet}$ (12km) | 0.35 | 0.27 |

**Lion**

| Full | Rank | P Rank | Extirpated | Rank | P Rank | Extinct Suitability | Rank | P Rank |
|---|---|---|---|---|---|---|---|---|
| Precip (3km) | 1 | 0.86 | % Human LC (12km) | 1 | 1 | % Human LC (12km) | 1 | 1 |
| % Human LC (12km) | 0.86 | 1 | MAT (12km) | 0.70 | 0.24 | Precip $_{dry\ season}$ (12km) | 0.73 | 0.47 |
| MAT (6km) | 0.85 | 0.36 | HPD (12km) | 0.55 | 0.54 | HPD (12km) | 0.73 | 0.64 |
| HPD (12km) | 0.68 | 0.79 | NDVI $_{dry}$ (500m) | 0.49 | 0.19 | NDVI StDv $_{dry}$ (12km) | 0.51 | 0.22 |
| Precip $_{seasonality}$ (6km) | 0.50 | 0.09 | Soil Carbon (12km) | 0.44 | 0.20 | Human Impact | 0.47 | 0.31 |
| Soil CEC (12km) | 0.48 | 0.24 | NDVI $_{wet}$ (12km) | 0.43 | 0.21 | Precip $_{seasonality}$ (3km) | 0.36 | 0.06 |
| NDVI StDv $_{dry}$ (3km) | 0.45 | 0.17 | Precip $_{seasonality}$ (6km) | 0.41 | 0.10 | MAT (12km) | 0.22 | -0.01 |
| Soil Carbon (12km) | 0.45 | 0.33 | Human Impact | 0.33 | 0.19 | NDVI StDv $_{wet}$ (12km) | 0.20 | -0.02 |

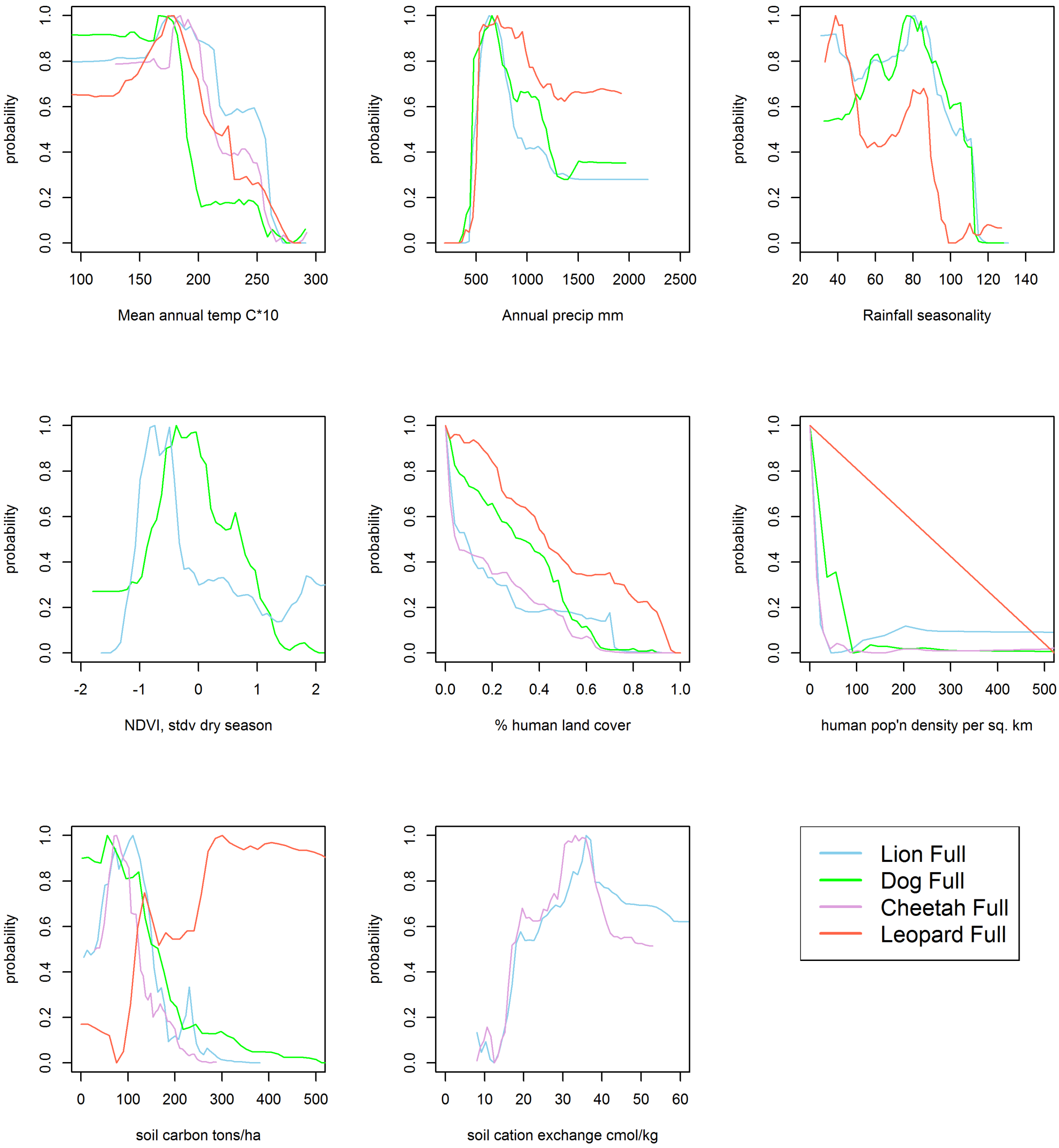


**Figure 5.6** Response curves for all species and the most frequently used variables in the Full modeling approach.

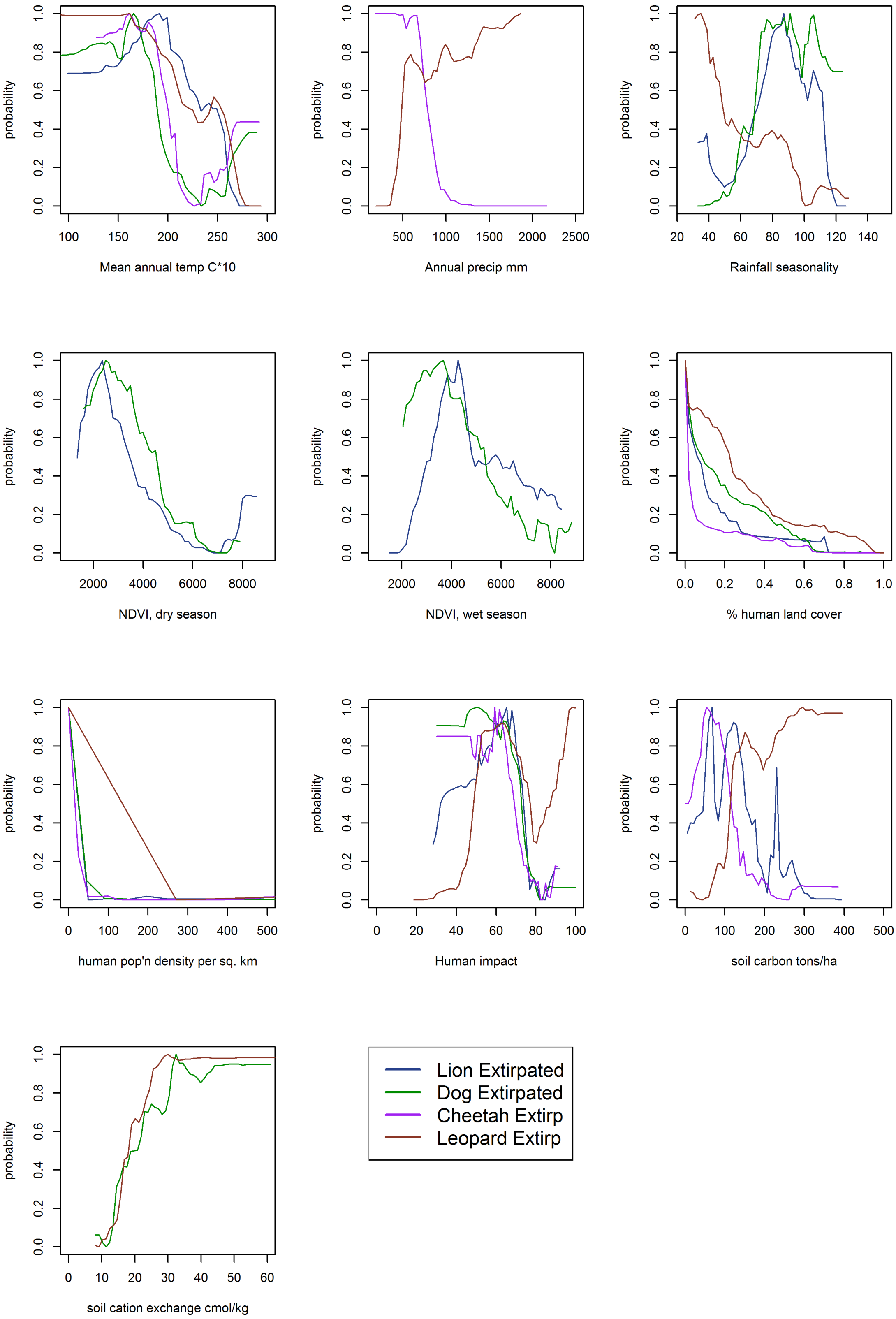


**Figure 5.7** **Response curves for all species and the most frequently used variables in the Extirpated modeling approach.**

The global concavity parameters representing the response curves for percent human land cover were significantly different across species in both modeling approaches (Full - F = 80.9, $p<0.0001$; and Extirpated - F = 92.4, $p<0.0001$) (Figure 5.8). When comparing individual species, the cheetah and lion had the steepest slopes (i.e., smallest values) and were not significantly different from each other. The wild dog was shallower than the lion and cheetah but steeper than the leopard. Both the wild dog and leopard were significantly different from every other species (all comparisons $p < 0.0001$ except for the Full option wild dog-cheetah $p = 0.0001$ and wild dog-lion $p = 0.002$). These results held in both the Full and Extirpated options. These results provide partial support for Prediction 2.

In support of Prediction 3, percent human land cover was the most influential variable in either one or both variable importance metrics for 10 of the 12 models. In the two models where land cover was not the most influential variable, the Full options for the cheetah and the leopard, percent human land cover was the 2nd most important variable. Percent human land cover was frequently about two times more important than HPD.

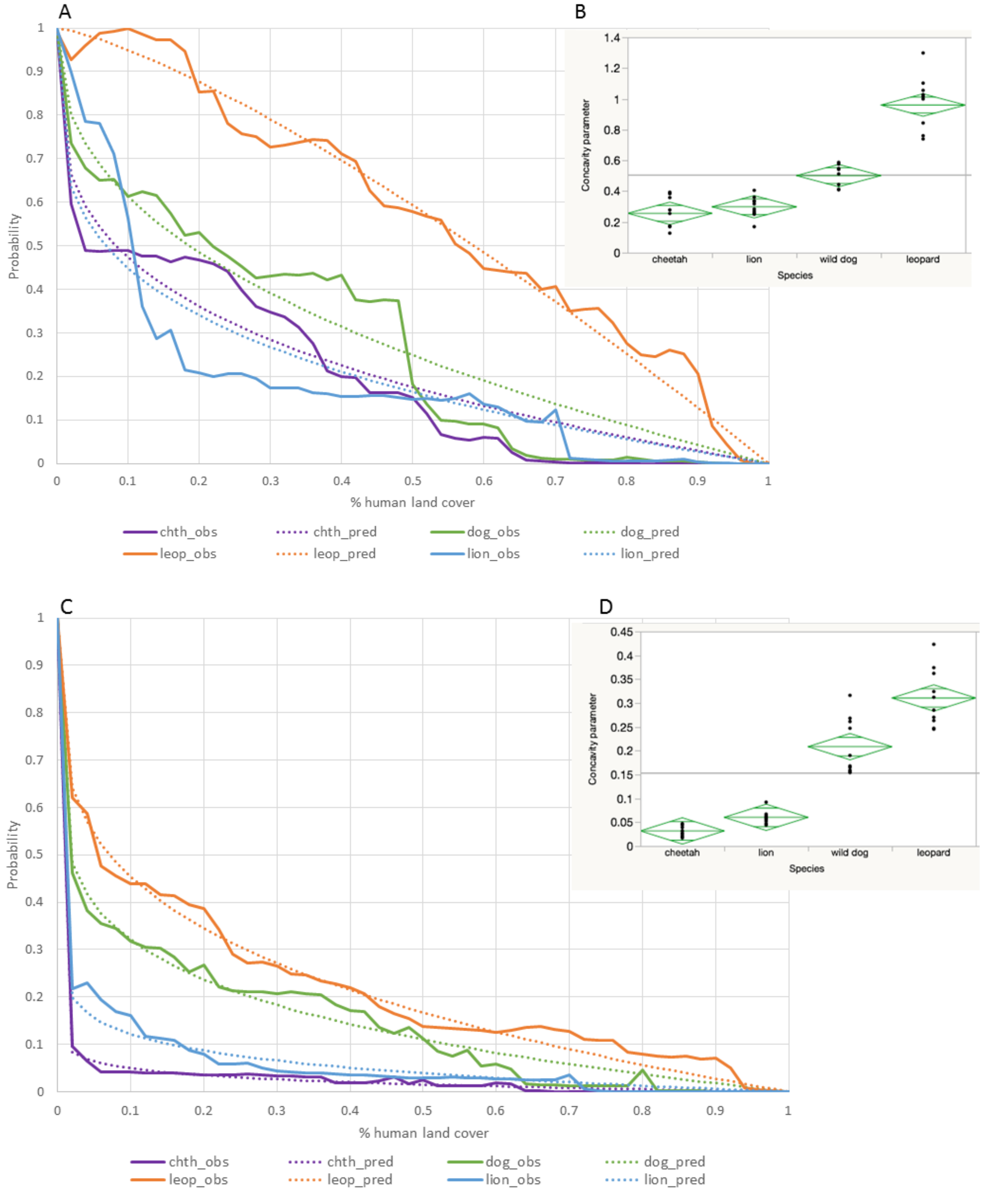


**Figure 5.8** **Comparison of response curves across species in the Full and Extirpated options. An example response curve and predicted curve for each species for the variable percent human land cover (A). Boxplot of the concavity parameters from the 10 PsA pulls for each species (B). C and D are the same but for the Extirpated option. A smaller concavity parameter represents a steeper slope, and steeper decline in probability of observation.**

## 5.4 Discussion

In this analysis, several ensemble RF models were used to identify correlates of cheetah, wild dog, leopard, and lion distribution in East Africa. Results supported the concept developed in Chapter 4 that posing various modeling questions via different combinations of presence and PsA data could strengthen ecological conclusions. This approach enabled comparison of the importance of HPD and human land cover, leading to the conclusion that human land cover is more important in restricting large carnivore presence than human populations per se. In addition, species have differential susceptibility to increasing levels of human land cover, with cheetah and lion most sensitive. The prominence of human impacts in restricting large carnivore distribution in East Africa is demonstrated, and results highlight the importance of the protected area network as habitat (Appendix 5A). Results also suggest the more remote parts of Kenya and Tanzania as potentially vital refugial habitat for all carnivores. A more detailed comparison of modeling output and species range maps can be found in Appendix 5B.

### 5.4.1 Modeling approach: evaluation and utility

I designed three modeling approaches to examine the distribution of four large carnivores in East Africa. Two options used presence data gathered from a variety of sources contrasted with PsA from the entire study area (Full) and from the extirpated range (Extirpated) to compare the importance of different predictors in carnivore range decline. The third approach, Extinct Suitability, used random "presence" points from the extirpated range, and contrasted them with PsA from the historical range of the species. In most circumstances, these multiple approaches are confirmatory and built support for conclusions to be discussed later.

A strength of these multiple modeling approaches is that response curves, among other metrics, can be readily compared. In several cases, a comparison of the response curves between the Full and Extirpated models highlight how a species has been lost from more favorable habitat (Figure 5.9). If the response curve from the Full option is interpreted as preferred, then a strong deviation from that in the Extirpated option suggests a change in the relative probability of observation between former and current habitat. For instance, the response curves of the cheetah to mean annual temperature suggest a depressed probability of observation in the Extirpated model between ~19 – 25 °C in the Extirpated option as compared to the Full. In contrast, there is elevated probability of observation above 25° C. Geographically this represents primarily the arid NE of Kenya. Hence, relatively more cheetah habitat in temperate climates has been lost in contrast to habitat at elevated temperatures. The wild dog exhibits a similar but less dramatic trend as the cheetah to temperature. Another example is with the leopard and annual precipitation. The response curves suggest a depressed probability in the Extirpated model between annual rainfall totals of ~500 – 1000 mm and an enhanced probability at higher precipitation levels. Leopards have been lost from much of the drier regions of their historic range (e.g. north Africa, the Middle East etc.) (Jacobson et al. 2016) and may be more persistent in higher productivity ecosystems, due to higher fecundity or better refuge in denser habitats. A third example is the response of the wild dog to rainfall seasonality. This variable is a measure of the amount of variation in rainfall between the wettest and driest month. Lower values indicate less variation in precipitation throughout the year, which likely contributes to year-round agriculture. The Full option of the wild dog shows dramatically higher probability than the Extirpated option at lower values but this is reversed at higher values. Thus, the wild dog has been lost from habitats with greater year-round consistency in rainfall but remained present in seasonally dry habitats like NE Kenya or SE Tanzania.

The models were evaluated in a number of ways including if they were ecologically realistic, robust, and had good spatial predictions. The models performed well across all these measures. The models largely matched ecological reality in terms of response curves (see Species Comparisons below for more information) and spatial predictions (below). While some of the response curves were unimodal as expected, some were truncated which could be due to few data from the extremes. Some variables may have been overfit as the response curves portrayed several breaks or multiple peaks in suitability. The models were relatively robust as compared across the 10 iterations, with all species achieving above 92% correlation in their Extirpated options. But the predictions do vary in spatial consistency based on the model approach; the Extinct Suitability option was least robust as the internal correlation varied from 0.76 to 0.87 across all species. The spatial predictions of the models were evaluated in several different ways, including comparison to each other, to expert-derived range maps and with traditional evaluation metrics like AUC.

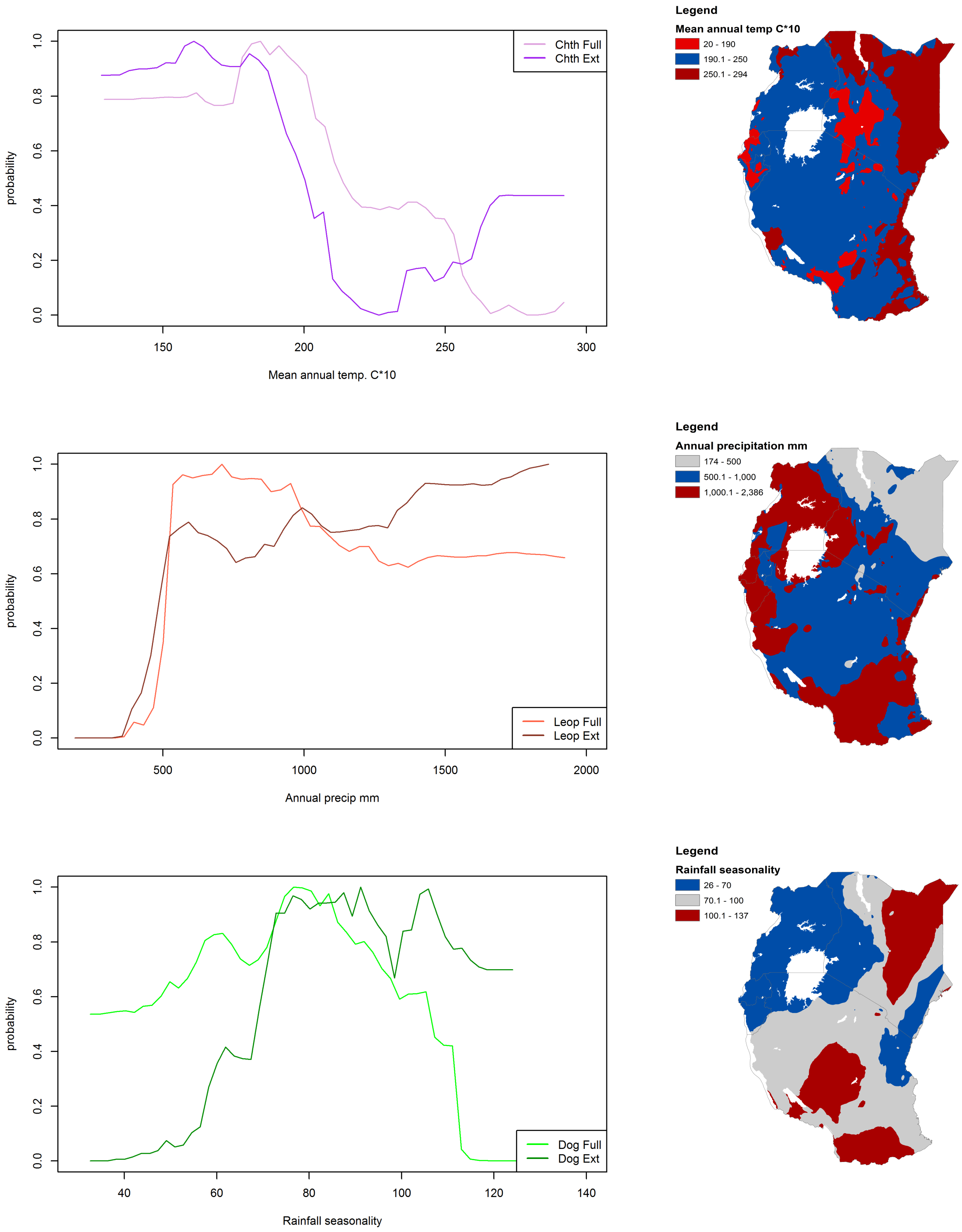


**Figure 5.9** In the first column, a comparison of response curves between Full and Extirpated options for select species and variables. In the second column, a map of the study area depicting the coverage of the key values. Variable ranges are consistently colored. Gray represents the range of values where the response curve is similar between the Full and Extirpated options. Blue represents the range of values where the response curve is lower in the Extirpated option than the Full; hence a decreased suitability. Red represents the range of values where the response curve is higher in the Extirpated option than the Full; hence an increased suitability.

Spatial predictions for each species largely matched those from expert-derived range maps and from biological knowledge. Thus, they provide an objective means of assessing the boundaries for these maps, and for filling in missing information, such as areas of range mapped as unknown or possible. The leopard, which historically ranged across the entire study area, had the highest mean likelihood of observation while the cheetah, with the least amount of historic range, had the lowest value. Matching expert knowledge, cheetah suitability is low in SE Tanzania in both Full and Extirpated options, despite the lack of human pressure in the Selous Game Reserve (IUCN/SSC 2007). That region only became more suitable in the Extinct Suitability option. In addition, cheetah was more broadly suitable in the arid NE of Kenya than any other species. However, there were omissions as well. The Selous Game Reserve which may be the largest single lion population in the world (Riggio et al. 2013), was not highly suitable in the Full option. Areas of wild dog resident range in NE Kenya (i.e. the Isiolo and Kora-Niktui range polygons) (IUCN/SSC 2007) were not very suitable in either Full or Extirpated options.

Yet, there were strong similarities between the range maps and spatial predictions. For all species and in many regions, the outline of the expert-derived resident range was mirrored by hard edges in the spatial predictions of the Extirpated option (e.g. along the western edge of the Serengeti ecosystem or eastern edge of the Ruaha ecosystem). Across the modeling approaches, the cheetah had the closest match to expert-derived distribution within resident, extirpated and outside historical range while the leopard had the worst. The mean likelihood of observation was higher in resident range and lower in extirpated range across all species for the Extirpated option compared with the Full option. A possible reason for this is because of sampling bias and some areas of resident range had few or no presence points (i.e., western Tanzania for the cheetah or the lion in Lamu and Garissa counties, eastern Kenya). However, these areas tended to have higher probabilities in the Extirpated option than the Full option, suggesting that they more closely resembled habitat than former habitat. Hence, the Extirpated approach may assist in identifying potential habitat even when sampling is biased across species' range.

In addition, the pattern of correlation among the species range maps largely matched the correlation in the spatial predictions. For instance, the correlation between the wild dog and lion or cheetah were high in both the range maps and modeled outputs; similarly the low correlation between leopard and cheetah. This suggests that the models were capturing the idiosyncrasies of the species' distributions.

Both the Full and Extirpated approaches had very high evaluation metrics with AUC values of >0.99 and OOB error rates <0.1 in several cases. The Extinct Suitability option had substantially worse evaluation metrics except in the case of the leopard, where the Boyce Index and OOB error were similar to the Full option. The leopard had the worst evaluation metrics overall of the Full and Extirpated options but the best Extinct Suitability metrics. Interestingly, the cheetah had many of the best evaluation statistics, as well as the spatial output that most closely resembled the expert-derived distribution.

### 5.4.2 Species comparisons

Species' response curves matched some of the known biological characteristics of the different species. For instance, the wild dog was the only species where topographic roughness was included in a model. Previous research has suggested that wild dogs use areas of greater topographical complexity, potentially as refuge from competitors and/or during denning (Woodroffe 2011; Jackson et al. 2014). When comparing the species responses to mean annual temperature, wild dogs were least tolerant to increasing temperatures; they showed the earliest and steepest decline as temperatures increased. Wild dogs are sensitive to overheating and higher daily temperatures are known to restrict ranging behavior and the duration of hunts (Woodroffe 2011; Hubel et al. 2016). Wild dog and cheetah showed more tolerance of low dry-season NDVI levels than lions, implying a capacity for survival in less productive environments. Although lions can live in some extreme environments (e.g., the Kalahari and Namib deserts (Nowell and Jackson 1996)), lion density is strongly related to herbivore biomass and rainfall (van Orsdol, Hanby, and Bygott 1985; Celesia et al. 2009). Although more true of the cheetah than wild dog, both species live at very low densities, and can subsist on small prey or at low prey densities (Woodroffe et al. 2007c; Belbachir et al. 2015; Sillero-Zubiri et al. 2015). Similarly, in terms of annual precipitation, cheetahs were most tolerant of low rainfall levels and least tolerant of higher precipitation levels. This matches their capacity to live in the Sahara desert at low densities (Belbachir et al. 2015; Sillero-Zubiri et al. 2015) and their historical absence from tropical forests (Wrogemann 1975). In addition, leopards, which are the only species in this group that live throughout African tropical forests, showed unique increasing suitability as annual precipitation increased. Finally, leopards had a dramatically different response to soil carbon (linked with plant biomass); leopard suitability increased while all other species declined as soil carbon increased. This may reflect leopard tolerance for African tropical forests.

Another interesting comparison among species' response curves was how leopard responded to human impact. The probability of observation for all species but the leopard decreased towards zero as human pressure increased, whereas leopard probability increased towards one. This suggests that suitability for leopards actually increased the closer they were to large human settlements and roads. This is unlikely to be true, as leopards are present in many remote, protected areas throughout their range (Jacobson et al. 2016) and this result may be an artifact of sampling design. In addition, this is contradictory to the leopard's response to percent human land cover or HPD, which decreases to zero as these values increases. But, this may suggest that leopards inhabit undeveloped or un-peopled areas that are close to cities and roads. In addition, leopards are known to live in highly human impacted regions throughout their range (Athreya et al. 2013; Jacobson et al. 2016). Finally, this matches ecological theory; generalist predators may become more abundant in fragmented habitats with greater access to human-derived food sources (Oehler and Litvaitis 1996), and mesopredators may be 'released' from competitive effects as apex predators (i.e. lions) disappear in more urban settings (Crooks and Soulé 1999; Caro and Stoner 2003; Vanak et al. 2013).

### 5.4.3 Ecological hypotheses

Results indicated that all species' suitability declined strongly as human populations and human land cover increased, supporting both Predictions 1a and 1b. Species showed different sensitivities to these predictors, yet all declined to zero suitability at some threshold level (although some rebounded to very low probabilities at higher HPD). Yet, large carnivores are making a comeback in several densely populated and fragmented landscapes, such as Europe (Chapron et al. 2014). More research is needed to determine what human actions are in greatest conflict with carnivore persistence (Ripple et al. 2014). Therefore, while increasing human populations and conversion of natural habitats threaten large carnivores, the situation is complex and impacts vary by species and local human actions (Burton et al. 2011).

Cheetahs, with the largest home range, were predicted to be the most vulnerable to increasing levels of human land cover, while wild dog, lion and leopard have decreasing home range sizes in that order (Prediction 2). Results confirmed the leopard was the least vulnerable, and cheetah the most vulnerable. However, there was no significant difference between lion and cheetah, and wild dog was less vulnerable than either of those two species. It's possible another biological factor (or a combination such as fecundity and habitat breadth) missing from this analysis modifies the relationship. Burton et al. (2011) examined the pattern of carnivore loss in Ghana's Mole NP, and found that susceptibility to extinction is highly complex and determined by interactions between intrinsic traits and extrinsic pressures.

The third prediction, that habitat loss is a greater factor impacting carnivore distributions than human populations, was also supported. Percent human land cover was the most important variable for nearly all species, modeling options, and both variable importance metrics. Percent human land cover was not the top predictor in either of the two metrics in only two instances (Cheetah, Full and Leopard, Full), but it was the 2nd most important predictor. In the Extirpated and Extinct Suitability options, it was the top predictor in both metrics in five of eight instances, and was frequently roughly two times more important than HPD. These results suggest that changes in land cover are more strongly associated with range decline for large carnivores than the expansion of human populations per se.

Human land cover is linked with lower prey densities, loss of vegetation for daytime cover/refuge, and overall loss of biodiversity (Fahrig 2003; Kiffner et al. 2014; Oriol-Cotterill et al. 2015b; Watson et al. 2015). These impacts may decrease hunting opportunities or hunting success rates, increase human-wildlife conflict events, and may increase the chance of detection from humans.

These results concur with Schuette, Creel, and Christianson (2013), who found that lions in a pastoral and unprotected landscape in southern Kenya achieved nearly the same density as lions in Serengeti NP. They believed the continued existence of the lions was aided by daytime refuges in dense vegetation and by spatiotemporal variation in human land uses. There was little cultivation in the area. They argued that maintaining pastoral traditions allowed coexistence while greater sedentarization and land conversion to agriculture could decrease suitability for lions and undermine coexistence.

On the contrary, HPD may not be useful as a predictor of carnivore presence, as species exhibit a wide range of tolerance. Oriol-Cotterill et al. (2015a) argue that in some cases, the mere presence of people may limit lion's use of the environment, while Dolrenry et al. (2014) found that the addition of HPD had no effect on an incidence function model exploring the probability of lion occurrence in East African habitat patches. Across leopard range, mean HPD in areas where leopards were extinct varied from 58 to 1,076 people and from 6 to 332 people/km$^2$ in extant populations (Jacobson et al. 2016). Henschel (2008) found no leopards in some study areas of Gabon where human density was only 1-2 people/km$^2$ and (Toni and Lodé 2013) found few leopards in areas of Cameroon with HPD >10. On the other hand, (Woodroffe 2000) calculated a critical threshold of 958 people/km$^2$ before leopards went extinct in Kenya. Taken together, these arguments suggest that HPD may be a rather poor predictor of carnivore extinction overall; it's government policy/legislation, tolerance of local people, trade, and suitable cover and prey that matter rather than HPD per se (Woodroffe 2000; Linnell, Swenson, and Andersen 2001; Athreya et al. 2015).

### 5.4.4 Caveats

As with any model, the training data strongly condition the outputs (Albert et al. 2010). Yet, like most SDM, these presence data are subject to sampling bias, imperfect detection and were collected opportunistically (Guillera-Arroita et al. 2015). These limitations mean that prevalence cannot be determined and the output of these SDMs is the relative probability of observation rather than output at higher information content levels, such as true occupancy (Guillera-Arroita et al. 2015). Indeed, the opportunistic collection of data may limit our ability to interpret response curves although its impact has not been well documented (Albert et al. 2010). In essence, model results shown here may lead to attributing too much importance to the variation in response curves between different species and models. However, care was taken to not infer too much at the extremes of response curves where data are likely more sparse.

For model training, all the presence data were lumped together. Problematically, a long collection period of presence data (15 years) could leave some older presence points in currently unsuitable habitat. However, the majority of data was from 2005 and newer. Presence records were also collated from numerous sources and collection types which varied in their spatial precision, and accuracy in species identification. Yet, SDMs are relatively robust to small spatial errors (up to 5 km) (Graham et al. 2008) and its unlikely the presence data had greater error than this. In addition, these carnivores are sympatric and, particularly leopard and cheetah, identification can be confused. Finally, previous carnivore SDMs have suggested that the gender (Conde et al. 2010), season (Takahata et al. 2014), activity (Guisan and Thuiller 2005; Roever et al. 2013), and if the individual is resident or dispersing (Elliot et al. 2014b; Jackson et al. 2016) can also influence habitat selection. However, this information was unavailable for the majority of the presence records and could not be incorporated into the analysis.

The choice of predictor variables and the appropriate scale are also important components of successful species distribution models in order to capture biological processes (Hirzel and Le Lay 2008; Bucklin et al. 2014; Jarnevich et al. 2015). These carnivores predominately eat meat (Hayward and Kerley 2008), however, I had no data on prey density, or even, habitat suitability for preferred prey species at the extent of East Africa. If data on prey biomass or distribution are poor, Carroll and Miquelle (2006) recommend the use of non-prey-based models. In addition, the decision to leave out a variable on protection status could have influenced the results. Appendix 5A documents the changes resulting from its inclusion. While it was an important predictor when included, spatial predictions and evaluation metrics were mostly unchanged, and support for the three predictions did not change substantially. Support for Prediction 1 was unchanged; large carnivore habitat suitability decreased as HPD and percent human land cover increased. Support for Prediction 2 was lessened slightly; in the Extirpated option the wild dog had as shallow a curve as the leopard in response to increasing percent human land cover. Hence, wild dogs with much larger home range sizes than the leopard, had similar susceptibility to increasing human land cover. Support for Prediction 3 was also lessened slightly after inclusion of distance to protected area as percent human land cover became a less dominant parameter. However, it was still the most important variable in several modeling options and it was consistently more important than HPD. Regarding scale, predictors averaged at the largest scale (12 km) were more commonly but not universally selected as the most important predictors. This demonstrates the importance of averaging variables at multiple scales and letting the model select the best ones (Mashintonio et al. 2014; Elliot et al. 2014b), and that values averaged at roughly the size of the home range of the target species work well (Kanagaraj et al. 2013).

The study species are all sympatric large carnivores that compete with one another, occasionally stealing food and even killing each other (Palomares and Caro 1999; Caro and Stoner 2003). Indeed, improved incorporation of interspecific competition into SDM is recommended (Austin 2002; Araújo and Guisan 2006; Austin 2007). Yet, the right spatial resolution, among other aspects, is critical if competition is included in SDM (Araújo and Guisan 2006; Austin 2007). The lion is the apex predator of the community and its movement and activity have the strongest impact on the other species (Linnell and Strand 2000; Hayward and Kerley 2008). Both cheetah and wild dog densities were found to be negatively associated with lion densities across protected areas (Laurenson 1995; Creel and Creel 1996), but recent research has suggested lion impacts are stronger on wild dog movements than the cheetah (Broekhuis et al. 2013; Swanson et al. 2014). Thus, I tested the inclusion of the output of the lion SDM as an input into the wild dog model. The lion SDM became the most important predictor in variable selection and in the final model (not shown). However, the response was positive, contrary to ecological theory. The model likely identified similarities in current and extirpated habitat between the species. As expected, the SDM was not at the appropriate scale to capture the competitive effects of the lion on the wild dog.

Jarnevich et al. (2015) cautions, results of all distribution models should be treated like hypotheses and subject to continued testing. Indeed, these results are from a correlative SDM, not a mechanistic or process-based model. The exact mechanism by which human land cover, or HPD, influences these carnivores in this region is unknown. Indeed, human land uses differ in their impact on ecological systems (Burdett et al. 2010) and croplands, urban areas etc. have all been collapsed into a single layer in the GE Grids process.

Finally, results from a regional model may differ from analyses done at the site level. For instance, Poessel, Gese, and Young (2017) found that coyotes were attracted to urban areas at a broad scale while Gehrt, Anchor, and White (2009) suggested that coyotes avoided urban areas within their home range. The scale at which species are attracted to or avoid humans is likely even influenced by the type of behavior (e.g. reproductive or not) (Wilmers et al. 2013).

### 5.4.5 Conservation recommendations

Given the caveats above, these analyses have important conservation implications. Results suggest that carnivores with larger home ranges are more sensitive to habitat loss than other carnivores. In this case, cheetah and lion are particularly sensitive among large African carnivores to expanding human land cover. Leopard, on the other hand, are less susceptible to the expansion of human land cover. Indeed, although this was not tested, leopards also appeared less susceptible to increasing HPD than any of the other carnivores. While I used 70% human land cover as a rough threshold when cheetah, wild dog, and lion would completely disappear, a value of only 20-30% is enough to drop probability of observation below 50% for these species.

In addition, these results suggest that human land cover is a more important predictor in species' distributions than HPD. This indicates that particular land use strategies, or zoning, could help improve habitat suitability for carnivores by limiting the expansion of agricultural development and urbanization in carnivore habitat. Since even low levels of human land cover strongly impact carnivore distribution it may be best to cluster settlements and agricultural land rather than allow patchy development to spread across large swathes of land (Lewis et al. 2015). Thus, a conservation program based on these results may be zoning of land uses that allows pastoralism but no more than a minimal amount of agriculture in conservation zones. Agriculture would be concentrated in human zones where large carnivores are excluded. Conflict mitigation efforts could be concentrated in buffer zones around PA's and in pastoral zones with the zones designed to be large enough for viable wildlife populations (Linnell et al. 2005). Local researchers have called for the widespread adoption of land use plans that respect agro-pastoral communities, integrate wildlife, and enable development (Msoffe et al. 2011; Watson et al. 2015). Recently, Kenyan organizations have created a Land Use Master Plan for the Athi-Kaputei Plains ecosystem, Kenya's first for a pastoral area. Combined with a wildlife lease program, these actions may have contributed to continued coexistence among pastoralists and wildlife, and an increasing lion population (Matiko 2014). Schuette, Creel, and Christianson (2013) also noted that a small area of land set aside by the community for conservation and grazing in times of drought, was a critical component to maintaining high densities of lions and low levels of HWC in southern Kenya.

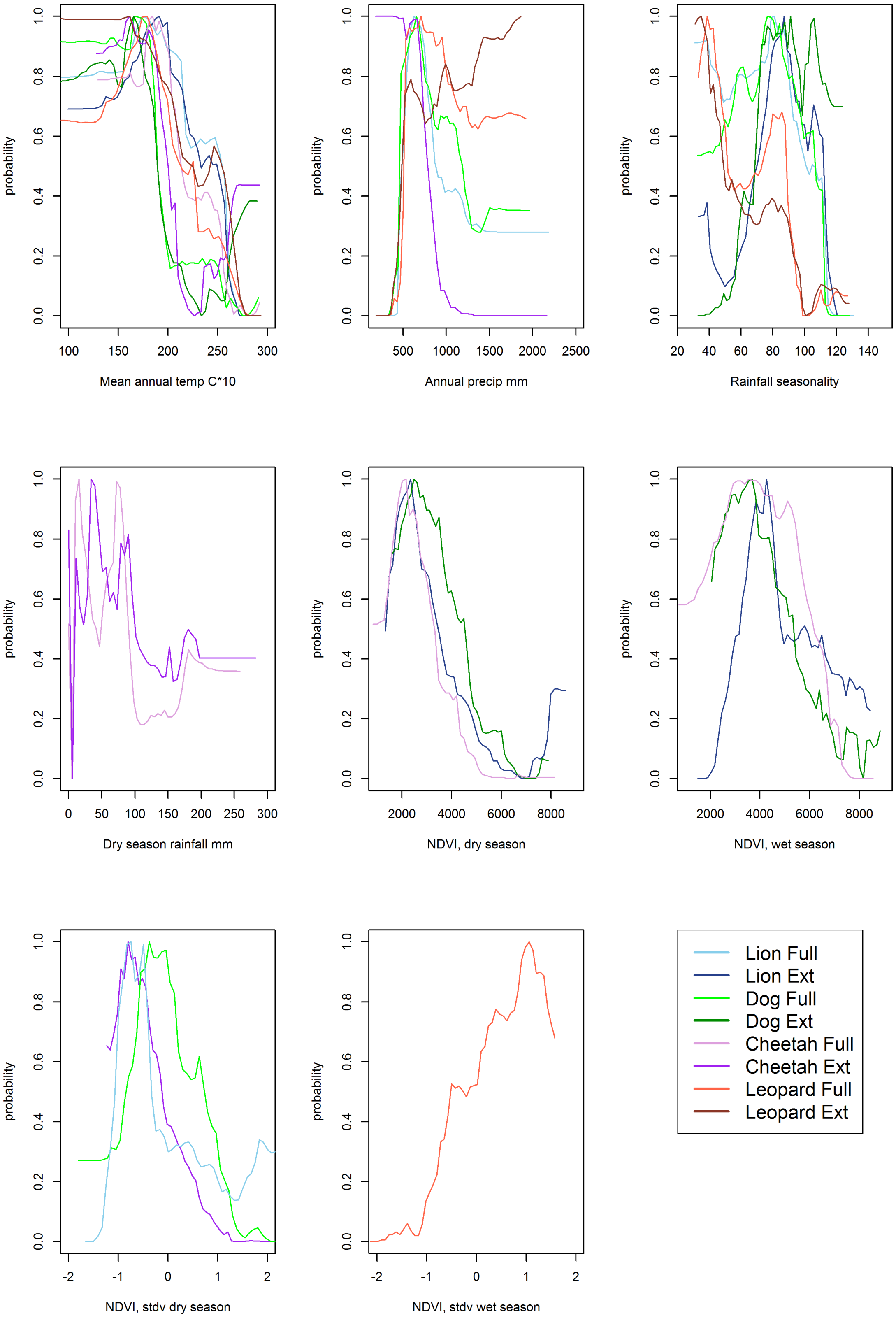
## 5.5 Conclusions

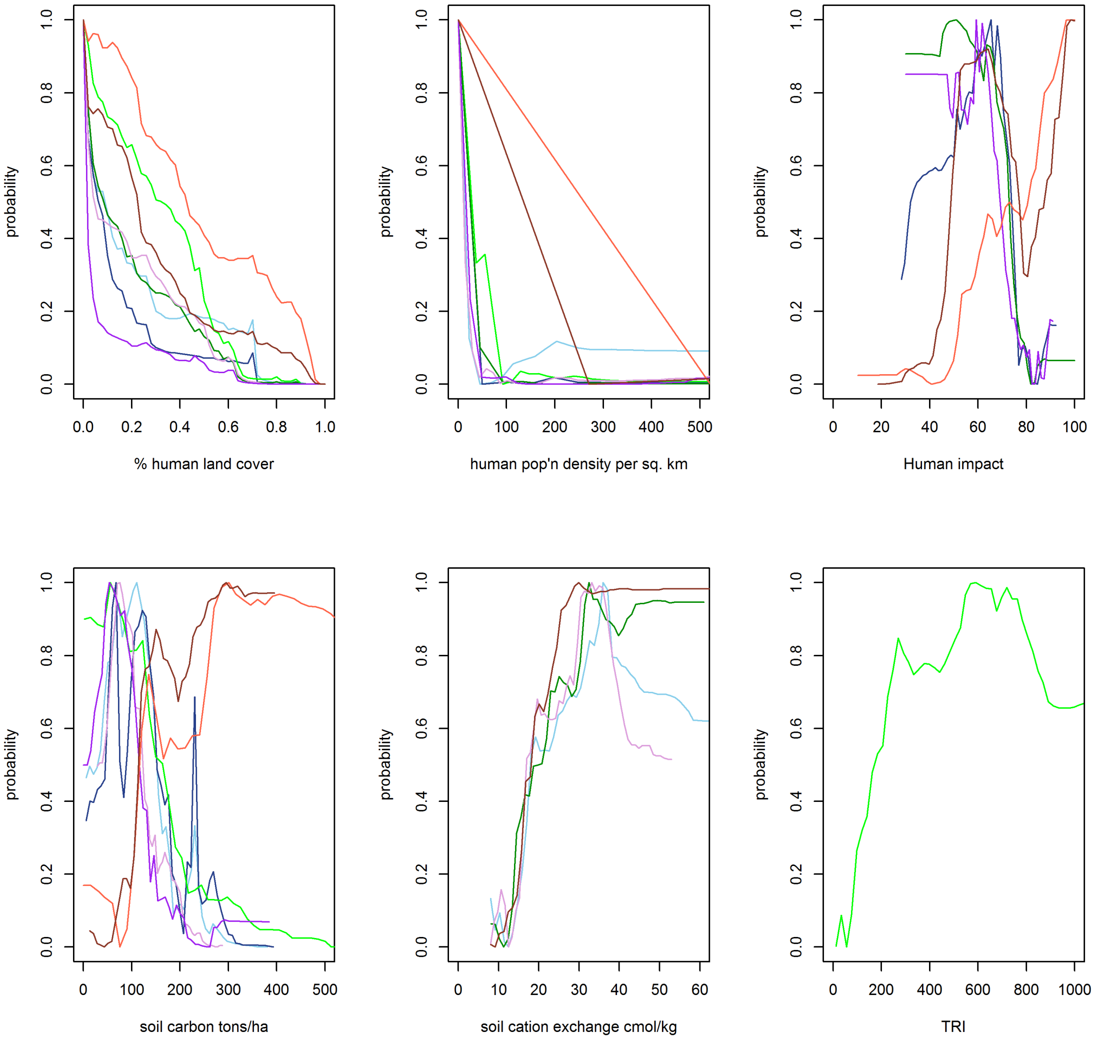
The multiple modeling approaches built inference to support the ecological hypotheses. Increasing human populations and human land cover threaten large carnivore populations in East Africa. Cheetah and lion appear most sensitive to increases in human land cover while leopard appear least sensitive to anthropogenic pressure. Finally, human land cover was a more important variable in determining species distribution than HPD, and that even low levels (20-30%) can substantially reduce habitat suitability for all carnivores.

Another utility of the approach is that training the model with PsA from areas of former habitat may help predict current habitat. In addition, comparison of response curves from different approaches helped identify regions that were losing species. Leopard is being lost from the drier regions, while the cheetah is more frequently found in hotter climates, and wild dog is more frequently found in hotter climates with greater seasonal variation in rainfall.

The protected area network is critical for these carnivores although there still remains high quality habitat outside the network and three regions (NE Kenya, SW Tanzania, and the Kenya/Tanzania border) have low Extinct Suitability for all species. Thus, for conservation, maintaining a strong network of large PA's (with minimal encroachment), and land use planning to designate multi-use landscapes that remain free of cultivation but allow pastoralism is recommended.

Human populations will continue to grow and land use will change to meet the demands of human development. But armed with the right information, political will, and adaptation by both humans and carnivores, coexistence is possible and carnivores will survive. Thankfully, carnivores are remarkably adaptable if given a chance.


probability
0.0
0.2
0.4
0.6
0.8
1.0
100
150
200
250
300
Mean annual temp C*10
500
1000
1500
2000
2500
Annual precip mm
20
40
60
80
100
120
140
Rainfall seasonality
0
50
100
150
200
250
300
Dry season rainfall mm
2000
4000
6000
8000
NDVI, dry season
2000
4000
6000
8000
NDVI, wet season
-2
-1
0
1
2
NDVI, stdv dry season
-2
-1
0
1
2
NDVI, stdv wet season
Lion Full
Lion Ext
Dog Full
Dog Ext
Cheetah Full
Cheetah Ext
Leopard Full
Leopard Ext



**Supplemental Figure 5.1** **Response curves for all species, variables and options**

# Appendix 5A Importance of protected areas in the analysis

This Appendix has two parts. The first is an examination of the inclusion of protected area status as a candidate predictor variable within the distribution modeling process, and its potential impact on the modeling output and hypotheses. The second is an examination of the importance of protected areas in terms of the habitat suitability within and outside protected areas.

## Part 1 Examining the inclusion of protected areas in distribution modeling

East Africa has an extensive network of protected areas including ~14% of the area, although this varies widely between countries (20% in Tanzania to 3.7% in Burundi). The primary source of spatial data on protected areas is the World Database on Protected Areas (WDPA), which is available on www.protectedplanet.com. The data are global and may not be accurate or up-to-date in local areas. East Africa also has various types of protection, varying from strictly protected areas, to multiple use areas which allow settlements and resource extraction. The WDPA data are therefore categorized with an IUCN category I-VII, with lower levels indicating greater level of protection. National Parks (NP) and Reserves (NR) within East Africa are IUCN category II while wildlife management areas are category IV. This data was initially excluded from the distribution modeling analysis. A majority of the sources contributing carnivore presence records centered their research within or in the buffer area of protected areas. While some sources were not tied to a specific protected area, and a few sources collecting sightings from various sources across an entire country, nevertheless, a significant majority of sources collected data tied to a protected area. I therefore left this variable out, as I felt it would likely inflate the importance of protected areas in the analysis. Here I test the impact of its inclusion on the results of the analysis in this Appendix. As expected, distance to protected area is an important variable and included in the final models for nearly all species and modeling options. However, the modeling outcomes are not substantially different from those in which the variable is not included. In addition, the conclusions regarding the three hypotheses do not change.

## Methods

Protected areas were downloaded from Protected Planet in 2015 (UNEP and IUCN 2015) for East Africa and the surrounding countries. Some edits were made to improve the relevance of protected areas for this analysis. I removed all marine protected areas, all 'proposed' protected areas, and internationally designated areas (e.g. Man and Biosphere Reserves, Ramsar sites). Minor edits were also made to the metadata to ensure consistency (e.g. to ensure all 'national parks' were given an IUCN category of II as opposed to N/A). Protected areas in Tanzania were updated based on spatial data from the Tanzania National Parks Authority.

I created three candidate predictors relating to protected status: a protection code given to the entire study area, a binary layer of protected or not, and distance to protected area. First, I classified all areas with a protection code from 1-6 according to their assumed effectiveness in wildlife protection and management of natural resources. The existing IUCN categories can poorly relate to actual protection and are often missing and/or not standardized across countries (Peter Lindsey pers. comm. and see Lindsey et al. 2017). Therefore, I chose not to use that classification and instead create my own protection code classification. National Parks and Reserves were given a designation of 1. Other federally managed protected areas not NPs or NRs were assigned a 2. Privately managed protected areas were assigned a 3. Wildlife management areas (i.e. community managed) were given a 4. Any other protected areas not included in the above categories were given a 5. All areas outside the protected area system were given a 6. This was my first candidate protection layer. Then, I set a size threshold of 225 km$^2$, roughly the average home range size of a lion (Celesia et al. 2009). All protected areas smaller than this (including Nairobi NP at ~115 km$^2$) were excluded. I selected only areas with a protection code 1 and 2 (i.e. all federally managed protected areas). This I set as a binary layer, protected or not, designating areas assumed to have the most effective wildlife protection that were large enough to contain at least the home range of a lion. Finally, from this binary layer, I calculated distance to protected area. I followed a similar process of designating areas within a 20 km buffer of the study area to include areas like Niassa NR in Mozambique before calculating the distance to PA. All analysis was conducted in ArcGIS 10.2, converted to raster with a cell size of 500 m, and projected in WGS 1984 Africa Albers Equal Area Conic projection.

Initially all three candidate predictor variables were included in the variable selection process for the lion and cheetah along with the other 62 predictors. The distance to protected area variable was substantially more important than the other two candidate predictors in these trial runs. The analysis was subsequently re-run for all species and options using only the distance to protected area predictor. The methods after inclusion of the new predictor followed the methods detailed in Chapter 5.

## Results/Discussion

The distance to protected area predictor was highly important (Table 5A.1). The predictor was included in the final model for all four species and three model options except in one case (wild dog, Full). It was relatively less important in the Extinct Suitability option for all four species, but was the most important variable in several instances (Lion, Full and Extirpated; Leopard, Full and Extirpated; Wild dog, Extirpated). Despite the inclusion of the new variable, the spatial output of each species and model option was highly correlated with the output from the original model; all correlations were between 0.88 and 0.98. A comparison of the spatial output and all three model options is shown using the cheetah as an example (Figure 5A.1). As expected, inclusion of the distance to protected area variable increased the probability of observation in and near protected areas, and decreased the probability in areas far from protected areas. Evaluation metrics after the inclusion of the new predictor were almost identical to the original metrics, often within the standard deviation of the original values (Table 5A.2). The only stronger consistent difference was with the lion whose evaluation metrics were slightly improved in both Full and Extirpated options with inclusion of the new predictor.

I also examined how the inclusion of the distance to protected area parameter may have affected the results and conclusions regarding the three predictions. In support of Prediction 1 (increasing human population densities and land cover negatively impact carnivore presence), increasing levels of human land cover and human population density still had a negative impact on carnivore persistence (Figure 5A.2). This was unchanged from the original result.

The original analysis gave partial support for Prediction 2 (species with larger home ranges such as cheetah and wild dog are more vulnerable to increases in human land cover). This was unchanged by the inclusion of the new parameter. The global concavity parameters representing the response curves for percent human land cover were significantly different across species in both modeling approaches (Full: $F = 52.5$, $p<0.0001$; and Extirpated: $F = 51.5$, $p<0.0001$). Wild dog and leopard had the highest concavity parameters in both the Full and

Extirpated Options in both the original and revised analysis (Figure 5A.3). Higher concavity parameters represent shallower curves as percent human land cover increased, i.e. a slower reduction in suitability as percent human land cover increases. Different from the original analysis, the wild dog had the highest concavity parameter, slightly larger than the leopard, in the Extirpated option.

After the inclusion of the distance to protected area predictor, percent human land cover became a less dominant variable than in the original analysis (Table 5A.1). Previously it was the most important predictor in 10 of the 12 models overall, but after the inclusion of the new predictor it was the most important predictor in only three options (wild dog, Full; and Extinct Suitability for lion and leopard). However, it was still the more dominant variable compared to human population density supporting Prediction 3 (human land cover is a more important variable restricting carnivore distribution than population density). In only two cases was population density more important than land cover, and they were the same instances as in the original analysis (wild dog, Extinct Suitability; leopard, Full). Thus, human land cover was still more important vis-à-vis population density, but its dominance was muted in the presence of the distance to protected area predictor.

**Table 5A.1 Variable importance for all species and model options after inclusion of distance to protected area as a candidate predictor. See Table. 4.3 as a guide to acronym meanings.**

**Cheetah**

| Full | Rank | Extirpated | Rank | Extinct Suitability | Rank |
|---|---|---|---|---|---|
| MAT (6km) | 1.00 | Precip (12km) | 1 | Precip (12km) | 1 |
| NDVI $_{dry}$ (3km) | 0.96 | % Human LC (12km) | 0.79 | % Human LC (12km) | 0.78 |
| % Human LC (12km) | 0.88 | NDVI StDv $_{dry}$ (12km) | 0.59 | HPD (12km) | 0.65 |
| Distance to PA | 0.87 | Distance to PA | 0.58 | Human Impact | 0.58 |
| NDVI $_{wet}$ (3km) | 0.81 | MAT (12km) | 0.42 | MAT (3km) | 0.53 |
| Soil Carbon (12km) | 0.53 | HPD (12km) | 0.42 | Soil Carbon (12km) | 0.43 |
| HPD (6km) | 0.49 | Soil Carbon (12km) | 0.41 | Precip $_{seasonality}$ (12km) | 0.41 |
| Soil CEC (12km) | 0.47 | Precip $_{dry\ season}$ (12km) | 0.40 | Distance to PA | 0.41 |

### Wild dog

| Full | Rank | Extirpated | Rank | Extinct Suitability | Rank |
|---|---|---|---|---|---|
| % Human LC (12km) | 1.00 | Distance to PA | 1.00 | Precip (12km) | 1 |
| MAT (3km) | 0.95 | % Human LC (12km) | 0.96 | HPD (12km) | 0.79 |
| Precip (12km) | 0.84 | MAT (12km) | 0.81 | % Human LC (12km) | 0.74 |
| HPD (12km) | 0.63 | Precip (3km) | 0.75 | NDVI StDv $_{dry}$ (12k)m | 0.68 |
| NDVI StDv $_{dry}$ (12km) | 0.61 | NDVI StDv $_{dry}$ (12km) | 0.72 | Human Impact | 0.59 |
| Precip $_{seasonality}$ (6km) | 0.59 | HPD (12km) | 0.60 | Precip $_{seasonality}$ (12km) | 0.48 |
| TRI (500m) | 0.40 | Human Impact | 0.57 | MAT (12km) | 0.34 |
| Soil Carbon (12km) | 0.37 | Precip $_{seasonality}$ (12km) | 0.54 | Distance to PA | 0.33 |

### Leopard

| Full | Rank | Extirpated | Rank | Extinct Suitability | Rank |
|---|---|---|---|---|---|
| Distance to PA | 1.00 | Distance to PA | 1.00 | % Human LC (12km) | 1 |
| Precip $_{seasonality}$ (12km) | 0.63 | % Human LC (12km) | 0.81 | Distance to PA | 0.89 |
| HPD (6km) | 0.55 | MAT (12km) | 0.45 | HPD (12km) | 0.76 |
| % Human LC (12km) | 0.51 | Precip (6km) | 0.45 | Precip (12km) | 0.72 |
| Precip (6km) | 0.49 | Precip $_{seasonality}$ (6km) | 0.40 | Precip $_{seasonality}$ (12km) | 0.70 |
| MAT (12km) | 0.48 | HPD (12km) | 0.31 | MAT (3km) | 0.65 |
| Soil Carbon (6km) | 0.30 | Human Impact | 0.27 | Human Impact | 0.52 |
| NDVI StDv $_{dry}$ (3km) | 0.29 | Soil Carbon (12km) | 0.24 | Soil CEC (12km) | 0.51 |

### Lion

| Full | Rank | Extirpated | Rank | Extinct Suitability | Rank |
|---|---|---|---|---|---|
| Distance to PA | 1 | Distance to PA | 1 | % Human LC (12km) | 1 |
| Precip (3km) | 0.74 | % Human LC (12km) | 0.93 | Precip $_{dry\ season}$ (12km) | 0.77 |
| MAT (6km) | 0.72 | MAT (12km) | 0.82 | HPD (12km) | 0.59 |
| % Human LC (12km) | 0.55 | NDVI $_{dry}$ (500m) | 0.57 | Human Impact | 0.43 |
| Precip $_{seasonality}$ (6km) | 0.41 | HPD (12km) | 0.55 | NDVI StDv $_{dry}$ (12km) | 0.41 |
| HPD (12km) | 0.38 | NDVI $_{wet}$ (12km) | 0.51 | Distance to PA | 0.35 |
| NDVI StDv $_{dry}$ (3km) | 0.35 | Precip $_{seasonality}$ (6km) | 0.43 | Precip (3km) | 0.34 |
| Soil Carbon (12km) | 0.31 | Soil Carbon (12km) | 0.41 | MAT (12km) | 0.21 |

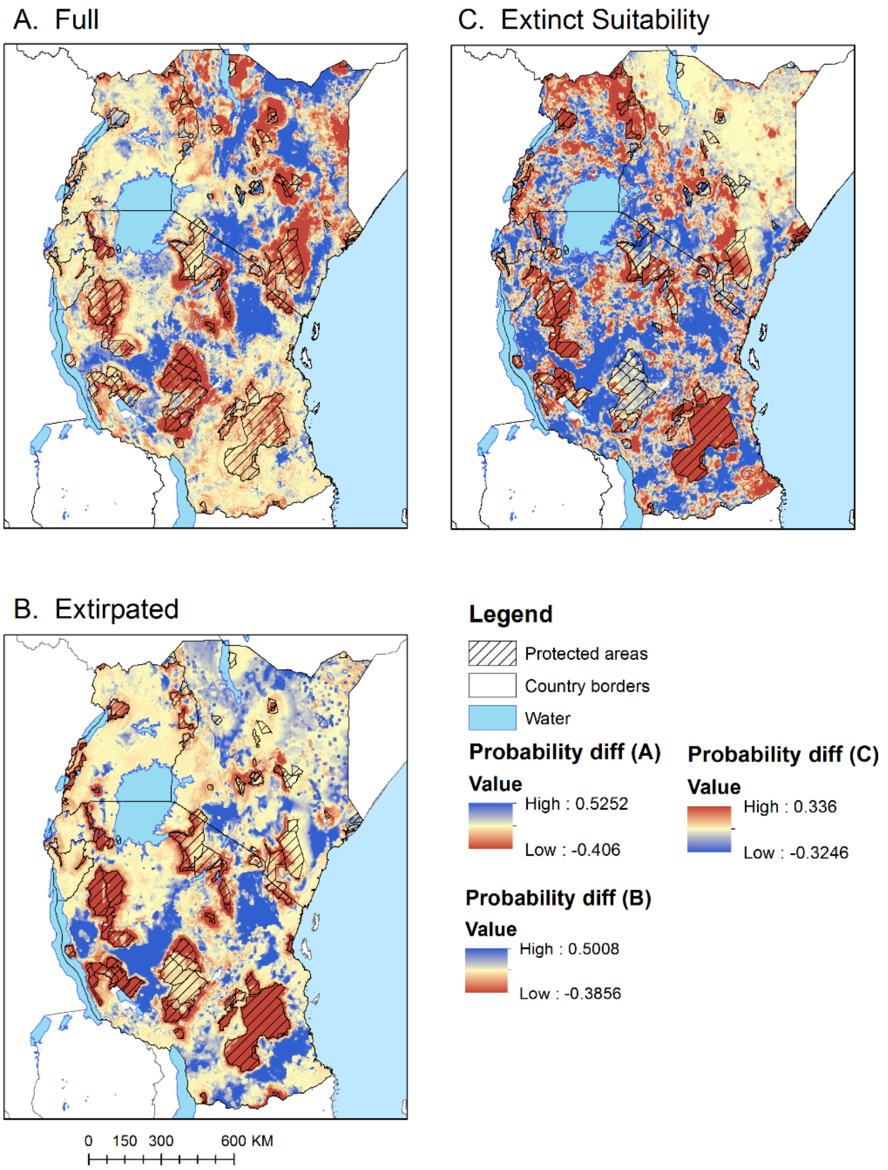


**Figure 5A.1** A spatial comparison for the cheetah for all three model options. The new output (with distance to PA as a variable) is subtracted from the original output to produce a difference map where positive difference values indicate higher probability of observation in the original than in the new output (in the Full and Extirpated options). Note that in the Extinct Suitability option, a higher value indicates increased likelihood of absence rather than presence. Thus, in the difference map (C), a more positive value represents a higher probability of absence in the original, and a higher probability of presence in the new output.

**Table 5A.2 Comparison of evaluation metrics with and without the addition of the distance to protected area predictor.**

| | Option | | Boyce Index | AUC | Out of bag error |
|---|---|---|---|---|---|
| **Cheetah** | Original | Full | 0.885 ± 0.014 | 0.987 ± 0.003 | 0.153 ± 0.018 |
| | | Extirpated | 0.878 ± 0.009 | 0.999 ± 0.001 | 0.051 ± 0.006 |
| | With distance to PA added | Full | 0.873 ± 0.007 | 0.985 ± 0.004 | 0.149 ± 0.012 |
| | | Extirpated | 0.891 ± 0.011 | 0.999 ± 0.001 | 0.049 ± 0.005 |
| **Wild dog** | Original | Full | 0.873 ± 0.008 | 0.988 ± 0.002 | 0.226 ± 0.009 |
| | | Extirpated | 0.885 ± 0.007 | 0.994 ± 0.002 | 0.111 ± 0.013 |
| | With distance to PA added | Full | 0.866 ± 0.006 | 0.986 ± 0.004 | 0.232 ± 0.018 |
| | | Extirpated | 0.879 ± 0.011 | 0.997 ± 0.001 | 0.102 ± 0.007 |
| **Lion** | Original | Full | 0.879 ± 0.004 | 0.987 ± 0.003 | 0.174 ± 0.011 |
| | | Extirpated | 0.885 ± 0.01 | 0.998 ± 0.001 | 0.078 ± 0.005 |
| | With distance to PA added | Full | 0.885 ± 0.008 | 0.99 ± 0.003 | 0.156 ± 0.011 |
| | | Extirpated | 0.891 ± 0.01 | 0.999 ± 0.001 | 0.061 ± 0.004 |
| **Leopard** | Original | Full | 0.859 ± 0.007 | 0.983 ± 0.006 | 0.275 ± 0.023 |
| | | Extirpated | 0.873 ± 0.009 | 0.997 ± 0.001 | 0.119 ± 0.013 |
| | With distance to PA added | Full | 0.866 ± 0.005 | 0.98 ± 0.005 | 0.261 ± 0.021 |
| | | Extirpated | 0.879 ± 0.008 | 0.998 ± 0.002 | 0.104 ± 0.011 |

A. Full option B. Extirpated option

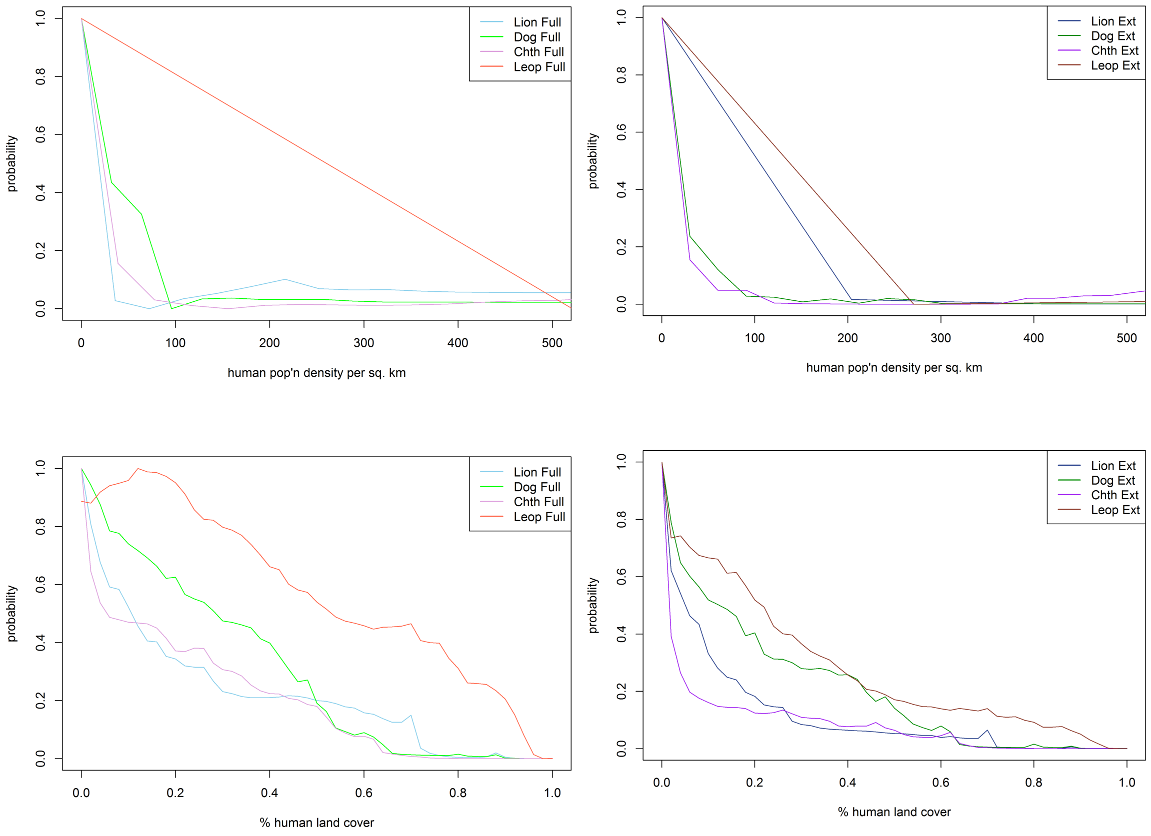
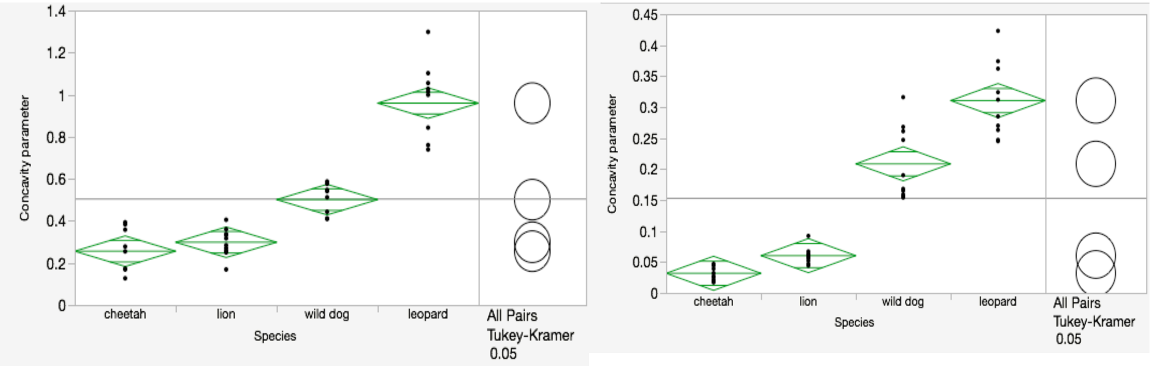
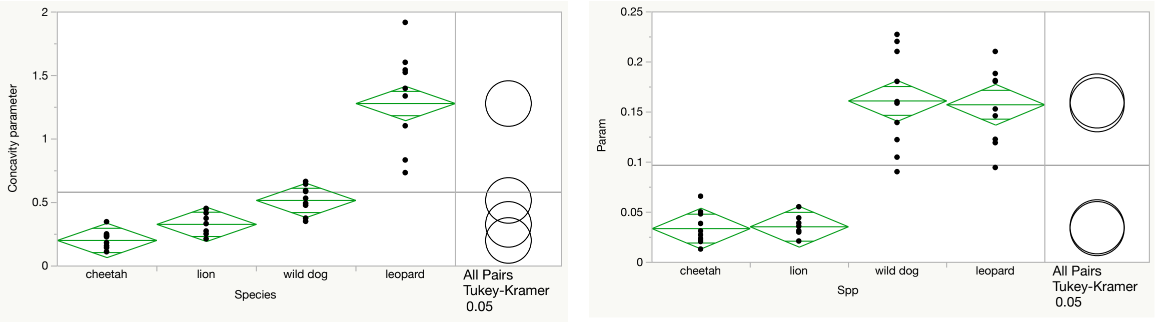


**Figure 5A.2 Response curves for all four species in response to percent human land cover and human population density in the Full (column A) and Extirpated (column B) options with the inclusion of distance to protected area predictor.**



A. Full, original analysis

B. Extirpated, original analysis



C. Full, with distance to PA variable added

D. Extirpated, with distance to PA added

**Figure 5A.3 Concavity parameters for all four species in both options, with (C & D) and without (A & B) the inclusion of distance to protected area predictor.**

## Conclusion

Distance to protected area is an important predictor in distribution models for all four large carnivore species in East Africa. In many cases, it became the most important predictor. This was expected as the majority of the sources for the carnivore presence records searched within protected areas or in their buffer zones. Yet, the spatial output was highly correlated to the original output and evaluation metrics nearly identical. Regarding the conclusions for the three predictions, none were changed with the inclusion either. However, there was only partial support for Prediction 2 in the original analysis and support was further lessened after the inclusion of distance to protected area. In addition, support for Prediction 3 was also lessened after the inclusion.

## Part 2 - Examining the importance of protected areas in the landscape

As a gauge to understand how important the protected area system is in relation to the outputs from the distribution modeling process, I looked at two species, the cheetah and the lion. A high percentage of cheetah range is outside the protected area system (Durant et al. 2016) whereas the lion probably has a relatively higher reliance on protected areas in this landscape than the other focal species. I assumed that if the protected area system is valuable for carnivores, it would contain a substantial proportion of the very high and high suitability classes. Alternatively, I expected the reverse in the Extinct Suitability option. The protected area system should hold a substantial portion of the low and very low absence suitability values. I examined if this was true.

To do this, I used the protection code developed previously (see Methods section above from this Appendix). Altogether, less than a quarter of the study area is under some form of protection code (Figure 5A.4). I then divided the distribution modeling outputs (see Chapter 5) into different suitability classes. I used five classes: very high suitability 1 - 0.8; high suitability 0.79 – 0.6; suitable 0.59 – 0.4; low suitability 0.39 – 0.2; and very low suitability 0.19 – 0. Therefore the very high suitability class in the Full and Extirpated options captured the areas with the highest probability of observation. This is reversed for the Extinct Suitability option, as the very high suitability represents very high suitability of absence. I then calculated the proportion of the suitability class within each protected area code. I repeated this for every suitability class. I further repeated this process for both the cheetah and the lion, and all three modeling options.

I assumed that the protected area system would be more valuable if it contained a substantial proportion of the very high and high suitability classes. Alternatively, I expected the reverse in the Extinct Suitability option. This indeed occurs. In the Full and Extirpated options for the lion, nearly 60% of the very high suitability class is within some form of protection (Figure 5A.5). As suitability declines, the percentage of area under protection also declines. Thus, the most suitable habitat tends to be under some form of protection, and the rate of protection decreases as suitability decreases. This pattern holds but is less pronounced for the cheetah than the lion. Hence, relatively more of suitable cheetah habitat is outside protected area. Indeed, this threat is widely acknowledged, and is a driving reason why Durant et al. (2016) recommend uplisting the cheetah on the IUCN Red List to Endangered.

The expectations for the Extinct Suitability options also hold. The pattern is even more pronounced than in the Full or Extirpated options. Virtually no area of very high absence suitability is under any form of protection. Slightly more areas of high absence suitability have some form of protection. As absence suitability declines, a larger portion of the area is found within the protected area system. Indeed, the higher classes of protection (e.g. National Parks and Reserves) contain almost no areas of higher suitability for absence (Figure 5A.6).

In sum, the protected area system does contain a substantial portion of suitable habitat. Over a third of the very high suitability area for the cheetah is found within national parks and reserves which together account for only 6% of the landscape (as identified in the Full modeling option). This is more pronounced for the lion than for the cheetah, with 43% of the very high suitability areas found within these types of protected areas (also with the Full option). In addition, the protected area system contains very little area thought to have a high chance for extinction for either the cheetah or lion.

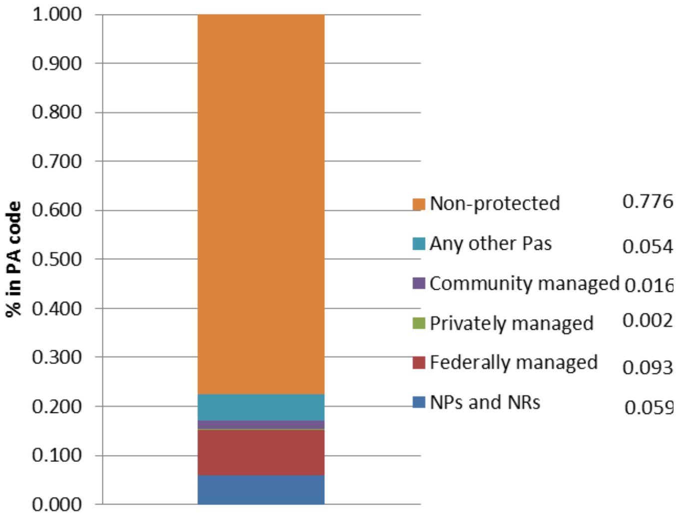
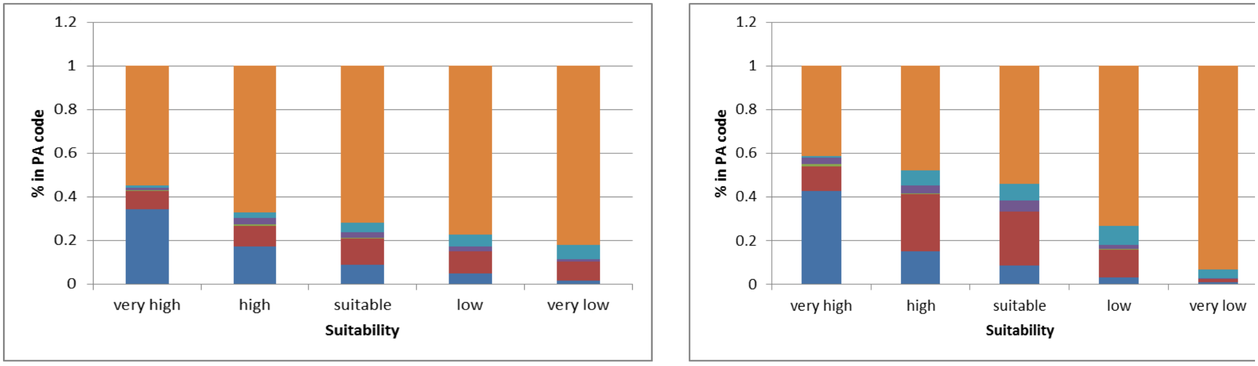


**Figure 5A.4** **Proportion of the area of the protection codes found within the study area.**

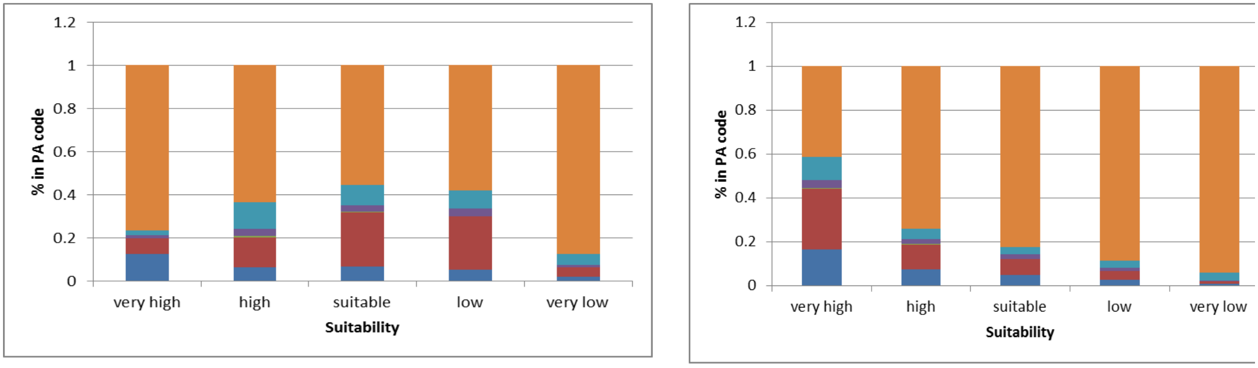
Cheetah

Lion

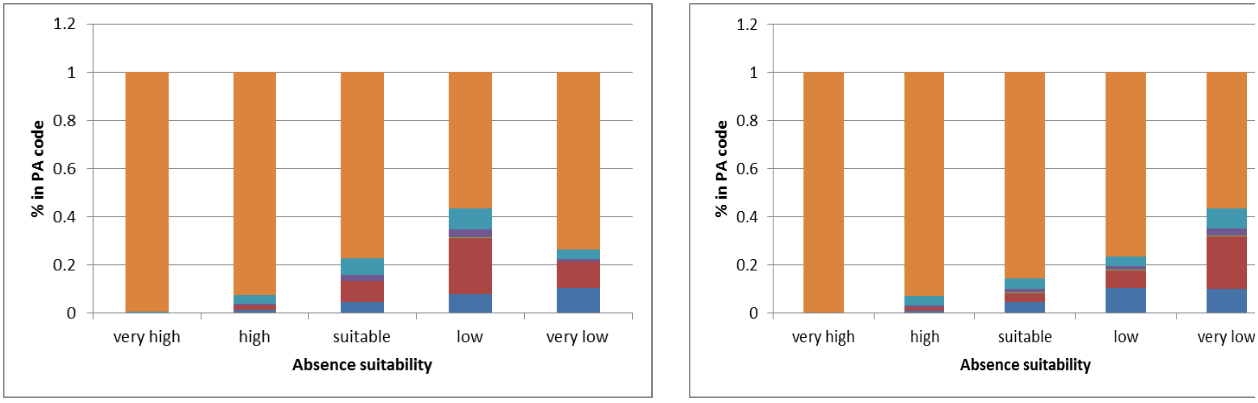
Full option



Extirpated option



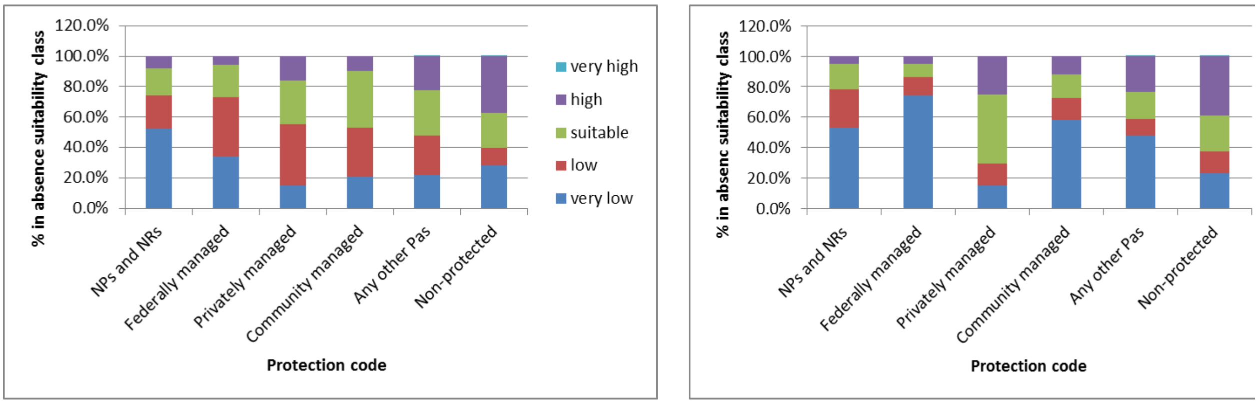
Extinct Suitability option



PA Code

🟧 Non-protected
🟦 Any other Pas
🟪 Community managed
🟩 Privately managed
🟥 Federally managed
🟦 NPs and NRs

**Figure 5A.5 Proportion of suitability classes within protected area codes for the cheetah and the lion. This is repeated for each of the three modeling options.**



Cheetah

Lion

**Figure 5A.6 Proportion of absence suitability in each protection code for the cheetah and lion. Almost no areas of very high or high suitability of absence are within national parks or reserves, or other federally managed protected areas.**

# Appendix 5B Areas of conservation interest

This Appendix identifies areas of conservation interest resulting from combining the spatial predictions of the distribution modeling and the expert-derived range maps. These are compared in two primary methods. The first is simply examining areas of high or low probability and the range maps. The second approach uses the response curves from the models and sets a threshold above which probability of observation is zero for both human population density (HPD) and percent human land cover. This threshold value is then applied to the HPD and percent human land cover variables to create a binary layer representing areas above and below these thresholds. This is the second layer I compared with the species range maps. Together, these approaches can help identify areas of conservation interest, such as areas of high suitability but that are not incorporated into extant range (i.e. representing potential extensions of extant range after field surveys), or areas of current range that are highly threatened.

## Methods

Definitions for range categories for cheetah and wild dog are defined in (IUCN/SSC 2007). The categories are:

1) Resident range: land where the species is known to be still resident.
2) Possible range: land where the species is still resident, but where residency has not been confirmed in the last 10 years.
3) Connecting range: land where the species may not be resident but which dispersing animals may use to move between occupied areas.
4) Unknown range: land where the species' status is currently unknown and cannot be inferred
5) Extirpated range: land where the species has been extirpated
    a. Recoverable range: currently extirpated land but where habitat and prey remain over sufficiently large areas that either natural or assisted recovery may be possible.

I merged possible and unknown range for simplification in this analysis.

For the leopard and lion, extant and extirpated categories are similar to the definitions above, the connecting range category was not used, and possible range represents temporary and/or possible occupation.

Comparing the areas of higher probability from the Full and Extirpated options against species range maps can suggest areas that are suitable habitat and may either be areas ripe for species reintroduction, or that contain residual, lesser-known populations. The Extinct Suitability option may be particularly useful in this regard as it would highlight areas that do not look like extirpated habitat within historical range. The Extirpated option is also highly useful, as it would identify areas that look like current habitat. So for instance, with the cheetah, the Extinct Suitability option indicates lower probabilities of absence in much of SE Tanzania. However, presence in the Selous GR and surrounding areas is uncertain (IUCN/SSC 2007), and the Extirpated option matches this low probability of observation. Yet, areas where both (or even better, all three) options agree suggest that the area may be highly suitable. On the contrary, areas of low probability in the Extirpated and Extinct Suitability options that are within resident (or extant) range indicate areas where the species may be under severe threat.

The second approach here sets threshold values for the HPD and percent human land cover layers above which species presence is negligible. The response curves suggest thresholds for percent human land cover and human population density above which the likelihood of species' observation is essentially zero (see Figures 5.6 & 5.7). Three species, cheetah, wild dog and lion, have very similar threshold values while the leopards' is substantially higher. For the Full option, the threshold for percent human land cover is ~0.7 for the three species, and ~0.95 for the leopard. For HPD, the threshold response was less consistent and so I selected two threshold values; ~50 people/km$^2$ and 100 people/km$^2$ for the three species, and ~280 people/km$^2$ and ~500 people/km$^2$ for the leopard. I applied these threshold values and created a binary layer for both percent human land cover and HPD. These layers were overlaid on each species' distribution to identify where the species may be gravely threatened.

## Results

After applying a threshold to percent human land cover and HPD, I overlaid these layers on each species' distribution map to identify areas of conservation importance (Figures 5B.1 – 4). For each species, some areas of resident or other non-extirpated range overlapped with values above the percent human land cover or HPD threshold.

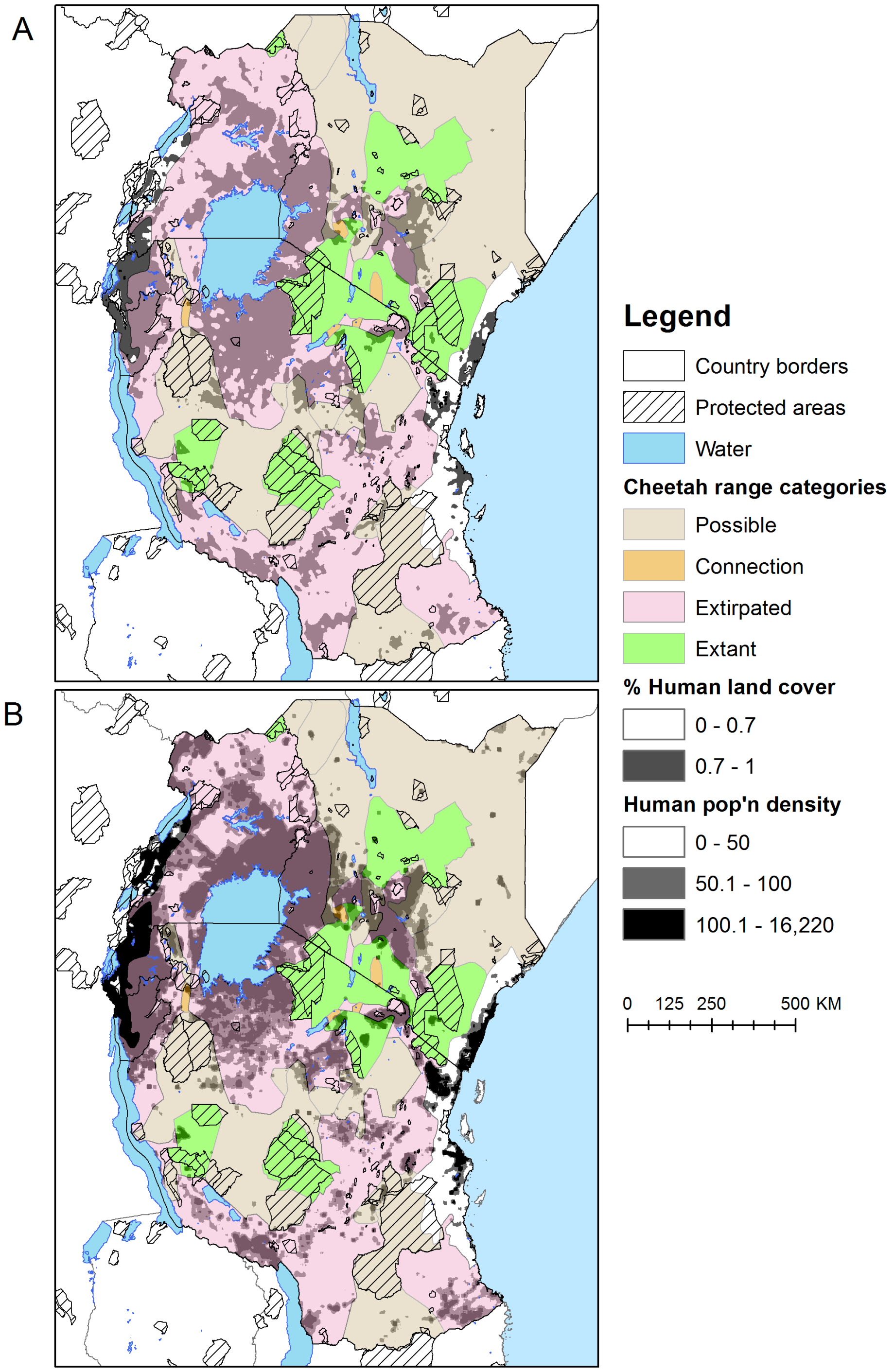


**Figure 5B.1** Thresholds for percent human land cover (A), and human population density (B), overlaid on expert-derived cheetah distribution.

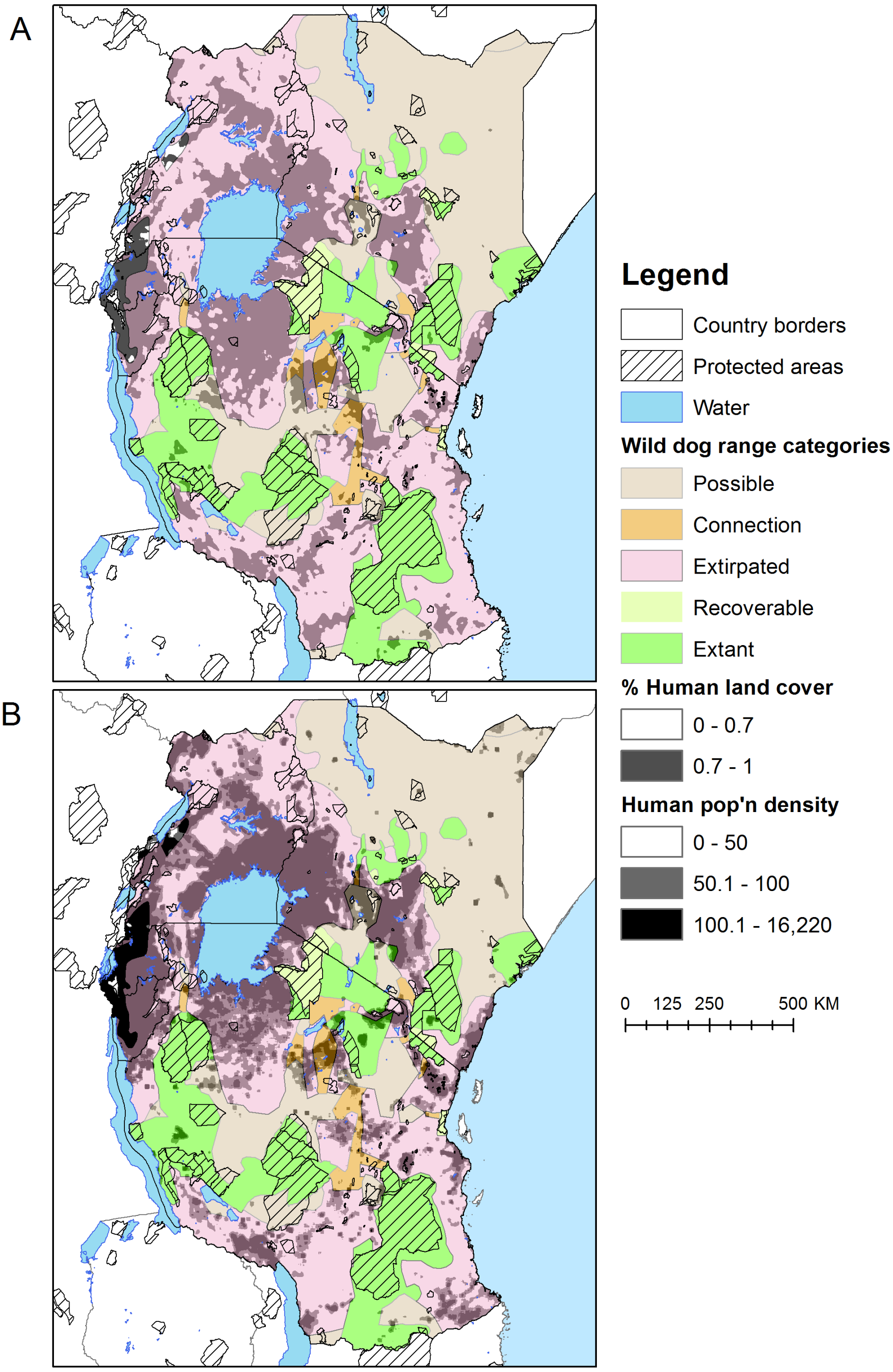


**Figure 5B.2 Thresholds for percent human land cover (A), and human population density (B), overlaid on expert-derived wild dog distribution.**

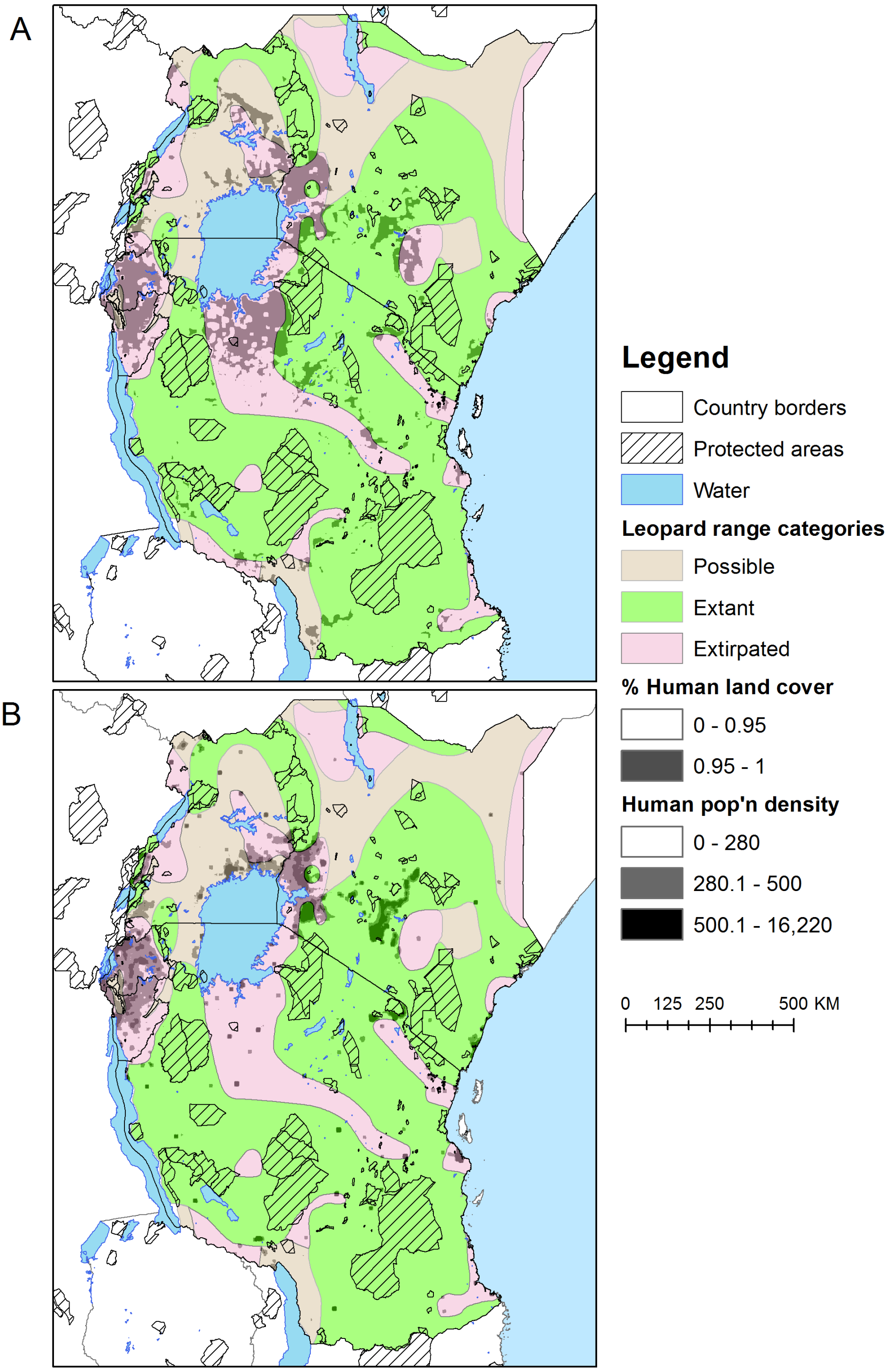


**Figure 5B.3** Thresholds for percent human land cover (A), and human population density (B), overlaid on expert-derived leopard distribution.

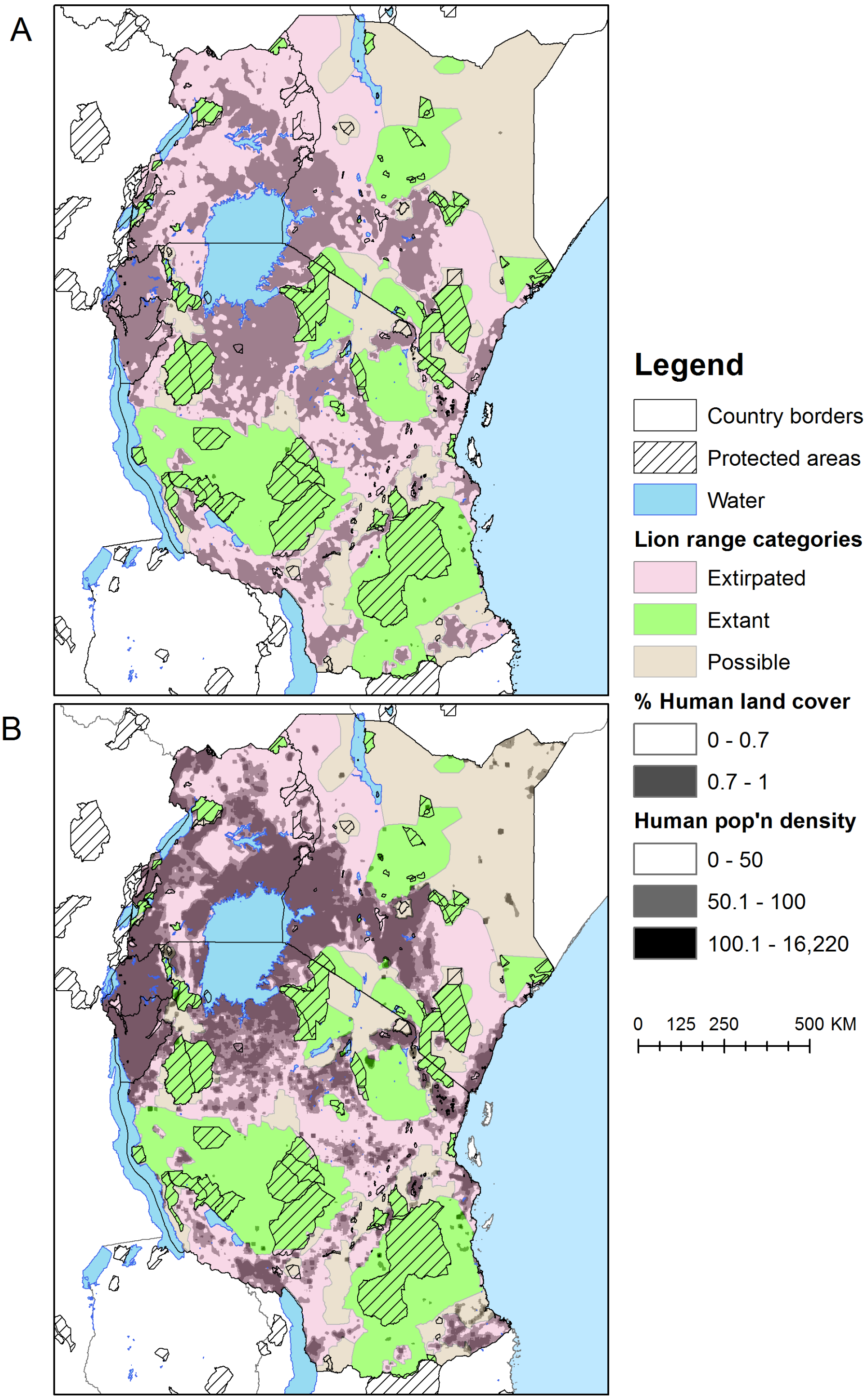


**Figure 5B.4** Thresholds for percent human land cover (A), and human population density (B), overlaid on expert-derived lion distribution.

## Discussion

Results from the species distribution models (SDMs) indicate that protected areas are important habitat for all species (see Appendix 5A). Yet, all four carnivores range beyond protected areas in East Africa and multi-use landscapes are also important (see Figures 5.2 – 5.6). North-central Kenya (e.g. Laikipia county) and the Kenya-Tanzania border, are also highly important for all species. However, the importance of these two regions may be artificially high since many researchers work there. SW Tanzania and NE Kenya also appear important although fewer researchers work there. Both areas have relatively low human pressures and hence are accentuated in the Extirpated and Extinct Suitability approaches. Although these two regions are not highly suitable for all species (e.g. lions in NE Kenya), they may provide good refugial habitat.

Problematically, outside of Kenya and Tanzania, the models suggest few large regions of higher carnivore probability. Within Burundi and Rwanda, there are higher probabilities in Nyungwe National Park (NP) (primarily for the leopard), and also along the border with Tanzania at Akagera NP (primarily cheetah and lion) and Kibira NP (primarily leopard). This is unsurprising as only the leopard has any non-extirpated range within those countries (although lion have been reintroduced at the partially fenced Akagera NP in Rwanda). In Uganda, the protected area network still contains lion, leopard, and cheetah (although cheetah are restricted to Kidepo Valley NP) (UWA 2010). Spatial predictions suggest almost no areas of higher probability for the cheetah or wild dog in Uganda, and higher probabilities for the lion are largely contained to the protected areas. Along the remote eastern border of Uganda and Kenya, some areas have retained higher probabilities for all species.

For the cheetah, Moyowosi and Kigosi Game Reserves (GRs), and all of NE Kenya represent potential habitat outside currently known range. In addition, east of Tarangire NP has higher probability of observation, as well as SW Tanzania between the Ruaha and Katavi ecosystems. These areas are currently identified as possible range (Durant et al. 2016), and may represent range extensions if the species is found there. On the other hand, resident range just south of Lake Nakuru NP in Kenya is above threshold values of HPD and human land cover, and has a low probability of observation. Thus, if cheetahs are still present here they are likely under severe threat. There are other pockets or edges of resident range that also have low probabilities, such as the western part of the Katavi ecosystem, and the northern and western part of the Tarangire ecosystem.

For the wild dog, there are many extensions of higher probabilities outside resident range including all of NE Kenya, but also parts of the Maasai Steppe and unprotected land in SW Tanzania (east of Ugalla Forest Reserve and west of Rungwa GR). The Serengeti ecosystem is identified as recoverable habitat (IUCN/SSC 2007) and these models suggest high probability of observation. However, the models do not account for interspecific competition, which was postulated in the disappearance of the wild dog from the national park (Ginsberg, Mace, and Albon 1995). Additionally, there are few areas in resident range that either have low suitability, or are above the anthropogenic thresholds. The same areas of northern and western Tarangire ecosystem are threatened as with the cheetah, along with a pocket of low probability just north of Katavi NP. The Full and Extirpated models also suggest there is a weakening connection between Mikumi NP and the Selous GR. Experts also identified a number of potential connections or possible habitat that linked many of the resident patches together (e.g. connecting habitat between the Laikipia patch and Maasai Steppe). Unfortunately, very few of these potential connections have high suitability, and most traverse through high levels of HPD or human land cover which dispersing wild dogs tend to avoid (Masenga et al. 2016).

For the leopard, much of central Uganda is outside of resident range (Jacobson et al. 2016) but has high probability of observation. This is unusual among the four species and seems likely due to the leopard's tolerance of higher levels of HPD and human land cover. In addition, a patch of western Kenya connecting Laikipia with eastern Uganda has higher suitability. In southern Tanzania near the border with Zambia, there are higher probabilities of observation in some areas outside resident range. Indeed there is some evidence that leopards persist in these areas of high probability, i.e. the Kipingere Range (aka Livingstone Mountains), Kitulo Plateau, and Mount Rungwe (Foley et al. 2014). The leopard is the only species where essentially all areas of high probability are connected. However, there are some areas of low suitability and/or above anthropogenic thresholds in resident range including parts of NE and coastal Kenya. In addition, parts of central Kenya and central Tanzania have many fragmented patches of high probability rather than continuous high probability, and potentially represent local population sinks.

Finally, for the lion, much of the Maasai Steppe and along the Tanzania-Kenya border has high probability of observation. However, the models do suggest increasingly tenuous connections between the Tarangire ecosystem to the NW with the Ngorongoro Conservation Area, and to the SE with Mikumi NP. There is a small patch of higher probability around Lake Eyasi in Tanzania that is outside of resident range (indeed this is true for cheetah and wild dog too). This overlaps with the range of the Hadzabe people, a hunter-gatherer tribe, and should be protected from encroachment. There are also some small extensions of high probability around the Selous ecosystem, the Tsavo ecosystem, and in northern Kenya near the border with Ethiopia and Somalia. Yet, there are pockets of low probability within many of the resident range patches, including much of the area east of Marsabit National Reserve, where lions may be increasingly threatened.

# Chapter 6

# Improving spatial risk modeling of point data in conservation: an example with livestock depredation from northern Tanzania

A Maasai woman indicates where a predator attacked her family's sheep. © Andrew Jacobson

# Chapter 6 Improving spatial risk modeling of point data in conservation: an example with livestock depredation from northern Tanzania

## Abstract

Human-wildlife conflict (HWC) is a widespread and global issue. Conflict can lead to retaliatory or pre-emptive killing of wildlife and direct killing is a leading threat to large carnivores globally. Researchers are increasingly applying species distribution modeling (SDM) techniques to HWC in the hopes of generating insight into risk factors and producing spatial risk maps. In turn, researchers hope conflict can be reduced and coexistence improved. However, when reviewing SDM practices as applied to HWC, two key aspects are largely neglected. The first aspect is that some biological and environmental factors affecting conflict (e.g. season or livestock species attacked) should be considered in regards to the research aim at the outset of modeling. The second aspect is that the background extent should be carefully set in order to draw appropriate inferences. I demonstrate the importance of these considerations by using Maxent, a distribution modeling tool, and livestock depredation data from northern Tanzania, to conduct spatial risk modeling. I applied lessons from Chapters 4 and 5 regarding the appropriate selection of background extent and generation of pseudo-absence points. I designed five comparisons to document if modeling output changes when considering these factors. Four comparisons evaluated ecological factors (event location, season, predator involved, and livestock species attacked) while one investigated changes resulting from different background extents. Contrasting model results between the different comparisons revealed that significant and substantial changes in model output could occur. Both key aspects, consideration of ecological factors before modeling and manipulating the background extent, affected model output, although output varied more substantially in three of the comparisons, location, season, and livestock. This example applied concepts from the larger SDM field to spatial risk modeling, and demonstrated their impact on model results. Addressing these issues can lead to more informative spatial risk models. This chapter addresses research aim IV and provides an improved framework for conducting spatial risk modeling.

## 6.1 Introduction

Large carnivores are found outside protected areas (PAs) in multiple-use landscapes worldwide. In these settings, carnivores frequently interact with people and domestic animals, and generate human-wildlife conflict (HWC) (Section 1.2). The type of interactions carnivores have with people can influence how people feel, and act towards them (Dickman 2010; Hazzah et al. 2013; Dickman and Hazzah 2016). Problematically, carnivores frequently attack and kill livestock where carnivores and livestock overlap. Livestock losses to carnivores can be economically and culturally significant, leading to retaliatory or pre-emptive killing of carnivores (Ogada et al. 2003; Madden 2004; Kissui 2008). Indeed, a leading threat to large carnivores is direct killing by humans (Section 1.4.4). A growing literature is documenting this worldwide problem, and attempting to identify which carnivores are responsible, what livestock are susceptible, what factors predispose livestock to attack, and what actions are effective in preventing loss (Inskip and Zimmermann 2009; Dickman 2010).

Researchers have increasingly applied correlative models developed from the rapidly advancing species distribution modeling (SDM) literature to HWC events. These models can help decipher the spatial pattern of risk to/from wildlife and the predictors that influence it. The rationale is that once factors precipitating high-risk of attack in space and time are identified, these locations can be avoided or the habitat manipulated to reduce the likelihood of conflict (Miller 2015). SDM has been used to model HWC events such as livestock depredations (Treves et al. 2004; Wilson et al. 2006; Abade, Macdonald, and Dickman 2014), wildlife attacks on humans (Kushnir et al. 2014), crop raiding (Ficetola et al. 2014), interactions with specific species such as elephant (Gubbi 2012), and wildlife poisoning (Mateo-Tomás et al. 2012; Márquez et al. 2013). Gathering the event or incident data can be onerous, although once collected, the ready availability of spatial data, geographic information systems (GIS) and modeling tools like Maxent (Phillips, Anderson, and Schapire 2006), make modeling and the creation of spatial risk maps relatively (and deceptively) easy.

However, researchers must consider the importance of different modeling parameters and the ecology implicit in the events at the outset of modeling. Failing to do so risks model results that may be uninformative or incorrect, and lead to the misallocation of precious conservation resources. I focus on two important aspects that are discussed in the larger SDM literature but largely overlooked in spatial risk modeling. The first is the consideration of various biological or environmental characteristics of the HWC event data in regards to how they may impact the research aim. The second is the background extent from which pseudo-absence (PsA) data are drawn.

A challenge to deriving informative and accurate spatial risk models has been that biological and environmental factors that potentially interact with depredation events are often overlooked. Within the larger SDM field, researchers have noted that gender (Conde et al. 2010), season (Takahata et al. 2014), activity (Guisan and Thuiller 2005; Roever et al. 2013), and if the individual is resident or dispersing (Elliot et al. 2014a; Jackson et al. 2016) can influence habitat modeling outcomes for large carnivores. For instance, Conde et al. (2010) demonstrate that single-sex models identify differences in habitat selection between the two sexes, but that selection is blurred if both sexes are pooled into a single model. In the context of spatial risk modeling, some of these same factors, as well as others, may influence the chance that carnivores and livestock interact, and that a livestock is attacked given the interaction (Table 6.1). For instance, carnivore species have a range of behavioral and life history traits that shape how they respond to human presence (Oriol-Cotterill et al. 2015b), and previous research has noted there are idiosyncrasies with which predators attack livestock (Ogada et al. 2003; Kissui 2008). Therefore, lumping all predator species together in a single depredation risk model may blur or misidentify risk factors or locations (but see Miller, Jhala, Jena 2015). Thus, a number of biological and environmental factors should be considered if they impact the research aim at the outset of modeling. If the factor is not explicitly considered, the analysis may not accurately model risk.

Secondly, the background extent from which PsA data are drawn is largely overlooked in HWC spatial risk modeling (but see Ficetola et al. 2014). This is one of the most important considerations in SDM (Phillips et al. 2009). VanDerWal et al. (2009) demonstrated how modeling outputs for 12 Australian species change as the background extent became larger. As shown in Chapter 4, manipulating the background extent and hence where PsA data are drawn from alters the modeling question and affects the ecological inferences that can be derived from the model. This issue deserves more explicit attention in spatial risk modeling. When modeling depredation risk, the background should represent areas where livestock and carnivores could have interacted. Clearly, both species must be present in order for a depredation to occur. An overly large extent would likely include areas where either one or the other is absent and hence PsA would be drawn from irrelevant locations. This could degrade the analysis and affect the desired modeling contrast between locations where depredations occurred and where they could have. Therefore, the background extent should be tightly set around the interaction zone of the predator and livestock, with some areas potentially further masked to ensure the right contrast.

The aim of this chapter is to demonstrate how these two issues, setting the background extent and considering some biological and environmental factors prior to modeling, can affect results and conclusions drawn from depredation risk modeling. To do this, I use Maxent, a widely-used SDM tool (Elith et al. 2006; Ahmed et al. 2015), to identify predictors of carnivore attacks on livestock in northern Tanzania and to predict the spatial distribution of risk. I design five comparisons and demonstrate how modeling outputs change after adjusting either background extent, or factors affecting depredation. Accounting for these aspects can improve risk modeling practices in wildlife conservation applications.

**Table 6.1 Biological and environmental factors potentially interacting with depredation risk; items in bold are included in this analysis. The other options were not included as either not all events had this data, or additional sub-setting would have resulted in very small sample size. The 5th comparison (Background) is not included here as it is not a factor (see section 6.2.3).**

| Factor | Reason |
|---|---|
| **Location of event (Comparison 1 – boma or pasture)** | The location and activity of a livestock in the environment change over time and alters its susceptibility to attack. Livestock may be corralled at night and herded during the day to graze. These activities are spatially dependent. The Maasai use herders to direct livestock to certain locations throughout the day for grazing, drinking and other activities (Homewood and Rodgers 1991). The Maasai locate a corral (or boma) for multiple reasons, broadly speaking by the availability of water and grazing resources (Western and Dunne 1979; Homewood and Rodgers 1991). At the local level, steep slopes and dense vegetation is avoided, and other factors such as soil and drainage may also be important (Western and Dunne 1979; Homewood and Rodgers 1991). Hence each activity likely has similar and particular environmental attributes. |
| **Predator (Comparison 2 – spotted hyena, leopard, lion)** | Carnivores have unique ecology and behavior that governs who, what, where, when and how they hunt. For instance, while a cheetah is usually diurnal, lions are nocturnal and thus their preferred behavior and habitat during the day, while livestock are in the landscape, may influence the frequency and location of predator-livestock encounters. Also, different predators are more or less prone to attacking corrals (Lichtenfeld, Trout, and Kisimir 2014) and in particular manners (e.g. going under, over or through fences). |

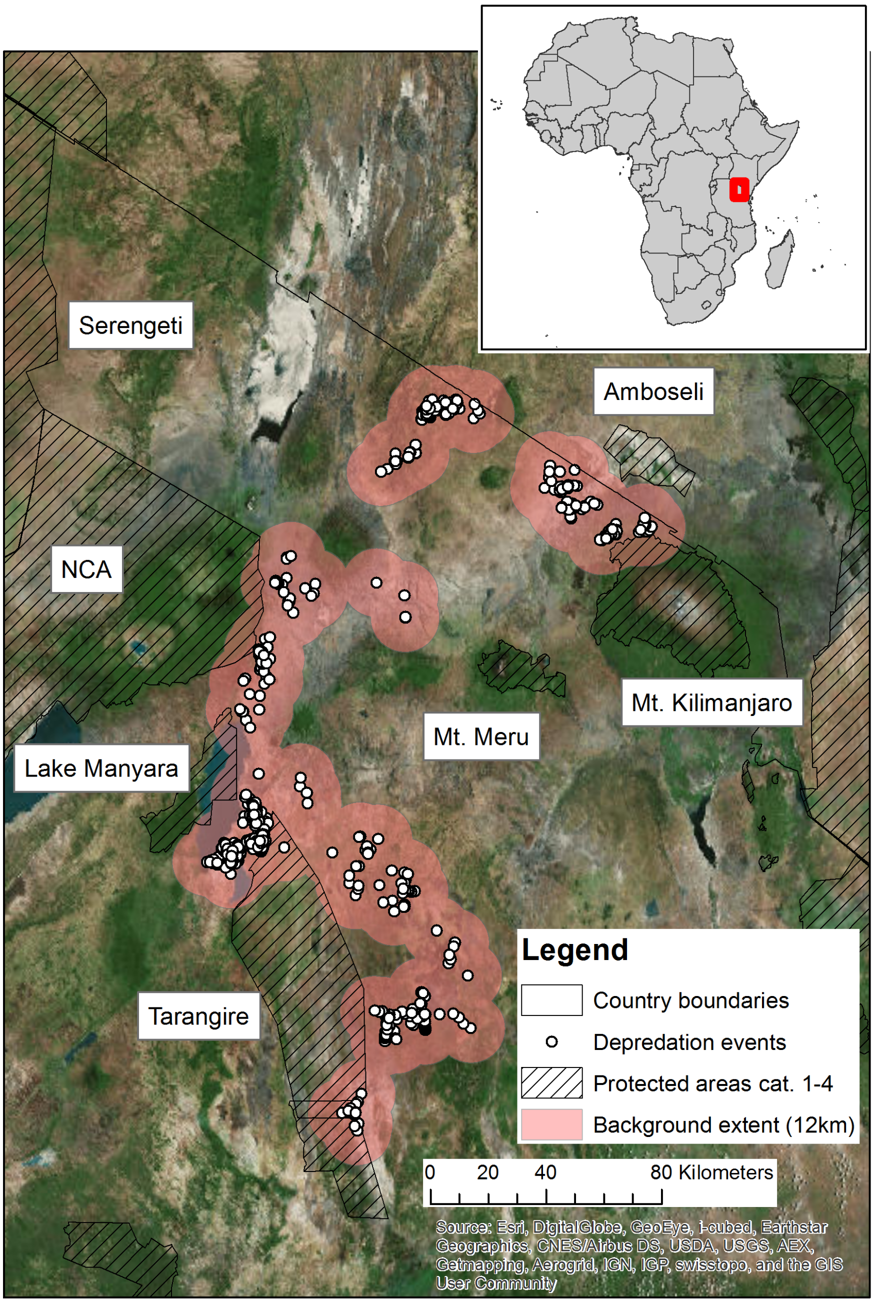
| | |
|---|---|
| **Season (Comparison 3 – dry, wet)** | A number of seasonally related factors may potentially interact to influence depredation. Maasai use different grazing areas at certain times of year (Homewood and Rodgers 1991; Butt 2010). Denser vegetation in the rainy season may make prey harder to catch, making livestock more attractive (Chardonnet et al. 2010). Herbivore vigilance and susceptibility to predation also changes throughout the year (e.g. pregnant females or newborns). Both herbivores and carnivores may move in/out of protected areas. Wild herbivore locations and densities fluctuate during the year (e.g. migratory species). Fewer or less available wild prey can impact carnivore foraging choices, including depredation (Valeix et al. 2012; Davidson et al. 2013; Lichtenfeld, Trout, and Kisimir 2014; Khorozyan et al. 2015). |
| **Livestock (Comparison 4 – cattle, shoat)** | Carnivores prefer wild prey of a certain size range (Hayward and Kerley 2008). This is likely true for livestock as well (Kissui 2008). The Maasai use different pastures for cattle, calves, and shoats (sheep and goats) to account for different nutritional needs and the distances that they can move away from the boma (Homewood and Rodgers 1991). Herder age and experience can vary between livestock type (Homewood and Rodgers 1991) which may impact susceptibility to predation. |
| Climate variability | Extreme weather events like drought influence the location of livestock (Homewood and Rodgers 1991; Goldman 2011), as well as herbivores and carnivores (Tuqa et al. 2014). Reduction in prey densities (i.e. from drought) can cause predators to roam outside protected areas searching for alternative prey (Chardonnet et al. 2010). |
| Individual identity and sex | Certain factors may influence the behavior and prey selection of individual predators (Pettorelli et al. 2011) such as: if the predator is resident or transient (Chardonnet et al. 2010), body condition (Bauer and Iongh 2005; Chardonnet et al. 2010), the length of tenure in an area, learned behavior, or individual proclivities (Linnell et al. 1999; Chardonnet et al. 2010), gender and reproductive status (Woodroffe et al. 2005; Cooper, Pettorelli, and Durant 2007; Conde et al. 2010; Loarie, Tambling, and Asner 2013), etc. |
| Time of day (or activity of livestock herd) | The Maasai use herders to direct livestock to certain locations throughout the day for grazing, drinking and other activities (Homewood and Rodgers 1991; Butt 2010). In addition, livestock may be "in transit" going to or from grazing areas, water etc. which can influence herd speed and spacing between individuals (i.e. if the herd is more tightly clumped or dispersed). Herder vigilance may also change throughout the day (i.e. alert in the morning/tired and hungry in the evening). |
| Conflict interventions | Actions taken to reduce HWC may alter predator behavior and influence (de)predation risk in a localized area (Lichtenfeld, Trout, and Kisimir 2014). |

## 6.2 Methods

### 6.2.1 Study area

This study was carried out in northern Tanzania, from Tarangire National Park (NP) north to the Kenyan border and from the Ngorongoro Conservation Area (NCA) east to Kilimanjaro NP (Figure 6.1). Active and extinct volcanoes pockmark the region, causing large isolated mountains including the continent's highest and 5th highest mountains (Mount Kilimanjaro and Meru), to rise out of the semi-arid plains. The Great Rift Valley, part of the East African Rift zone, runs north-south along the western edge of the study area. Historic volcanic activity distributed lava flows and ash across the region and, when combined with higher rainfall, leads to areas of high productivity and/or cultivation. Most rangeland soils in contrast are relatively low in fertility (Pratt and Gwynne 1977). The majority of the study area receives between 500 – 1,000 mm of rain. (Pratt and Gwynne 1977). The area is in a transition zone between a single (starting about November) and double rainy season with a long dry period beginning in May or June until the rains begin in November (Pratt and Gwynne 1977). Rainfall is highly irregular on a year-to-year basis with periodic droughts (Pratt and Gwynne 1977; Gelorini and Verschuren 2012).

Two land uses dominate the region, pastoralism and wildlife protection, with limited cultivation. The Maasai are the predominant pastoral tribe in the region (Fratkin 2001). They live in family groups, herding primarily cattle, sheep, and goats (small stock or collectively, shoats) (Homewood and Rodgers 1991). Donkeys are also grazed in small numbers alongside other livestock. Livestock are actively grazed during the day and corralled at night within a homestead (*nkang*, or in Swahili *boma*). Herds are split, with adult cattle herded separately from calves and shoats (Homewood and Rodgers 1991). Older and more experienced individuals herd the more valuable cattle. Cattle also travel further from the boma than other livestock units (Homewood and Rodgers 1991). Dedicated and separate grazing areas are reserved for cattle, calves, and shoats, and vary by season (Homewood and Rodgers 1991). Bomas are typically constructed by piling up thorny branches gathered from the surrounding area (Homewood and Rodgers 1991; Lichtenfeld, Trout, and Kisimir 2014). They tend to be a circular shape with many divided "rooms" to separate livestock and people (Figure 6.2). The thorn branches used for the walls and gate disintegrate over time, leading to holes and other weaknesses that heighten predation success (Lichtenfeld, Trout, and Kisimir 2014). Historically, the Maasai were transhumant with routes cutting across political boundaries, however, mobility is increasingly restricted by protected areas (where grazing is restricted), cultivation and privatization of land (Homewood and Rodgers 1991; Goldman 2011).



Source: Esri, DigitalGlobe, GeoEye, i-cubed, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo, and the GIS User Community

**Figure 6.1** **Study area and extent of collected livestock depredation events by Tanzania People & Wildlife, an NGO, in Northern Tanzania. The primary protected areas in the region are labeled. NCA – Ngorongoro Conservation Area**

**Figure 6.2 A typical Maasai boma as seen from the air. Photo courtesy of Tanzania People & Wildlife.**

The PA network of northern Tanzania is globally known, contributes importantly to the national economy (Prins 1987; Lindsey, Roulet, and Romanach 2007) and contains important populations of threatened wildlife (Fund 2013; Riggio et al. 2013; Chase et al. 2016). PAs were primarily designated around permanent water sources and contain impressive congregations of large mammals (although they tend to spill out into unprotected buffer lands during rainy seasons). The NPs exclude human settlement, grazing, hunting, and resource extraction while the NCA allows settlements, grazing and limited resource extraction. Other PAs in the region, such as forest reserves or wildlife management areas, are less stringently protected and more lenient in allowing a range of human activities.

All members of the large carnivore guild can cause HWC, including the lion, leopard, cheetah, African wild dog, spotted hyena, and striped hyena (*Crocuta crocuta* and *Hyaena hyaena* respectively). Smaller carnivores, particularly jackals, can also cause HWC (Gusset et al. 2009). The spotted hyena is more likely a threat to livestock in the region than the striped (Ogada et al. 2003; Kolowski and Holekamp 2006; Lyamuya et al. 2014), which is primarily a scavenger (Mills and Hofer 1998). However, for both jackal and hyena, respondents did not reliably identify the individual species and they were labeled generically as jackal or hyena. While herbivores occasionally kill livestock (Western and Dunne 1979), these occurrences were not included. Hyena, leopard and lion were the most common predators of livestock and are the focus of the analysis. Spotted hyena and leopard are present throughout the analysis extent (see below) (Bohm and Honer 2015; Jacobson et al. 2016), while lion range (both permanent and temporary) is ~12,450 km$^2$ (89%) of analysis extent (Amy Dickman et al. unpublished data).

### 6.2.2 Data collection

Depredation monitoring began in 2003 in the community of Loibor Siret, just east of Tarangire NP, the local base for Tanzania People & Wildlife (TPW), a non-governmental organization. Over time, TPW has expanded its reach to increase coverage of northern Tanzania. As of 2016, TPW employed roughly 30 trained conflict officers in 20 communities. The officers are local Maasai men each responsible for covering one community/village. Officers are typically informed of a wildlife attack on livestock by phone call. The conflict officers then go to meet the livestock owner (and/or herder) to collect a detailed verbal record of the depredation event, a GPS location, and photographs of the incident (when possible). They respond to requests within a day (at most 3 days) and use field and other corroborative evidence to assess the cause of the incident and veracity of the herder report. If the attack occurred at a boma, a GPS point is taken at the boma, but if the attack occurred outside the boma, a GPS point is taken where the livestock was found. In some cases this can be different from the location of the kill as predators, such as leopards, sometimes drag prey elsewhere to be eaten (Sunquist and Sunquist 2002). Incidents are recorded in a standardized database and given a reliability score by the officer. The reliability score is a subjective attempt to rate events based on the veracity of the account. Occasionally livestock are found dead from unknown causes and/or the herder may misrepresent accounts of the death. There is no financial compensation for killed livestock and no obvious reason to misrepresent losses. Not all losses are reported. Only the most recent data was used in this analysis, a ~ 3 year period from 2013-2016, to minimize potential variation in record keeping practices between different officers. The length of record keeping varied between communities with officers operating in some areas for less than a year.

Officers collected information on a variety of characteristics related to the attack. This included date, time of day, season, weather, predator involved, livestock and herd information, location, and if the livestock was lost prior to attack. Season was split into rainy (November through April) and dry seasons. Predator was identified as cheetah, hyena, jackal, leopard, lion or wild dog. Livestock was identified as cattle, donkey, sheep, or goat. Location was split into boma or pasture events.

I screened the data, and all duplicates and events with the lowest reliability score were removed. Two conflict events at bomas were several km inside PA boundaries and removed as this location is unlikely to be accurate. Only events with spatial information were included in the analysis.

### 6.2.3 Ecological predictors

Candidate predictors were identified using a combination of ecological knowledge, prior use in similar studies, and data availability (Appendix Figures 6A.1-3). Predictors were split into two categories, bioclimatic, and human disturbance (Table 6.2). Predictors were tested for correlation and only one was retained if correlation was > 0.7 (Dormann et al. 2013). All data sets were converted to raster format and projected into the WGS 1984 Africa Albers Equal Area Conic coordinate system. Data were processed using ArcGIS 10.2 (ESRI 2014) at 250 m resolution using a snap raster to align all grids.

Bioclimatic features include distance to water, Terrain Ruggedness Index (TRI) (Riley, DeGloria, and Elliott 1999), percent tree cover, and dry and wet season NDVI. Surface water is an important determinant of predator and prey distribution in semi-arid savannah (Valeix et al. 2009). Lions in particular are known to preferentially hunt in proximity to water holes and river confluences (Hopcraft, Sinclair, and Packer 2005; de Boer et al. 2010; Davidson et al. 2013). Distance to water was calculated using surface water from the regional land cover dataset Africover (Alinovi, Di Gregorio, and Latham 2000) and WWF Hydrosheds (Lehner, Verdin, and Jarvis 2006). Lions (Hopcraft, Sinclair, and Packer 2005) and leopards (Balme, Hunter, and Slotow 2007) are known to select habitat for prey success rather than prey abundance. Areas of topographical complexity, like erosion embankments, can improve predation success. Elevation data from the USGS Shuttle Radar Topography Mission (Jarvis et al. 2008) were converted into a measure of the topographic variation in an area, or TRI. Slope was also tested but not included, as it was highly correlated with TRI and less important in preliminary Maxent models that included both variables. Vegetation abundance and density can also impact prey accessibility and hunting preferences (Hopcraft, Sinclair, and Packer 2005; Balme, Hunter, and Slotow 2007; Loarie, Tambling, and Asner 2013). All seven bands of a MODIS image were used to construct a proportional estimate of woody vegetation cover for 2010 (percent tree cover from the vegetation continuous field collection) (Townshend et al. 2011). While no data were available on prey distribution, the Normalized Difference Vegetation Index (NDVI) is directly correlated to vegetation productivity and indirectly correlated to forage availability for herbivores (Pettorelli 2013). Given that NDVI changes were shown to correlate with herbivore abundance (Pettorelli et al. 2009), it was used as a proxy for herbivore distribution (Pettorelli et al. 2005; Pettorelli et al. 2009; Swanepoel et al. 2013). Both wet and dry season averages from the years 2008-2012 were used (see Chapter 4).

**Table 6.2 Predictor variables used in the modeling approach. The grain of the analysis is 250 m. A local difference data set was also created for all variables except the 'distance to' variables.**

| Variable (category) | Variable (specific) | Source | Mean; std dev | Range | Original resolution |
|---|---|---|---|---|---|
| Bioclimatic | April (wet season) NDVI averaged from 2008-2012 | MODIS NDVI | 5,336; 1,556 | 9,820 – -2,000 | 250 m |
| | August (dry season) NDVI averaged from 2008-2012 | MODIS NDVI | 2,829; 1,355 | 8,592 – -1,977 | 250 m |
| | Elevation | SRTM | 1353, 406 | 4905 - 703 | 250 m |
| | Terrain Ruggedness Index | SRTM | 260, 199 | 1,712 - 0 | 250 m |
| | Distance to river and lake (m) | WWF hydrosheds & Africover (water) | 1,709; 1,420 | 9,397 - 0 | NA |
| | Percent tree cover 2010 | MODIS VCF | 8; 15 | 86 - 0 | 250 |
| Human disturbance | Human population density in 2015 (people/km$^2$) | Landscan | 21; 72 | 5,588 - 0 | 30 seconds |
| | Distance to protected area (m) | WDPA | 20,902; 19,580 | 82,177 - 0 | NA |
| | Distance to roads (m) | gROADS v1 | 7,515; 7,260 | 35,501 - 0 | NA |

Human disturbance features included distance to roads and protected areas, and human population density (HPD). Livestock are occasionally herded along roads as easier routes although human activity and disturbance is also greater along roadways. These factors can lead carnivores into a kind of push-pull relationship with roads, access to prey yet avoidance of people (Rogala et al. 2011). After review in Google Earth, the gROADS data set (CIESIN 2013) appeared best for the study area. Tanzania has a well-developed network of PAs that include many different classifications (UNEP and IUCN 2015);

however, these variably protect wildlife (Stoner et al. 2007a) or prevent encroachment (Jason Riggio et al. unpublished data). Thus, only distance to PA with the highest protection categories (NPs and the NCA) was calculated. This variable was included as livestock grazing is excluded in PAs (except the NCA), and wildlife move in/out of PAs over the course of a year. Human population density is positively associated with carnivore extinction risk (Cardillo et al. 2008) and human presence affects the behavior and distribution of carnivores (Rogala et al. 2011; Schuette, Creel, and Christianson 2013; Oriol-Cotterill, et al. 2015a, Chapter 5). Naturally, also, there are no livestock where people are absent. In order to determine the most appropriate population density data set for the study area, I compared WorldPop (Linard et al. 2012) and Landscan (Bright et al. 2015). WorldPop appeared to spread population densities across political areas creating discontinuities between regions. For this reason Landscan was deemed the better data set for the study area. The data represent the ambient 2015 population count per grid cell. Distance to settlement, distance to boma, and livestock densities were also considered as additional variables and excluded. Not all bomas were mapped, only those where conflict was reported had a GPS point. Not all settlements were recorded either. Local livestock densities were not known and global livestock data sets (e.g. Gridded Livestock of the World) were too coarse.

For all predictors except the 'distance-to' variables, the difference of a particular cell from a local mean was calculated. As many predators, including lion (Hopcraft, Sinclair, and Packer 2005) and leopard (Balme, Hunter, and Slotow 2007), prefer areas of high prey accessibility over high prey density, it was crucial to attempt to identify areas of high prey accessibility. Areas of increased prey accessibility include features such as dense vegetation or steep slopes (Laundré, Hernández, and Altendorf 2001; Laundré, Hernández, and Ripple 2010). To calculate the local difference value, the mean value within a radius of 1.5 km was subtracted from the actual value of the focal cell. This distance represents the diameter of roughly half the daily distance moved by the lion (Elliot et al. 2014b). The leopard and hyena likely move farther than this. Thus, this process produced a data set that represented areas of relatively higher or lower values than the surrounding accessible area. If the focal cell was a dense thicket surrounded by open savanna, or the edge of an escarpment, this process would identify these areas as being different from the local mean.

### 6.2.4 Maxent

The widely-used SDM tool Maxent identified relationships between environmental predictors and depredation locations (Elith et al. 2006; Merow, Smith, and Silander 2013; Ahmed et al. 2015). Maxent was chosen as it performs well even with few training data (Phillips and Dudik 2008) and for its flexible model parameters (see below). Depredation points are used in place of the more typical presence data in SDM. Maxent is a machine learning (i.e. data-driven) method that minimizes the relative entropy in multivariate space between two probability densities, one derived from the background and the other provided by the presence data (Phillips, Anderson, and Schapire 2006; Elith et al. 2011). Using 10,000 background points (i.e. the pseudo-absence data), and combined with the presence data, Maxent generates a probability density distribution of the background environmental conditions for each variable. The pseudo-absence points are drawn randomly from the entire study area to sample the entire background. The probability density distribution of the presence data is then compared to the background conditions to ascertain species' response. In so doing, Maxent identifies species response curves, predictor importance scores, and calculates a relative occurrence rate for each cell i.e. generating a spatial prediction of relative likelihood.

Maxent default parameters may be inappropriate in some situations, particularly when ecological interpretation and biologically appropriate responses are key (Merow, Smith, and Silander 2013). Features are the different classes of response types that Maxent can use to fit response curves to the presence data. As most ecological relationships are unimodal (Austin 2007), only linear and quadratic features were used. Features allowing more complex fits to the data (i.e. product, threshold, and hinge) were not used. The number of iterations was increased from the default 500 to 5,000 as models were occasionally not converging after the default.

Maxent version 3.3.3 was run via the DISMO package (Hijmans et al. 2016) in R Studio (R 2016). Each model was repeated 100 times, starting with a different random seed each time. I selected 100 iterations as a balance between potential model variation and computation time. Predictor importance, response curves and spatial predictions were all averaged across the 100 runs. Other parameters were kept at default values.

### 6.2.5 Analysis

Five comparisons were designed to investigate how modeling outputs change given consideration of different biological and environmental factors, and various background extents. They are:

- Comparison 1 investigated the location of attack by comparing events at boma and at pasture. Comparison 1A uses all events, only splitting by location. Another comparison (1B), uses a more specific subset of the data, only including hyena attacks in the wet season on cattle;
- Comparison 2 investigated differences due to the predator, by contrasting hyena, leopard and lion attacks at pasture in the wet season on all livestock species;
- Comparison 3 investigated differences due to the season by contrasting hyena attacks on shoats at pasture in the wet and dry seasons;
- Comparison 4 investigated differences due to the livestock species by contrasting hyena attacks in the wet season at pasture on cattle and shoats; and
- Comparison 5 investigated differences due to the background extent by contrasting lion attacks on cattle in the wet season at three different extents.

All comparisons first analyzed the full data set before analyzing the two (or three) data subsets. For instance, in Comparison 1A, the first model was run using all events as training data. Then, two additional models were run, using the two data subsets as training data, i.e. only attacks at boma or at pasture. In sum, Comparison 1A used three different training data sets, all events, all events at boma, and all events at pasture. Differences were evaluated between the events at boma and at pasture (see below). The only exception to this was Comparison 5 because the event data were not subset; the same data were analyzed using different backgrounds. The data subsets used in each comparison were selected to isolate the criteria of interest and keep the other variables constant while also maximizing the number of available data points.

All comparisons (except Comparison 5) were conducted within a 12 km buffer set around all boma events. Thus defined, the analysis extent is 13,938 km$^2$. The average transit distance (or grazing orbit) for Maasai cattle in East Africa is 12-15 km per day (Bekure et al. 1991), suggesting a radius from the boma of about 7 km. A 12 km buffer is thus slightly larger than this radius and captured all pasture events (the furthest from a conflict boma was 10.5 km). A single buffer set from boma locations was used under the assumption that it included all (or at least the vast majority) of livestock grazing areas in the communities covered by TPW since actual livestock movements, grazing areas and all used (and not attacked) boma locations are unknown. A second important assumption is that conflict officers were capable of reaching all parts of the analysis extent.

Comparison 5 compared three different extents: a 12 km buffer with areas masked out where lion and cattle were unlikely to spatially overlap (“clipped”); within a 25 km buffer of all events (“expanded”); and within a rectangular buffer 140 km x 160 km (“full”). The larger buffer sizes were chosen to represent a slightly and substantially exaggerated buffer size respectively and have no particular ecological justification. The “expanded” buffer is 28,517 km$^2$ and “full” buffer is ~275,000 km$^2$. The “clipped” product started with the 12 km buffer and, in order to restrict the extent to regions where both cattle and lions could occur, protected areas (IUCN categories 1-4) (1,431 km$^2$) and areas outside of permanent or temporary lion range were removed . The “clipped” buffer is 11,274 km$^2$ (or 81% of original). As distribution maps suggest hyenas and leopards are present throughout the analysis extent, no areas were masked in Comparisons 1-4.

### 6.2.6 Evaluating potential differences in output

Maxent produces spatial predictions of risk, potential environmental niches, predictor importance scores, response curves, and evaluation metrics. These outcomes, except response curves, were contrasted to determine potential differences in outcome due to the specified comparison for each of the five comparisons. To determine if the comparisons resulted in differences in spatial predictions, Pearson correlation coefficients from the 100 replicate runs (i.e. the internal consistency of the model) were compared with correlation coefficients between the averaged spatial predictions. To determine if the comparisons resulted in differences in environmental niche, the Identity Test, developed in Warren et al. (2008), investigated SDM niche similarity (see below). Training and testing AUC values were used as evaluation metrics with test percentage set at 10% using subsampling. I used two-tailed two sample t-tests with unequal variance to determine if predictor importance and AUC values were different between the data subsets. Since I had 100 replicates for every model, I could use the t-tests to see if predictor importance or AUC values differed significantly in each comparison. I used an alpha value of 0.05 to determine statistical significance. I reviewed response curves of the three most important variables in Comparison 1A to highlight potential changes in response curves (Appendix 6A). In Comparison 5, there was no I test as the event data were the same for the three models, and no testing AUC value as this represented only one data point.

The Identity Test, calculated via ENMTools in R, revealed whether the species distribution models produced by two populations were similar (Warren, Glor, and Turelli 2010). The test compared niche similarity in environmental space produced by the two models (the observed I value) with a null distribution of niche differences. To produce the null distribution, the test pools the presence (depredation) data from the two models, splits the data back into two samples the same size as the originals, and then generates two new distribution models from the sample. It then calculates a null I value for this randomized sample. This was repeated 100 times to produce a null distribution of similarity values. The observed value of I is then compared to the null distribution of test I values to see if it was significantly lower than the similarity values expected from the pseudo-replicated data sets. While the Pearson correlation coefficients were calculated on the spatial predictions (geographic space), I was calculated on the environmental space of the model.

## 6.3 Results

TPW collected a total of 1,019 depredation events in northern Tanzania from August 2013 to July 2016 (Figure 6.3). Roughly two thirds of attacks occurred at the boma as opposed to at pasture, and two thirds occurred during the rainy season. Hyenas perpetrated roughly three quarters of attacks while leopards, lions, cheetahs, jackals, and wild dogs were also responsible in decreasing order of importance. Shoats suffered over three quarters of attacks, many more than either cattle or donkeys. Of the 973 depredation events after data screening, 905 had spatial information and were included in the analysis (Figure 6.4).

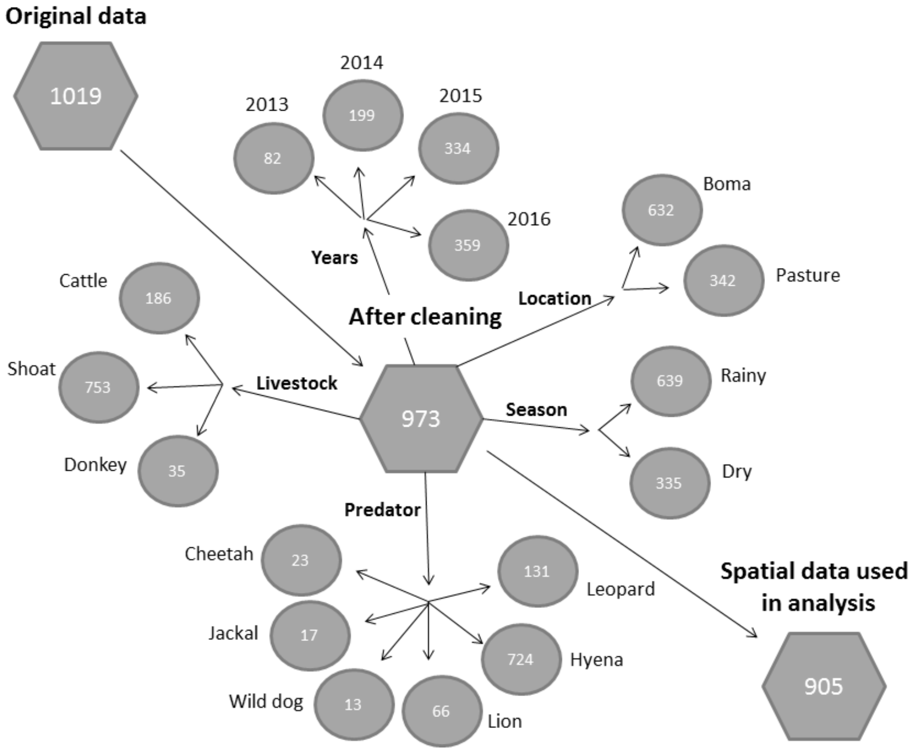


**Figure 6.3** **Summary of livestock depredation data. After data cleaning and removal of duplicates, 973 depredation points remained of the original 1,019. Not all these incidents had spatial information, so only 905 were used in the analysis.**

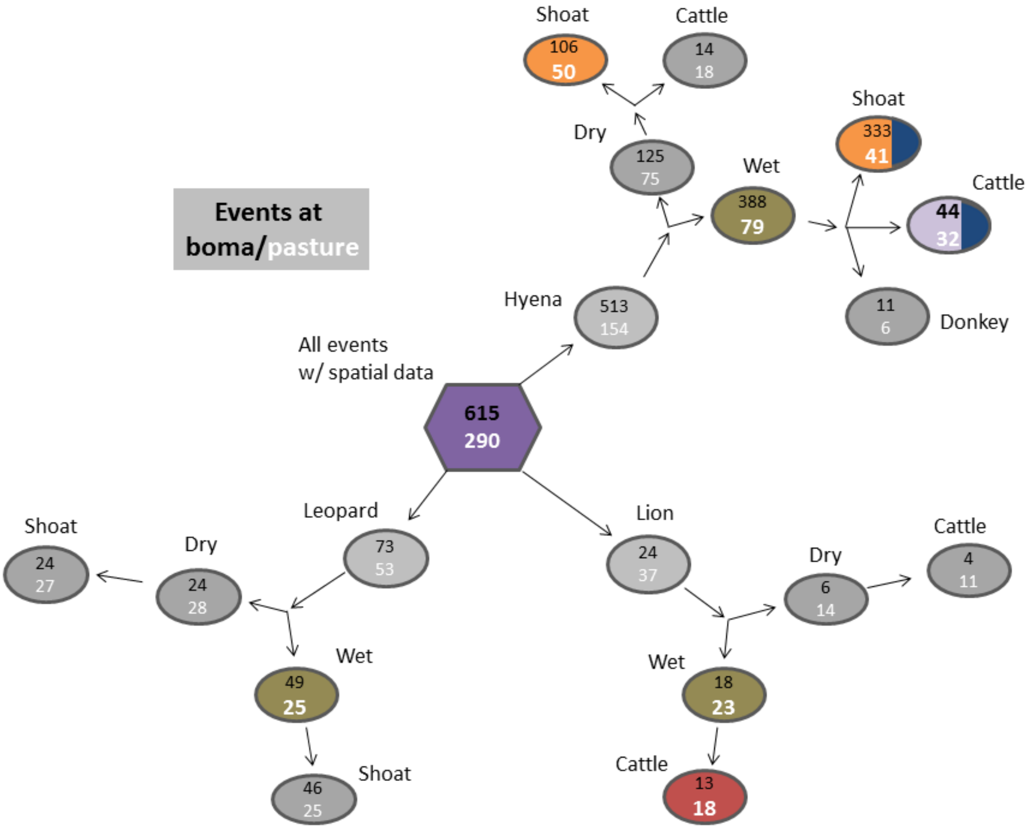


**Figure 6.4 Summary of livestock depredation data used in the analysis*. All comparisons use only pasture events other than the 1st comparison. Bold values indicate the number of events used in the analyses. Comparison 1A and 1B investigated the location of attack, and are indicated by purple and lilac colors; Comparison 2 investigated the predator involved in the attack, and is indicated by brown; Comparison 3 investigated the season of attack, and is indicated by orange; Comparison 4 investigated the livestock involved in the attack, and is indicated by blue; and Comparison 5 investigated the analysis extent, and is indicated by red. Note the number of events is different than the number of training samples as numbers here represent total number of events recorded by TPW of that type whereas the training samples exclude data points that are used for testing.**

*Note: This Figure should be read starting at the center and working outwards. For instance, in Comparison 3, indicated by orange ovals, hyena attacks in the dry season on shoats at pasture (50 events) are compared with hyena attacks in the wet season on shoats at pasture (41 events).

Across the five comparisons, there could be substantial and significant variation in variable importance, spatial prediction of risk, environmental niche, and evaluation metrics. Variation in response curves was only reviewed visually, an example is shown in Appendix Figure 6A.4.

Nearly every predictor had a significantly different level of importance in each of the five comparisons (Table 6.3). In each comparison, every predictor involved in the model was compared between the two (or three) subsets using a two-sample t-test. There were 120 variable pairs compared this way, only 11 were found to have non-significant differences, the rest were all significantly different from each other. In some cases, the dominant variables could shift completely within the comparison. For instance, in Comparison 5, evaluating the effect of the background, distance to road was dominant when the background was clipped but became much less important at larger buffer sizes. The opposite pattern occurred for percent tree cover, as it became the most dominant variable at larger extent sizes. In general, Comparison 2 and Comparison 5, evaluating the effects of predator and background respectively, had the most number of non-significant variable pairs (four and six). In contrast, every variable pair was significantly different in Comparisons 1A, 1B and 3.

The difference in spatial output in each comparison was compared via Pearson correlation coefficients (Table 6.4). The variation could be quite substantial. In Comparison 1B, evaluating the effect of the event location, the correlation between the two outputs was only 0.38. The correlations in Comparison 2 (predator) were much higher, varying between 0.68 for the leopard and lion, and 0.79 for the hyena and lion. The correlations in spatial output in Comparison 5 (background) were even higher, varying only from 0.86 – 0.96. The internal correlation coefficients, indicating the similarity of the output across the 100 model replicates, were all very high, between 0.95 and 0.99. This suggests the model output was robust to slight changes in model inputs. The spatial output from Comparison 1A is shown as an example (Figure 6.5). The remaining spatial comparisons are shown in Appendix 6A (Appendix Figures 6A.5-10).

There was also substantial variation when comparing the environmental niche (Table 6.5). Two of the comparisons had significant p-values at the alpha 0.05 level; they were Comparison 1A (location), and Comparison 3 (season). The p-values for the other comparisons varied between 0.21 and 0.66 suggesting the environmental niches between comparisons were relatively more similar.

Nearly all training and test AUC values were significantly different across all five comparisons (Table 6.6). Of the 19 training and test AUC value pairs, only four of them were not significantly different. In Comparison 3 (season), neither training nor test AUC were significantly different. In Comparison 5 (background), the training AUC did not reach statistical significance.

**Table 6.3 Comparison of predictor importance scores and their standard deviation for the five different comparisons. All comparisons of predictor importance between data subsets are statistically significant at alpha = 0.05 except for those marked with an * and ^. These symbols indicate the subset pairs that are not significantly different.**

| Species | Distance to PA | Distance to road | Distance to water | Wet season NDVI | Local diff, Wet season NDVI | Dry season NDVI | Local diff, Dry season NDVI | Human pop'n density |
|---|---|---|---|---|---|---|---|---|
| **Comparison 1A: Location** | | | | | | | | |
| All boma events | 33.87 ± 1.45 | 31.24 ± 2.22 | 4.56 ± 0.49 | 1.05 ± 0.47 | 0.17 ± 0.15 | 0.26 ± 0.11 | 0.04 ± 0.03 | 8.88 ± 2.3 |
| All pasture events | 9.22 ± 1.14 | 52.89 ± 2.52 | 8.7 ± 1.41 | 8.76 ± 1.74 | 1.48 ± 0.6 | 0.69 ± 0.28 | 0.23 ± 0.15 | 0.23 ± 0.3 |
| **Comparison 1B: Location** | | | | | | | | |
| Hyena attacks in wet season on cattle… | | | | | | | | |
| … at boma | 15.5 ± 2.69 | 28.43 ± 3.53 | 8.89 ± 1.83 | 10.55 ± 2.28 | 0.11 ± 0.14 | | | 18.24 ± 3.33 |
| … at pasture | 23.39 ± 6.68 | 6.16 ± 7.85 | 4.46 ± 2.48 | 0.38 ± 0.53 | 8.92 ± 6.29 | | | 5.44 ± 2.58 |
| **Comparison 2: Predator** | | | | | | | | |
| Attacks at pasture in wet season on all livestock by… | | | | | | | | |
| … Only hyena | 4.16 ± 1.41* | 53.01 ± 3.56 | 4.48 ± 1.56 | 0.09 ± 0.14 | 0.08 ± 0.24 | | | 6.98 ± 1.52 |
| … Only leopard | 3 ± 1.04 | 56.49 ± 3.29 | 5.86 ± 1.46 | 7.79 ± 3.03 | 2.28 ± 1.42 | | | 2.01 ± 1.21 |
| … Only lion | 4.66 ± 3.29* | 71.78 ± 5.27 | 1.38 ± 1.13 | 1.89 ± 1.35 | 1.57 ± 0.71 | | | 0.31 ± 0.27 |
| **Comparison 3: Season** | | | | | | | | |
| Hyena attacks at pasture on shoats… | | | | | | | | |
| … in dry season | 4.76 ± 1.57 | 44.23 ± 2.1 | 8.6 ± 1.82 | | | 0.06 ± 0.1 | 0.29 ± 0.42 | 0 ± 0 |
| … in wet season | 0.62 ± 0.32 | 62.3 ± 3.77 | 3.24 ± 1.34 | 0.07 ± 0.19 | 0.1 ± 0.13 | | | 7.32 ± 2.05 |
| **Comparison 4: Livestock** | | | | | | | | |
| Hyena attacks at pasture in wet season… | | | | | | | | |
| … on cattle | 23.39 ± 6.68 | 6.16 ± 7.85 | 4.46 ± 2.48 | 0.38 ± 0.53 | 8.92 ± 6.29 | | | 5.44 ± 2.58 |
| … on shoat | 0.62 ± 0.32 | 62.3 ± 3.77 | 3.24 ± 1.34 | 0.07 ± 0.19 | 0.1 ± 0.13 | | | 7.32 ± 2.05 |
| **Comparison 5: Background** | | | | | | | | |
| Lion attacks at pasture in wet season on cattle… | | | | | | | | |
| … at clipped extent | 5.08 ± 2.81 | 69.02 ± 5.05 | 0.09 ± 0.17* | 1.94 ± 1.32* | 6.14 ± 3.25 | | | 3.38 ± 1.43 |
| … at expanded extent (25 km) | 3.27 ± 1.92 | 58.78 ± 3.52 | 0.46 ± 0.5 | 1.92 ± 0.87* | 3.31 ± 1.38 | | | 0.88 ± 0.19 |
| … at full extent (150 km) | 13.2 ± 2.46 | 21.49 ± 1.62 | 0.05 ± 0.14* | 3.99 ± 1 | 0.13 ± 0.37 | | | 0.01 ± 0.01 |

Table 6.3 continued

| Species | Elevation | Local diff, Elevation | % Tree cover | Local diff, % Tree cover | Terrain Roughness Index | Local diff, TRI |
|---|---|---|---|---|---|---|
| **Comparison 1A: Location** | | | | | | |
| All boma events | 7.27 ± 1.22 | 0.01 ± 0.05 | 10.63 ± 1.01 | 0.58 ± 0.35 | 0.06 ± 0.06 | 0.29 ± 0.18 |
| All pasture events | 2.5 ± 0.62 | 0.28 ± 0.22 | 12.57 ± 1.39 | 0.82 ± 0.44 | 0.31 ± 0.2 | 1.29 ± 0.5 |
| **Comparison 1B: Location** | | | | | | |
| Hyena attacks in wet season on cattle… | | | | | | |
| … at boma | 0.52 ± 0.68 | 2.41 ± 1.18 | 5.22 ± 1.19 | 1.56 ± 0.87 | 0.88 ± 0.61 | 1.11 ± 1.19 |
| … at pasture | 4.63 ± 2.95 | 19.81 ± 3.65 | 14.05 ± 2.58 | 4.01 ± 1.58 | 5.71 ± 2.37 | 2.24 ± 1.78 |
| **Comparison 2: Predator** | | | | | | |
| Attacks at pasture in wet season on all livestock by… | | | | | | |
| … Only hyena | 3.75 ± 1.54 | 3.02 ± 0.96 | 10.13 ± 1.39 | 4.75 ± 1.35 | 0.08 ± 0.25* | 9.38 ± 1.97 |
| … Only leopard | 10.14 ± 1.77 | 2.26 ± 1.31 | 3.78 ± 0.63 | 3.13 ± 0.88 | 0.13 ± 0.21*^ | 3.14 ± 2.33* |
| … Only lion | 0.96 ± 0.71 | 0.38 ± 1 | 12.72 ± 2.04 | 0.32 ± 0.42 | 0.37 ± 1.33^ | 3.51 ± 1.19* |
| **Comparison 3: Season** | | | | | | |
| Hyena attacks at pasture on shoats… | | | | | | |
| … in dry season | 0.76 ± 0.73 | 0.01 ± 0.05 | 35.2 ± 2.7 | 0 ± 0 | 0.44 ± 0.4 | 5.31 ± 1.24 |
| … in wet season | 3.13 ± 1.18 | 1.39 ± 0.41 | 8.02 ± 1.18 | 3.96 ± 1.29 | 0.75 ± 0.76 | 9.01 ± 1.71 |
| **Comparison 4: Livestock** | | | | | | |
| Hyena attacks at pasture in wet season… | | | | | | |
| … on cattle | 4.63 ± 2.95 | 19.81 ± 3.65 | 14.05 ± 2.58 | 4.01 ± 1.58* | 5.71 ± 2.37 | 2.24 ± 1.78 |
| … on shoat | 3.13 ± 1.18 | 1.39 ± 0.41 | 8.02 ± 1.18 | 3.96 ± 1.29* | 0.75 ± 0.76 | 9.01 ± 1.71 |
| **Comparison 5: Background** | | | | | | |
| Lion attacks at pasture in wet season on cattle… | | | | | | |
| … at clipped extent | 1.94 ± 1.07 | 0.24 ± 1.34*^ | 10.1 ± 2.54 | 0.2 ± 0.39 | 0.37 ± 1.25* | 0.61 ± 0.67* |
| … at expanded extent (25 km) | 0.74 ± 0.42 | 0.61 ± 1.59* | 26.35 ± 2.33 | 0.5 ± 0.69 | 0.54 ± 1.63* | 2.13 ± 0.69 |
| … at full extent (150 km) | 5.61 ± 0.76 | 0.02 ± 0.05^ | 54.95 ± 2.7 | 0.05 ± 0.09 | 0.09 ± 0.3 | 0.41 ± 1.15* |

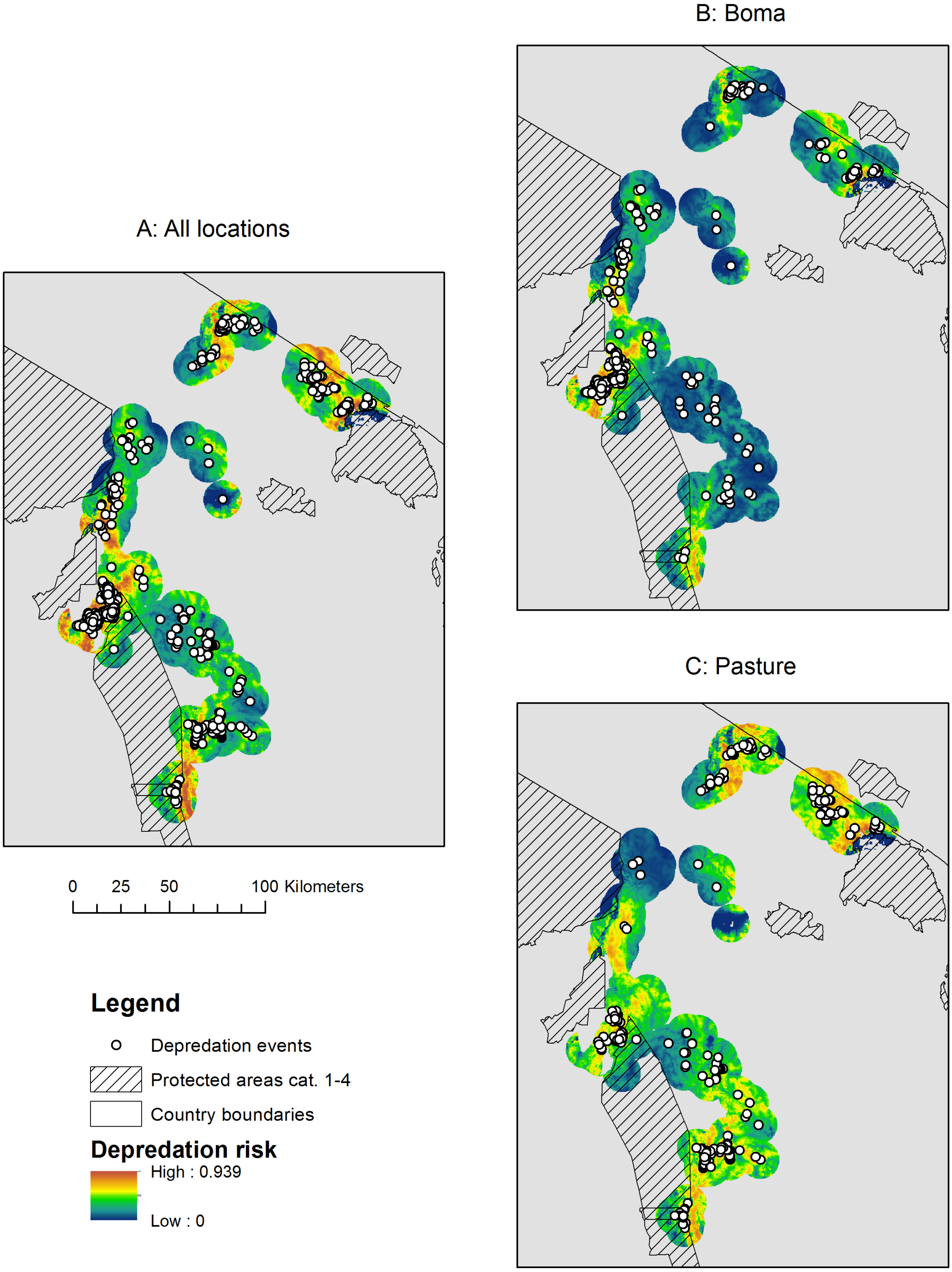


**Figure 6.5 Spatial predictions of depredation risk in Comparison 1A: location. Overall depredation risk using all event data is shown in subfigure A, while the spatial predictions trained using just events at boma or at pasture are in subfigures B and C. These represent the ensemble output, averaged across the 100 model replicates.**

**Table 6.4 Pearson correlation coefficients of internal model consistency and between data subsets for each of the five comparisons. The values on the 1:1 line are the internal correlations of the spatial predictions among 100 replicates. The values below the 1:1 line are the correlations in spatial predictions between the different models.**

| | | | |
|---|---|---|---|
| **Comparison 1A: Location** | All boma | All pasture | |
| All boma events | 0.99 ± 0.00 | | |
| All pasture events | 0.69 | 0.99 ± 0.00 | |
| **Comparison 1B: Location** | … at boma | … at pasture | |
| Hyena attacks in wet season on cattle... | | | |
| … at boma | 0.98 ± 0.02 | | |
| … at pasture | 0.38 | 0.95 ± 0.05 | |
| **Comparison 2: Predator** | Only hyena | Only leopard | Only lion |
| Attacks at pasture in wet season on all livestock by… | | | |
| … Only hyena | 0.98 ± 0.01 | | |
| … Only leopard | 0.7 | 0.99 ± 0.01 | |
| … Only lion | 0.79 | 0.68 | 0.98 ± 0.01 |
| **Comparison 3: Season** | … in dry season | … in wet season | |
| Hyena attacks at pasture on shoats… in both seasons | | | |
| … in dry season | 0.98 ± 0.01 | | |
| … in wet season | 0.62 | 0.98 ± 0.01 | |
| **Comparison 4: Livestock** | ... on cattle | … on shoats | |
| Hyena attacks at pasture in wet season... on all livestock | | | |
| … on cattle | 0.95 ± 0.05 | | |
| ... on shoat | 0.42 | 0.98 ± 0.01 | |
| **Comparison 5: Background** | ..., clip | ..., expanded | ..., full |
| Lion attacks at pasture in wet season on cattle… | | | |
| … at clipped extent | 0.99 ± 0.01 | | |
| … at expanded extent (25 km) | 0.96 | 0.98 ± 0.01 | |
| ... at full extent (150 km) | 0.86 | 0.89 | 0.99 ± 0.01 |

**Table 6.5 P-values of the I test for the five different comparisons. The I test compares the predicted environmental niche between models.**

| | | | |
|---|---|---|---|
| **Comparison 1A: Location** | All boma | All pasture | |
| All boma events | | | |
| All pasture events | 0.02 | | |
| **Comparison 1B: Location** | … at boma | … at pasture | |
| Hyena attacks in wet season on cattle… | | | |
| … at boma | | | |
| … at pasture | 0.66 | | |
| **Comparison 2: Predator**<br>Attacks at pasture in wet season on all livestock by… | … only hyena | … only leopard | … only lion |
| … only hyena | | | |
| … only leopard | 0.42 | | |
| … only lion | 0.25 | 0.35 | |
| **Comparison 3: Season** | … in dry season | … in wet season | |
| Hyena attacks at pasture on shoats… | | | |
| … in dry season | | | |
| … in wet season | 0.03 | | |
| **Comparison 4: Livestock**<br>Hyena attacks at pasture in wet season… | ... on cattle | … on shoats | |
| … on cattle | | | |
| ... on shoat | 0.21 | | |

**Table 6.6 Evaluation metrics and their standard deviation for all models. All comparisons of AUC values between data subsets are statistically significant at alpha = 0.05 except for those marked with an *. This symbol indicates the subset pairs that are not significantly different.**

| Species | Training AUC | Testing AUC |
|---|---|---|
| **Comparison 1A: Location** | | |
| All boma events | 0.84 ± 0.00 | 0.84 ± 0.03 |
| All pasture events | 0.76 ± 0.00 | 0.73 ± 0.04 |
| **Comparison 1B: Location** | | |
| Hyena attacks in wet season on cattle... | | |
| … at boma | 0.87 ± 0.01 | 0.82 ± 0.09 |
| … at pasture | 0.72 ± 0.02 | 0.57 ± 0.16 |

| | | |
|---|---|---|
| **Comparison 2: Predator** | | |
| Attacks at pasture in wet season on all livestock by… | | |
| … Only hyena | 0.79 ± 0.01 | 0.75 ± 0.08* |
| … Only leopard | 0.88 ± 0.01 | 0.8 ± 0.08 |
| … Only lion | 0.85 ± 0.02 | 0.71 ± 0.07* |
| **Comparison 3: Season** | | |
| Hyena attacks at pasture on shoats... | | |
| … in dry season | 0.82 ± 0.01* | 0.76 ± 0.08* |
| … in wet season | 0.82 ± 0.01* | 0.78 ± 0.08* |
| **Comparison 4: Livestock** | | |
| Hyena attacks at pasture in wet season... | | |
| … on cattle | 0.72 ± 0.02 | 0.57 ± 0.16 |
| ... on shoat | 0.82 ± 0.01 | 0.78 ± 0.08 |
| **Comparison 5: Background** | | |
| Lion attacks at pasture in wet season on cattle... | | |
| … at clipped extent | 0.85 ± 0.01 | NA |
| … at expanded extent (25 km) | 0.88 ± 0.01* | NA |
| … at full extent (150 km) | 0.89 ± 0.01* | NA |

## 6.4 Discussion

Species distribution models paired with depredation events can be an important tool to identify factors correlated with attack, the shape of response curves, and produce spatial estimates of risk. Risk modeling can ameliorate HWC if predictors of conflict are understood and those conditions avoided by pastoralists. Yet, to generate good models, researchers should be cognizant of lessons from the larger SDM literature. This chapter focuses on two aspects of distribution modeling that are apparent when considering modeling from an ecological perspective and applies them to spatial risk modeling. These aspects are the consideration of various biological or environmental characteristics of the HWC event data in regards to how they may impact the research aim, and constraining the background extent to where carnivores and livestock interact. Results from the five comparisons demonstrate that modeling output can change when considering these different aspects. In this example, variations in attack location, season, and livestock led to more substantial differences in model output than the predator involved. The size of the background extent can also affect model results, and its importance in drawing the appropriate ecological inference is explained (see below).

### 6.4.1 Livestock depredations in northern Tanzania

Livestock depredations by large carnivores are a widespread and serious conservation issue in Tanzania and Kenya (Ogada et al. 2003; Patterson et al. 2004; Okello 2005; Woodroffe et al. 2005; Kolowski and Holekamp 2006; Kissui 2008; Ikanda and Packer 2008; Kaswamila 2009; Mwakatobe, Nyahongo, and Røskaft 2013; Mponzi, Lepczyk, and Kissui 2014; Okello, Kiringe, and Warinwa 2014). All livestock are vulnerable to attack throughout the year, both in and outside the boma. Similar to other nearby studies, hyenas were the most problematic carnivore in this study (Kissui 2008; Mwakatobe, Nyahongo, and Røskaft 2013; Schuette, Creel, and Christianson 2013; Mponzi, Lepczyk, and Kissui 2014). Depredations can spur retaliatory attacks, particularly against lions (Lichtenfeld 2005; Kissui 2008), and affect the attitude of pastoralists and threaten carnivore coexistence (Dickman 2010; Hazzah et al. 2013; Dickman and Hazzah 2016).

Some potential errors or biases may remain although TPW data collection techniques are as or more rigorous than similar HWC studies (e.g. Kolowski and Holekamp 2006; Kissui 2008; Abade, Macdonald, and Dickman 2014). A trained conflict officer verifies all events in person while evidence and the respondent's memory is fresh. A potential error is misidentification of kills. In some cases, the cause of death may not be witnessed and the event may be attributed to a predator that may have scavenged or chased a different predator off the kill. Alternatively, some predators may be confused with one another, particularly the cheetah and leopard that are roughly similar in pelage and size. Skillful examination of field evidence should minimize these errors. Other biases may result if either of the two key assumptions is violated – if the background extent does not represent areas where livestock and carnivores interact, or if conflict officers cannot reach all areas to document depredations. These biases would degrade the analysis by affecting the desired contrast between areas where livestock were predated vs. the interaction zone where they could be predated. Both assumptions should be true at the smallest background extent, but are unlikely to be true at the larger spatial extents in Comparison 5.

### 6.4.2 Consideration of event data characteristics

Prior research from the SDM field has shown that various biological or environmental characteristics of the presence data, such as gender (Conde et al. 2010), season (Takahata et al. 2014), and activity (Guisan and Thuiller 2005; Roever et al. 2013), can affect modeling output and conclusions. These and other studies caution that certain characteristics should be considered in the context of the research aim before modeling in order to draw more accurate and relevant conclusions. While not all potential characteristics can feasibly be considered for every model, the research aims should provide some guidance. For instance, Conde et al. (2010) model landscape occupancy of jaguars in relation to potential human-jaguar conflict areas. As there are strong differences in conflict rates between the male and female jaguars, creating sex-specific models made sense. Indeed, they found that male jaguars occupied grazing and agricultural lands more than females. I list a number of potential factors to consider when conducting depredation risk models for large carnivores (Table 6.1). For instance, livestock have different nutritional needs and movement capacities which likely preconditions particular locations for grazing (i.e. sheep and cattle may be grazed in different habitats and distances from the boma). Indeed, Miller, Jhala, and Jena (2015) found buffalo and cattle were attacked in similar environments and these environments were different than where sheep were attacked.

I designed four comparisons to demonstrate that considering depredation-related factors can substantially affect model results and conclusions. Without discussing the differences in model output for every comparison, significant and substantial variations could occur. In terms of predictor importance, nearly every variable in every comparison was significantly different between the data subsets. This indicates that the model for each subset is identifying unique depredation risk characteristics. In addition, in some cases the variation in predictor importance between different data subsets was substantial. For instance, distance to road changed substantially between pasture and boma subsets in Comparisons 1A and 1B. Distance to PA also varied substantially in Comparisons 1A and 4. Importantly, in each case, predictor importance for the non-subset data was in between the importance for the two data subsets (Appendix 6A). This suggests a similar situation to what Conde et al. (2010) demonstrate, that a combined model can blur results that are identified in factor-specific models.

Other modeling output also changed. Response curves could change more subtly (e.g. percent tree cover) or substantially (e.g. distance to road). Spatial predictions between comparisons also variably differed. Comparison 1B and 4 had substantially lower spatial correlations (~0.4) than the other comparisons, whereas Comparison 5 had the highest correlation (0.96 between clipped and expanded models). Although several comparisons between data subsets had high correlations (<0.96 and >0.6), all were outside the range of internal correlation from the 100 model replicates. Thus, accounting for specific factors did produce their own unique spatial prediction. Finally, in terms of differences in the environmental niches of depredations in environmental space, there was again substantial variation across the five comparisons. The most significant differences in environmental niche according to the Identity Test were in Comparison 1A and 3. Interestingly, these comparisons were not the same as the models that had the lowest geographic correlation (1B and 4). This may be due to the way that environmental characteristics are geographically distributed. Large differences in environmental space may not correspond proportionally to large differences in geographic space, as some conditions are more or less represented on the landscape.

Overall, in terms of predictor importance and spatial predictions, changes in event location, season, and livestock attacked (Comparisons 1, 3, and 4) resulted in relatively greater differences in model outputs than those resulting from predator or background extent (Comparisons 2 and 5). This suggests that of the three predators tested, leopard, lion, and hyena, the characteristics of their attacks at pasture in the wet season did not differ very much. Indeed, all three species are capable of attacks on either cattle or shoats although only the lion predominately attacked cattle. Also, all three carnivores are primarily nocturnal and many attacks at pasture occur at night on lost livestock.

It should be noted however that these results may be unique to this data set and other factors may be more/less important in other study sites. For instance, there may be more substantial differences between predators with an alternative carnivore guild, or less substantial differences between seasons in less seasonal habitats.

### 6.4.3 Setting the background extent

Risk mapping from other fields, such as crime mapping, can be relevant in understanding the importance of setting the background extent. The first crime maps were based on the volume of events, simple choropleth (i.e. display quantity of things in areas) maps plotting the number of events across different geographic boundaries (Chamard 2006). However, these maps can hide the relative level of risk that people actually experience. A more nuanced approach involves turning volumes into rates and involves understanding where there are more/less incidences than expected in relation to the distribution of people. Consider an example from crime mapping whereby police want to model the risk of pickpocketing. Simply plotting pickpocketing instances will likely indicate where people congregate (i.e. busy markets). However, since there are more people in particular locations, like a market, the risk per person of being pickpocketed may actually be lower. Thus, crime mappers use the term 'denominator' to denote the appropriate reference background (Eck et al. 2005; Piquero and Weisburd 2010). While typical crime mapping approaches are not correlative and instead are based on clustering or interpolation methods (e.g. choropleth maps, kernel density estimators etc.), the point remains that accounting for the reference background is key to accurate assessment of risk.

Similar to the example from crime mapping, is that carnivores and livestock likely encounter one another more often than just when there is a depredation. The key to risk modeling is to understand what turns an encounter into a deadly attack. There may be fewer attacks at some places on the landscape simply because they are less used by livestock and/or carnivores. Importantly, this does not mean that those areas are less dangerous; but because the frequency of livestock-carnivore interactions is unknown and the proportion of time spent by livestock and carnivores in different habitats is also unaccounted for, true risk is difficult to model. Detailed movement information on the predator and livestock involved is necessary to get closer to an estimate of true risk, but this is difficult to do (but see Laporte et al. 2010; Muhly et al. 2010). However, setting the background to constrain PsA selection only from areas where both livestock and carnivores overlap can be a useful approximation for movement data and a necessary first step in the analysis.

In the SDM literature, the importance of the background is well documented, as it can be the most influential modeling parameter (Phillips et al. 2009) and in most cases should be set tightly around the historical extent of the species to enable relevant comparisons with the presence data (VanDerWal et al. 2009; Elith et al. 2011; Merow, Smith, and Silander 2013). In addition, setting the proper background and hence where PsA points are drawn from influences the modeling question and is key to drawing ecological inferences (see Chapters 4 and 5). In the application to spatial risk modeling, the background should be restricted to areas where both livestock and carnivores are likely to spatially and temporally overlap. Alternative extents and methods of drawing PsA, and hence different modeling questions, are less likely to be useful in risk modeling than in the larger SDM field. Drawing PsA from locations where livestock do not go, or where carnivores are absent, degrades the analysis and affects the desired contrast between areas where carnivores could and ultimately do predate on livestock. An overly large background risks emphasizing predictors that only coarsely identify where livestock and carnivores overlap. In addition, more fine-scale variables known to affect predation risk (such as dense vegetation or steep slopes) (Laundré, Hernández, and Altendorf 2001; Laundré, Hernández, and Ripple 2010) may be overlooked at larger extents. But setting a tight spatial extent may still only identify broad scale patterns of depredation.

The analysis extent may need to be further masked to draw finer-scale conclusions. Using depredation data with a large spatial extent may ultimately function as a distribution model for livestock, and only identify where the livestock go rather than what areas have greater/lesser risk. For instance, Miller, Jhala, and Jena (2015) found that open pastures were attack hotspots for sheep, but this turned out to be their grazing location. So if sheep were spending the majority of their time grazing in pastures, then it's logical to expect more attacks there as well. This is similar to the example from crime mapping in which more pickpocketing occurs in busy markets because there are more people there, not because it is inherently more dangerous. To draw more fine-scale conclusions, a more appropriate comparison could be which pastures are more frequently attacked than other pastures, or where within grazing pastures, are sheep attacked. Drawing PsA from the general landscape would not assist in this question, only when the analysis extent is restricted to pastures could this question be resolved. In particular, this demonstrates why considering ecological factors, like the ones highlighted above, is important. Since wet and dry season pastures may be different, or sheep may graze in different locations to cattle, then only those pastures in the appropriate season or for the appropriate livestock should be included.

### 6.4.4 Caveats

Scale is an important consideration in SDM and in predation risk modeling specifically (Austin 2007; Miller et al. 2015; Rostro-García et al. 2016). Variables may be important for habitat selection or affect predation risk at one scale and be irrelevant at another. Predictors should be at biologically-relevant scales and indeed, it may be prudent to let the model select the most appropriate scale from a range of candidate scales (Boyce 2006; Mashintonio et al. 2014). While it is easy to do focal averaging and smooth predictors to achieve coarser resolutions, an overriding challenge is to obtain data at a fine enough resolution that is actually relevant for predation. Indeed, the length of a stalk for stalking predators may only be a few tens of meters, much finer than most data sets for instance. This study was conducted at the finest resolution possible; however, alternative scales could have been more appropriate. Yet, the conclusions regarding the importance of considering data subsets and the background are likely scale independent.

### 6.4.5 Best practices

When constructing a spatial risk model using distribution modeling, a few things should be considered. The researcher should identify the research question at the outset and gather abundant, unbiased event data. The research question will motivate the appropriate predictors, scale(s), and modeling algorithm(s). The research question will also influence the choice of extent, if any regions should be masked out, and how pseudo-absences are drawn. The researcher should also consider and account for relevant ecological factors that influence depredation, such as livestock species or season. Including predictors such as distance to carnivore habitat (e.g. protected area, forest cover etc.) or livestock location (e.g. boma) may be useful proxies if true carnivore and livestock distributions are unknown. However, some caution should be used, as these variables are likely to become more important as extent expands at the expense of other factors that may influence local depredation risk.

## 6.5 Conclusions

This study demonstrated how two key aspects of SDM relate to correlative risk models and how they can help to draw more informative conclusions from spatial risk models. These aspects are the consideration of various biological or environmental characteristics of the HWC event data in regards to how they may impact the research aim (in this case, depredation risk), and properly setting the background extent. Comparisons were designed to test how considering these aspects affected results from depredation risk modeling. Results demonstrated that outputs can significantly and substantially change. Thus, as in the larger SDM field, considering these aspects at the outset is necessary to conduct more useful and accurate spatial risk models.

# Appendix 6A

This appendix is split into two parts. The first contains various extra Figures and Tables germane to Chapter 6. The second part contains information related to an evaluation of if depredation events occurred in risky habitats as predicted by predation theory.

## Part 1

**Table 6A.1 Comparison of the number of training samples, variable importance scores, and evaluation metrics for Maxent models between the five different comparisons. This table is an expanded version of the one in Chapter 6; it also shows the results for the model before the data was subset.**

| Species | # Training samples | Training AUC | Testing AUC | Distance to PA | Distance to road | Distance to water | Wet season NDVI | Local diff, Wet season NDVI | Dry season NDVI | Local diff, Dry season NDVI |
|---|---|---|---|---|---|---|---|---|---|---|
| **Comparison 1A: Location** | | | | | | | | | | |
| All attack | 559 | 0.79 | 0.79 ± 0.03 | 25.84 ± 1.1 | 34.1 ± 1.28 | 5.6 ± 0.64 | 3.14 ± 0.76 | 0.56 ± 0.29 | 0.23 ± 0.12 | 0.06 ± 0.04 |
| Boma | 357 | 0.84 | 0.84 ± 0.03 | 33.87 ± 1.45 | 31.24 ± 2.22 | 4.56 ± 0.49 | 1.05 ± 0.47 | 0.17 ± 0.15 | 0.26 ± 0.11 | 0.04 ± 0.03 |
| Pasture | 224 | 0.76 | 0.73 ± 0.04 | 9.22 ± 1.14 | 52.89 ± 2.52 | 8.7 ± 1.41 | 8.76 ± 1.74 | 1.48 ± 0.6 | 0.69 ± 0.28 | 0.23 ± 0.15 |
| **Comparison 1B: Location** | | | | | | | | | | |
| Hyena attacks in wet season on cattle in both locations | 62 | 0.82 | 0.77 ± 0.09 | 13.06 ± 1.9 | 19.07 ± 2.82 | 11.09 ± 1.89 | 5.41 ± 1.12 | 1.05 ± 0.51 | | |
| … at boma | 36 | 0.87 | 0.82 ± 0.09 | 15.5 ± 2.69 | 28.43 ± 3.53 | 8.89 ± 1.83 | 10.55 ± 2.28 | 0.11 ± 0.14 | | |
| … at pasture | 27 | 0.72 | 0.57 ± 0.16 | 23.39 ± 6.68 | 6.16 ± 7.85 | 4.46 ± 2.48 | 0.38 ± 0.53 | 8.92 ± 6.29 | | |
| **Comparison 2: Predator** | | | | | | | | | | |
| All predator attacks at pasture in wet season on all livestock | 117 | 0.8 | 0.78 ± 0.05 | 8.38 ± 1.25 | 61.83 ± 2.53 | 6.04 ± 1.29 | 0.23 ± 0.24 | 0.05 ± 0.08 | | |
| Hyena | 64 | 0.79 | 0.75 ± 0.08 | 4.16 ± 1.41 | 53.01 ± 3.56 | 4.48 ± 1.56 | 0.09 ± 0.14 | 0.08 ± 0.24 | | |
| Leopard | 22 | 0.88 | 0.8 ± 0.08 | 3 ± 1.04 | 56.49 ± 3.29 | 5.86 ± 1.46 | 7.79 ± 3.03 | 2.28 ± 1.42 | | |
| Lion | 20 | 0.85 | 0.71 ± 0.07 | 4.66 ± 3.29 | 71.78 ± 5.27 | 1.38 ± 1.13 | 1.89 ± 1.35 | 1.57 ± 0.71 | | |
| **Comparison 3: Season** | | | | | | | | | | |
| Hyena attacks at pasture on shoats in both seasons | 77 | 0.79 | 0.75 ± 0.07 | 5.71 ± 1.45 | 52.46 ± 3.16 | 8.71 ± 1.53 | 7.33 ± 2.56 | 1.18 ± 0.73 | 0.4 ± 0.34 | 0.45 ± 0.62 |
| … in dry season | 45 | 0.82 | 0.76 ± 0.08 | 4.76 ± 1.57 | 44.23 ± 2.1 | 8.6 ± 1.82 | | | 0.06 ± 0.1 | 0.29 ± 0.42 |
| … in wet season | 35 | 0.82 | 0.78 ± 0.08 | 0.62 ± 0.32 | 62.3 ± 3.77 | 3.24 ± 1.34 | 0.07 ± 0.19 | 0.1 ± 0.13 | | |
| **Comparison 4: Livestock** | | | | | | | | | | |
| Hyena attacks at pasture in wet season on all livestock | 64 | 0.79 | 0.75 ± 0.08 | 4.16 ± 1.41 | 53.01 ± 3.56 | 4.48 ± 1.56 | 0.09 ± 0.14 | 0.08 ± 0.24 | | |
| … on cattle | 27 | 0.72 | 0.57 ± 0.16 | 23.39 ± 6.68 | 6.16 ± 7.85 | 4.46 ± 2.48 | 0.38 ± 0.53 | 8.92 ± 6.29 | | |
| … on shoat | 35 | 0.82 | 0.78 ± 0.08 | 0.62 ± 0.32 | 62.3 ± 3.77 | 3.24 ± 1.34 | 0.07 ± 0.19 | 0.1 ± 0.13 | | |
| **Comparison 5: Background** | | | | | | | | | | |
| Lion attacks at pasture in wet season on cattle, clip | 17 | 0.85 | NA | 5.08 ± 2.81 | 69.02 ± 5.05 | 0.09 ± 0.17 | 1.94 ± 1.32 | 6.14 ± 3.25 | | |
| …, expanded (25 km) | 17 | 0.88 | NA | 3.27 ± 1.92 | 58.78 ± 3.52 | 0.46 ± 0.5 | 1.92 ± 0.87 | 3.31 ± 1.38 | | |
| …, full (150 km) | 17 | 0.89 | NA | 13.2 ± 2.46 | 21.49 ± 1.62 | 0.05 ± 0.14 | 3.99 ± 1 | 0.13 ± 0.37 | | |

**Table 6A.1 continued**

| Species | Human pop'n density | Elevation | Local diff, Elevation | % Tree cover | Local diff, % Tree cover | Topographic Roughness Index | Local diff, TRI |
|---|---|---|---|---|---|---|---|
| **Comparison 1A: Location** | | | | | | | |
| All attack | 10.99 ± 1.89 | 0.22 ± 0.09 | 0 ± 0 | 16.43 ± 1.26 | 1.53 ± 0.48 | 0 ± 0 | 0.02 ± 0.02 |
| Boma | 8.88 ± 2.3 | 7.27 ± 1.22 | 0.01 ± 0.05 | 10.63 ± 1.01 | 0.58 ± 0.35 | 0.06 ± 0.06 | 0.29 ± 0.18 |
| Pasture | 0.23 ± 0.3 | 2.5 ± 0.62 | 0.28 ± 0.22 | 12.57 ± 1.39 | 0.82 ± 0.44 | 0.31 ± 0.2 | 1.29 ± 0.5 |
| **Comparison 1B: Location** | | | | | | | |
| Hyena attacks in wet season on cattle in both locations | 29.5 ± 3.94 | 0.97 ± 0.43 | 7.32 ± 1.46 | 7.13 ± 0.96 | 2.13 ± 1.01 | 1.21 ± 0.85 | 0.57 ± 0.74 |
| … at boma | 18.24 ± 3.33 | 0.52 ± 0.68 | 2.41 ± 1.18 | 5.22 ± 1.19 | 1.56 ± 0.87 | 0.88 ± 0.61 | 1.11 ± 1.19 |
| … at pasture | 5.44 ± 2.58 | 4.63 ± 2.95 | 19.81 ± 3.65 | 14.05 ± 2.58 | 4.01 ± 1.58 | 5.71 ± 2.37 | 2.24 ± 1.78 |
| **Comparison 2: Predator** | | | | | | | |
| All predator attacks at pasture in wet season on all livestock | 0.89 ± 0.58 | 4.29 ± 1 | 1.49 ± 0.49 | 10.67 ± 1 | 3.53 ± 0.87 | 0.23 ± 0.34 | 2.36 ± 0.7 |
| Hyena | 6.98 ± 1.52 | 3.75 ± 1.54 | 3.02 ± 0.96 | 10.13 ± 1.39 | 4.75 ± 1.35 | 0.08 ± 0.25 | 9.38 ± 1.97 |
| Leopard | 2.01 ± 1.21 | 10.14 ± 1.77 | 2.26 ± 1.31 | 3.78 ± 0.63 | 3.13 ± 0.88 | 0.13 ± 0.21 | 3.14 ± 2.33 |
| Lion | 0.31 ± 0.27 | 0.96 ± 0.71 | 0.38 ± 1 | 12.72 ± 2.04 | 0.32 ± 0.42 | 0.37 ± 1.33 | 3.51 ± 1.19 |
| **Comparison 3: Season** | | | | | | | |
| Hyena attacks at pasture on shoats in both seasons | 0.11 ± 0.12 | 2.53 ± 0.81 | 0.08 ± 0.11 | 14.17 ± 2.77 | 0.38 ± 0.23 | 0.41 ± 0.26 | 6.04 ± 1.27 |
| … in dry season | 0 ± 0 | 0.76 ± 0.73 | 0.01 ± 0.05 | 35.2 ± 2.7 | 0 ± 0 | 0.44 ± 0.4 | 5.31 ± 1.24 |
| … in wet season | 7.32 ± 2.05 | 3.13 ± 1.18 | 1.39 ± 0.41 | 8.02 ± 1.18 | 3.96 ± 1.29 | 0.75 ± 0.76 | 9.01 ± 1.71 |
| **Comparison 4: Livestock** | | | | | | | |
| Hyena attacks at pasture in wet season on all livestock | 6.98 ± 1.52 | 3.75 ± 1.54 | 3.02 ± 0.96 | 10.13 ± 1.39 | 4.75 ± 1.35 | 0.08 ± 0.25 | 9.38 ± 1.97 |
| … on cattle | 5.44 ± 2.58 | 4.63 ± 2.95 | 19.81 ± 3.65 | 14.05 ± 2.58 | 4.01 ± 1.58 | 5.71 ± 2.37 | 2.24 ± 1.78 |
| … on shoat | 7.32 ± 2.05 | 3.13 ± 1.18 | 1.39 ± 0.41 | 8.02 ± 1.18 | 3.96 ± 1.29 | 0.75 ± 0.76 | 9.01 ± 1.71 |
| **Comparison 5: Background** | | | | | | | |
| Lion attacks at pasture in wet season on cattle, clip | 3.38 ± 1.43 | 1.94 ± 1.07 | 0.24 ± 1.34 | 10.1 ± 2.54 | 0.2 ± 0.39 | 0.37 ± 1.25 | 0.61 ± 0.67 |
| …, expanded (25 km) | 0.88 ± 0.19 | 0.74 ± 0.42 | 0.61 ± 1.59 | 26.35 ± 2.33 | 0.5 ± 0.69 | 0.54 ± 1.63 | 2.13 ± 0.69 |
| …, full (150 km) | 0.01 ± 0.01 | 5.61 ± 0.76 | 0.02 ± 0.05 | 54.95 ± 2.7 | 0.05 ± 0.09 | 0.09 ± 0.3 | 0.41 ± 1.15 |

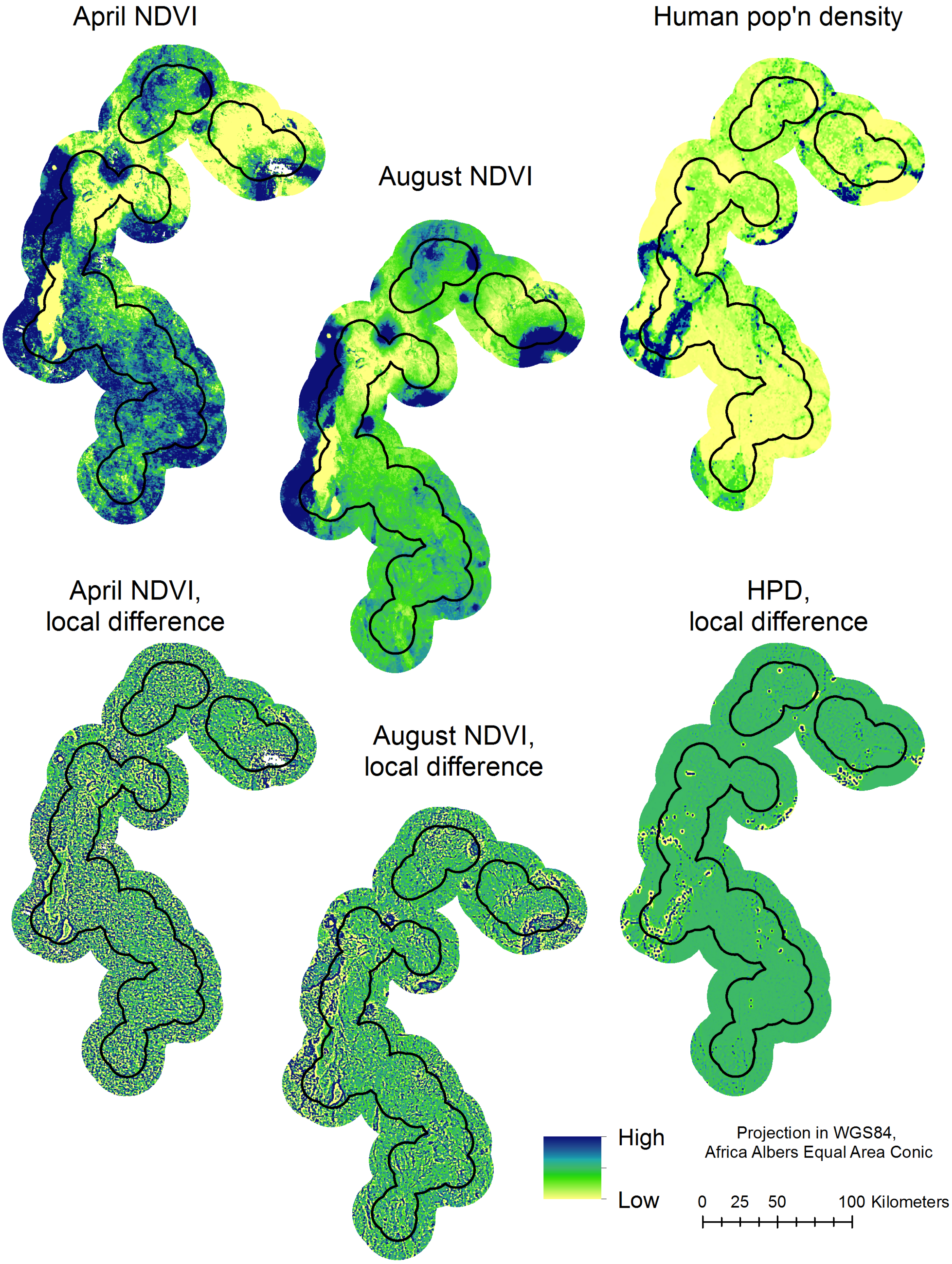


**Figure 6A.1** **Predictor variables with the 12 km buffer shown in black outline.**

]

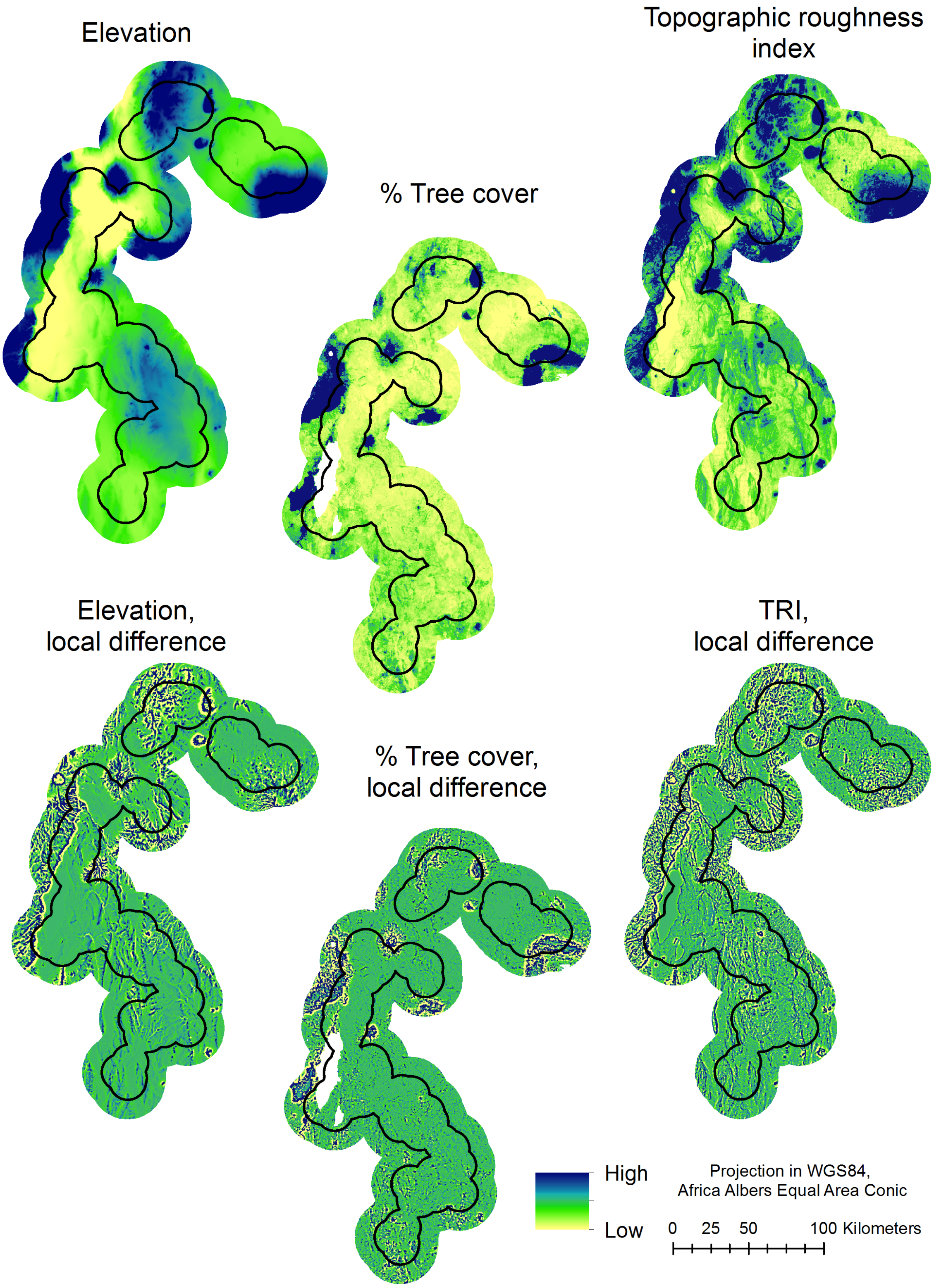
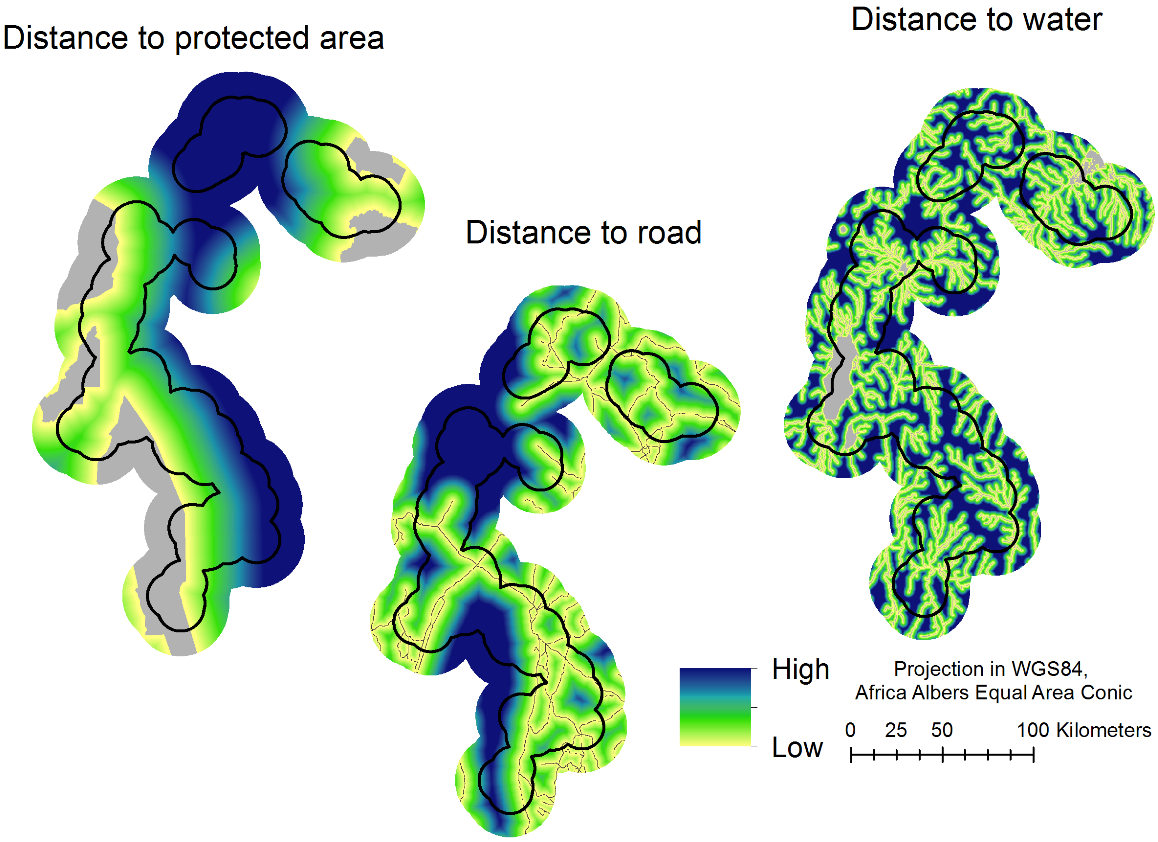


**Figure 6A.2** **Predictor variables with the 12 km buffer shown in black outline.**



Distance to protected area

Distance to road

Distance to water

**Figure 6A.3** **Predictor variables with the 12 km buffer shown in black outline.**

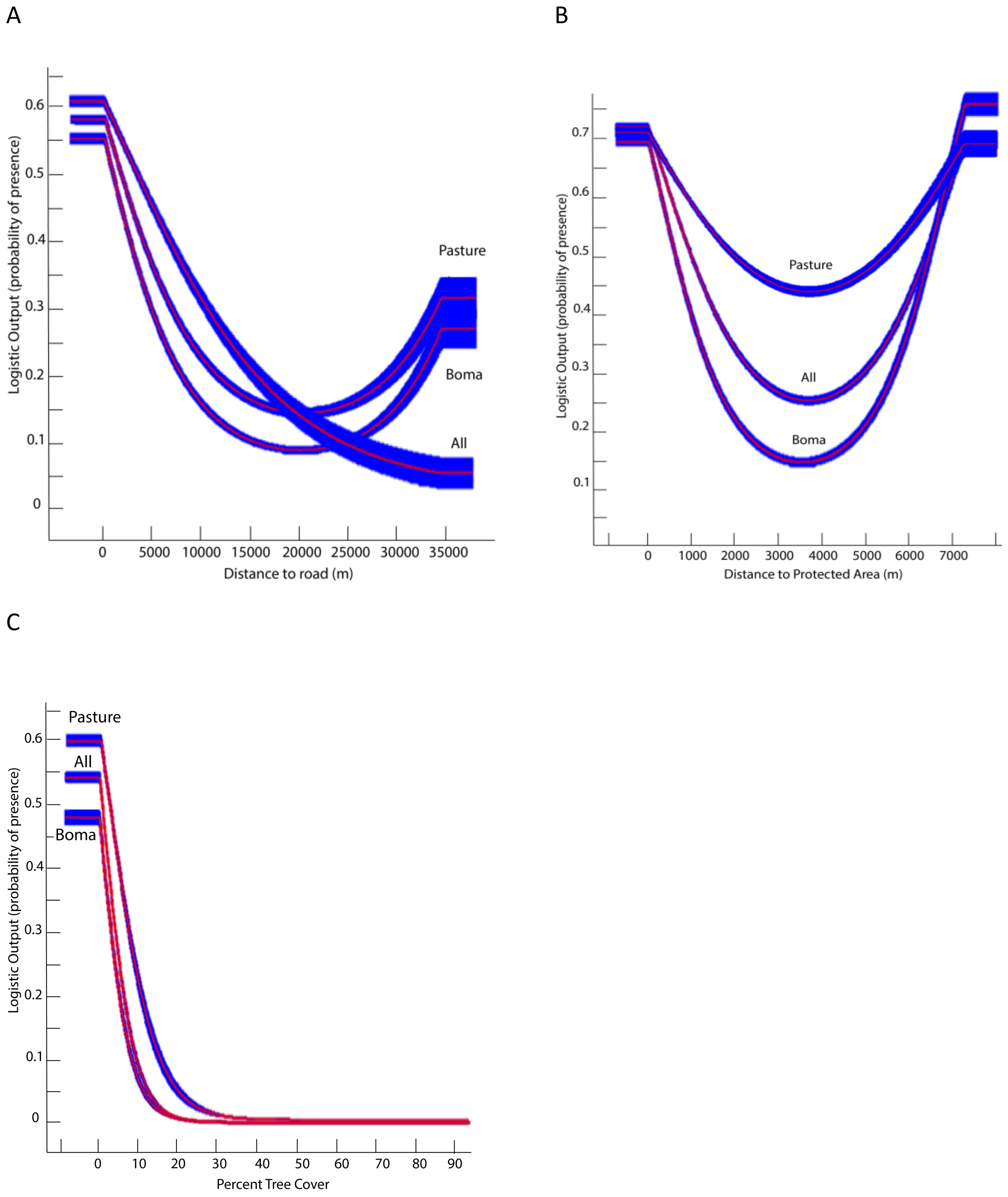


**Figure 6A.4** **Variation in response curves in Comparison 1A for the three most important variables: A) distance to road, B) distance to protected area, and C) percent tree cover.**

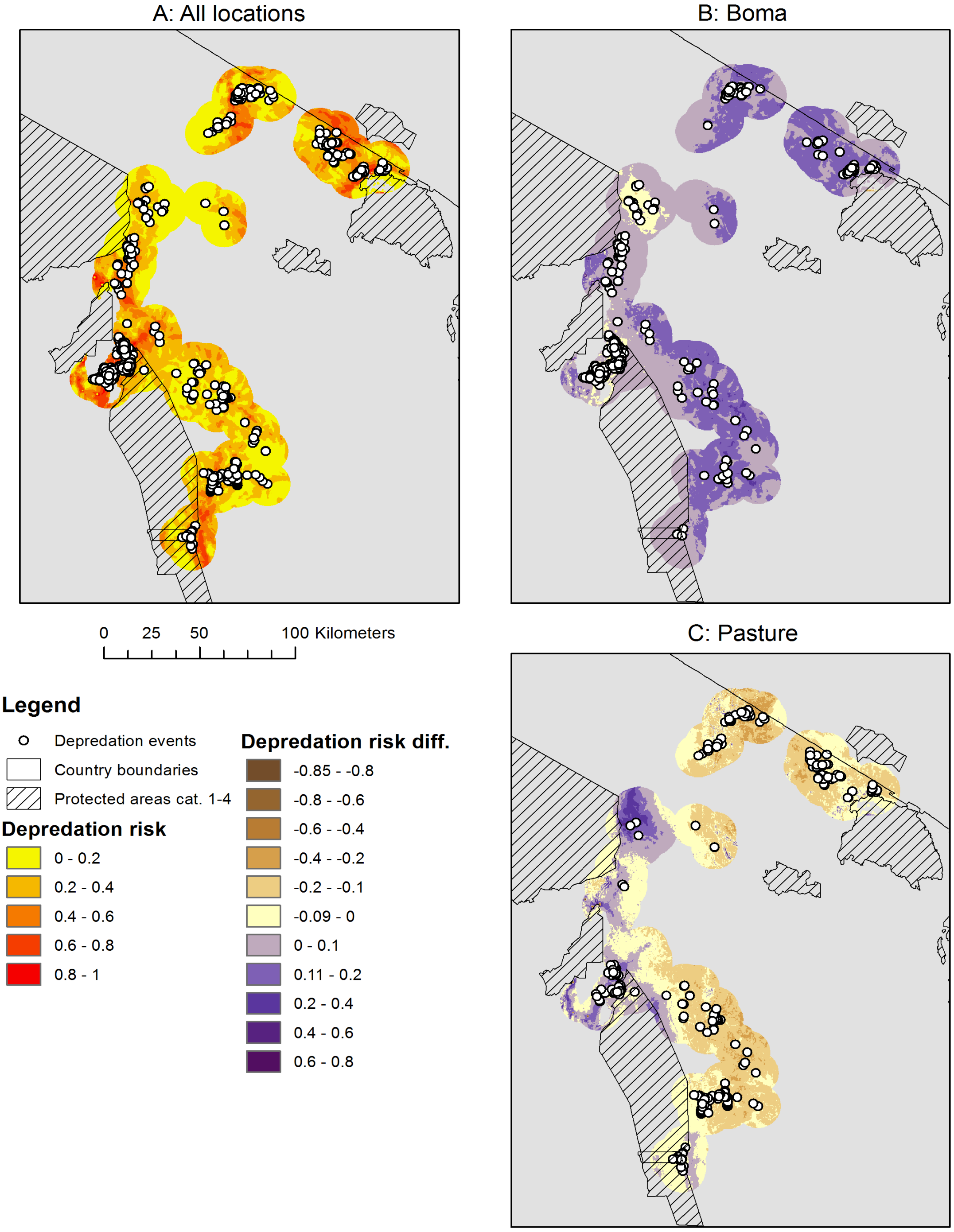


**Figure 6A.5 Spatial predictions of depredation risk in Comparison 1A: location. Overall depredation risk using all event data is given in Figure A. The difference in risk from A is shown in B and C (i.e. the models trained on the data subsets). A positive difference value indicates a higher depredation risk in the full model than in the subset model.**

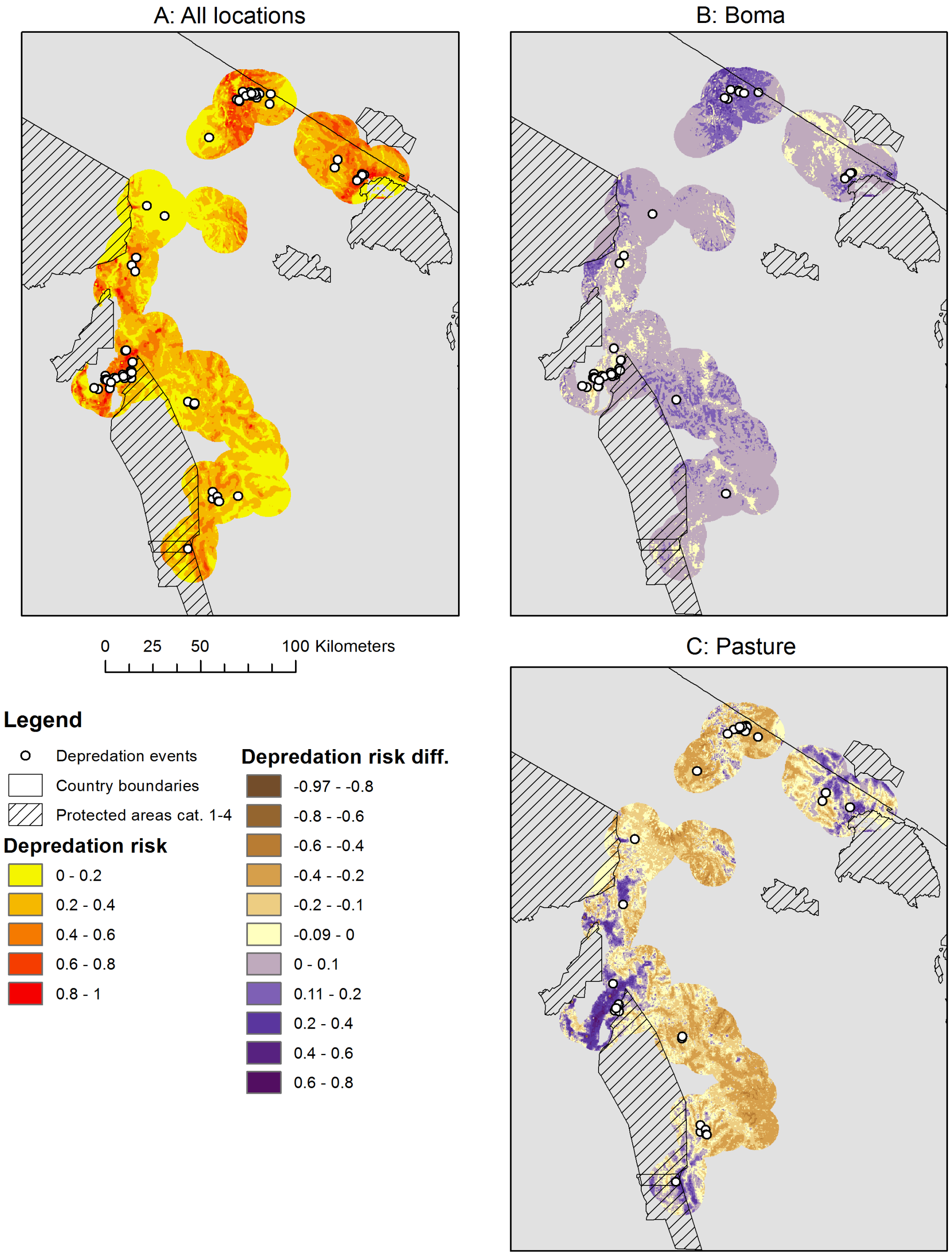


**Figure 6A.6. Spatial predictions of depredation risk in Comparison 1B: hyena attacks in the wet season on cattle at either location. Overall depredation risk is given in Figure A, and the difference in risk from A in the boma and pasture model is shown in B and C respectively.**

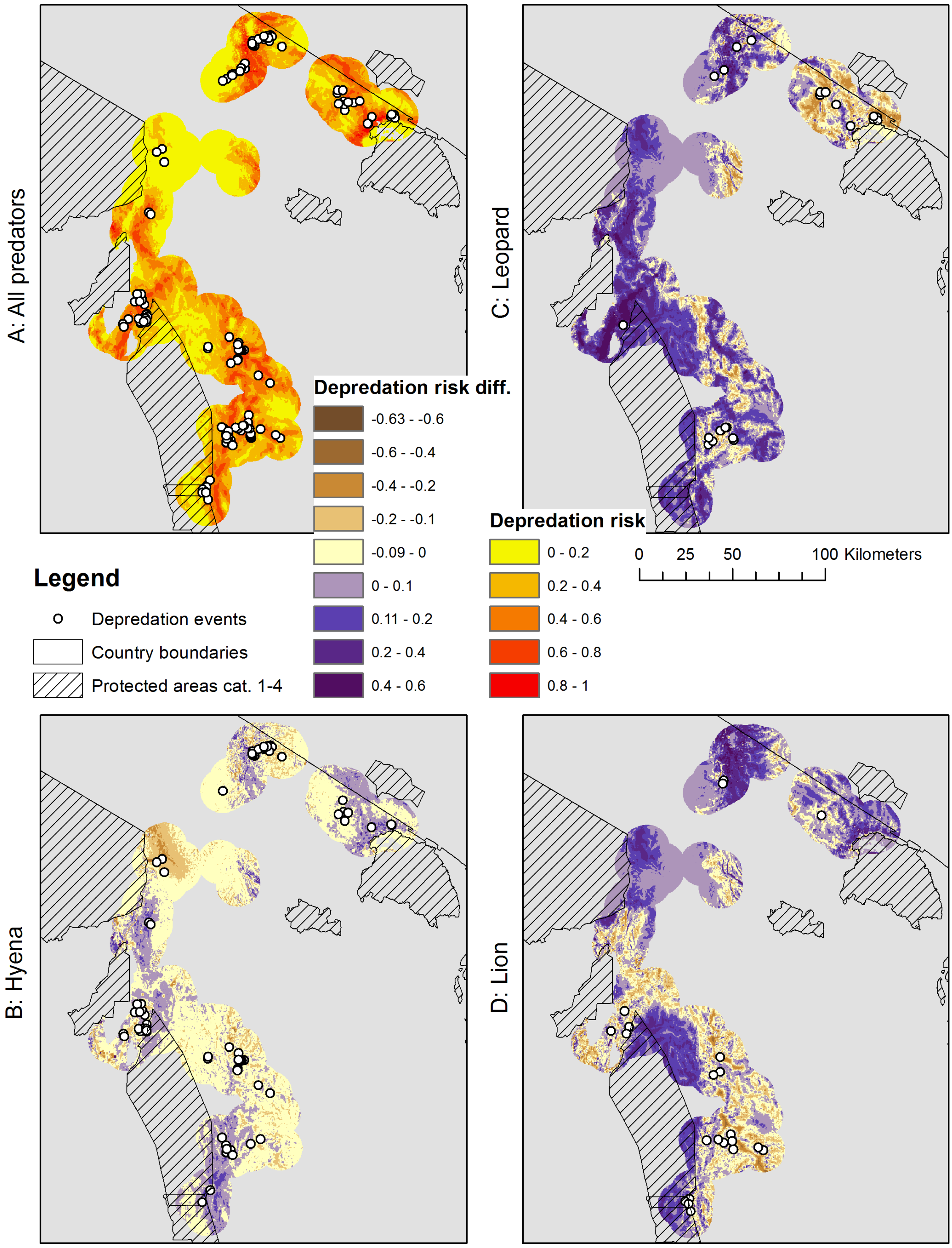


**Figure 6A.7** **Spatial predictions of depredation risk in Comparison 2: attacks at pasture in the wet season on all livestock by different predators. Overall depredation risk is given in Figure A, and the difference in risk from A in the hyena, leopard, and lion models is shown in B, C and D.**

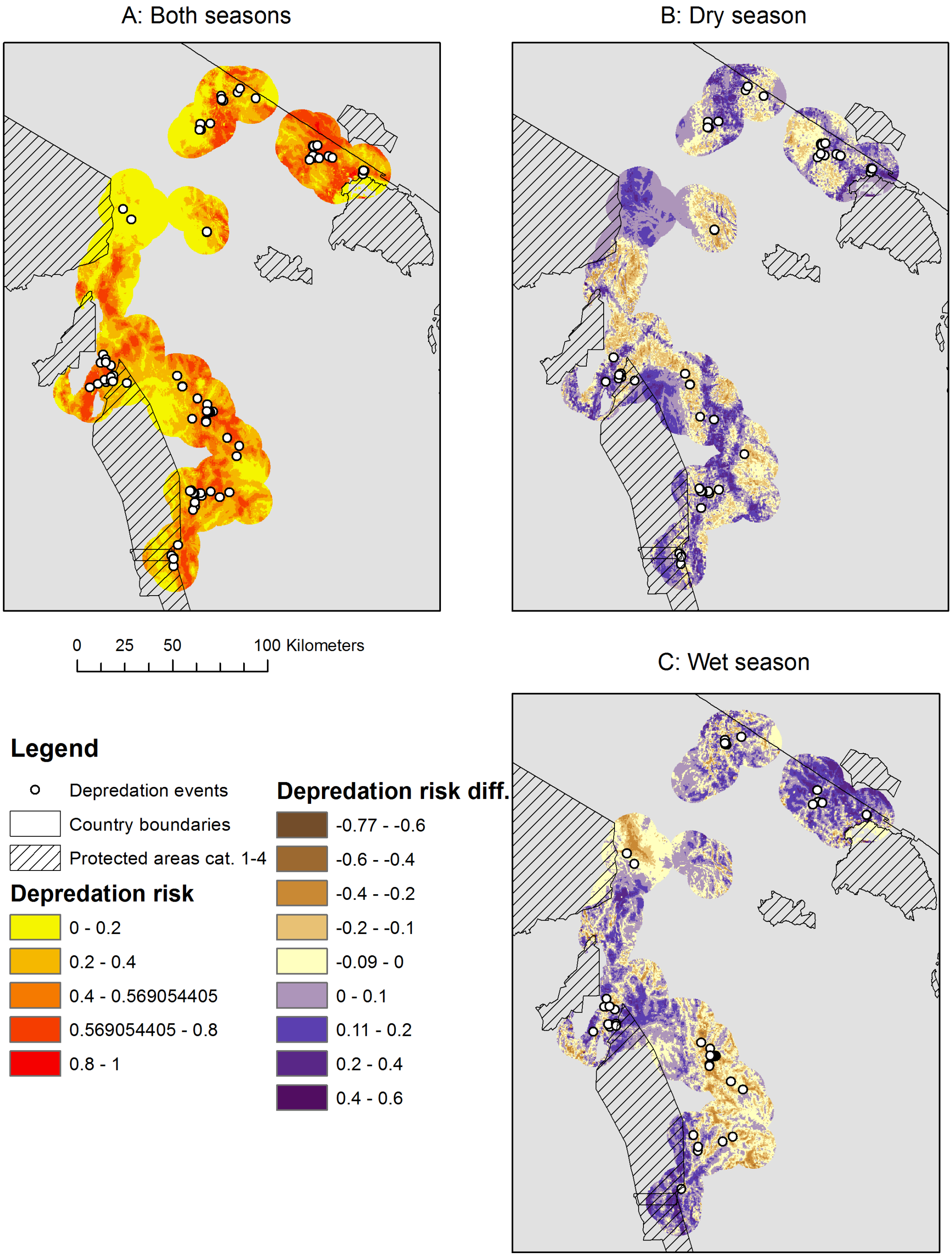


**Figure 6A.8 Spatial predictions of depredation risk in Comparison 3: hyena attacks at pasture on shoats in different seasons. Overall depredation risk is given in Figure A, and the difference in risk from A in the dry and wet season models is shown in B and C respectively.**

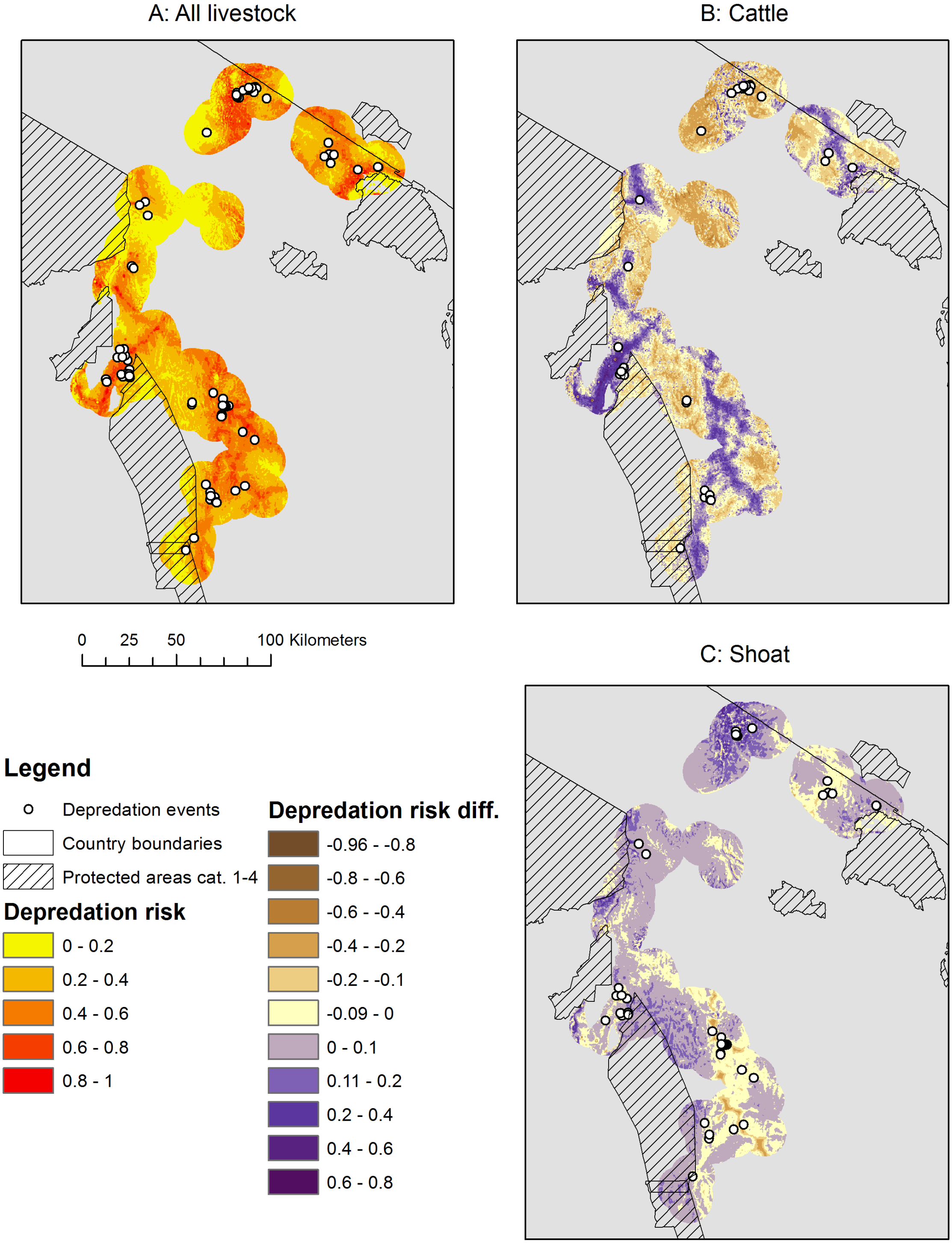


**Figure 6A.9 Spatial predictions of depredation risk in Comparison 4: hyena attacks at pasture in wet season on different livestock. Overall depredation risk is given in Figure A, and the difference in risk from A in the cattle and shoat model is shown in B and C respectively.**

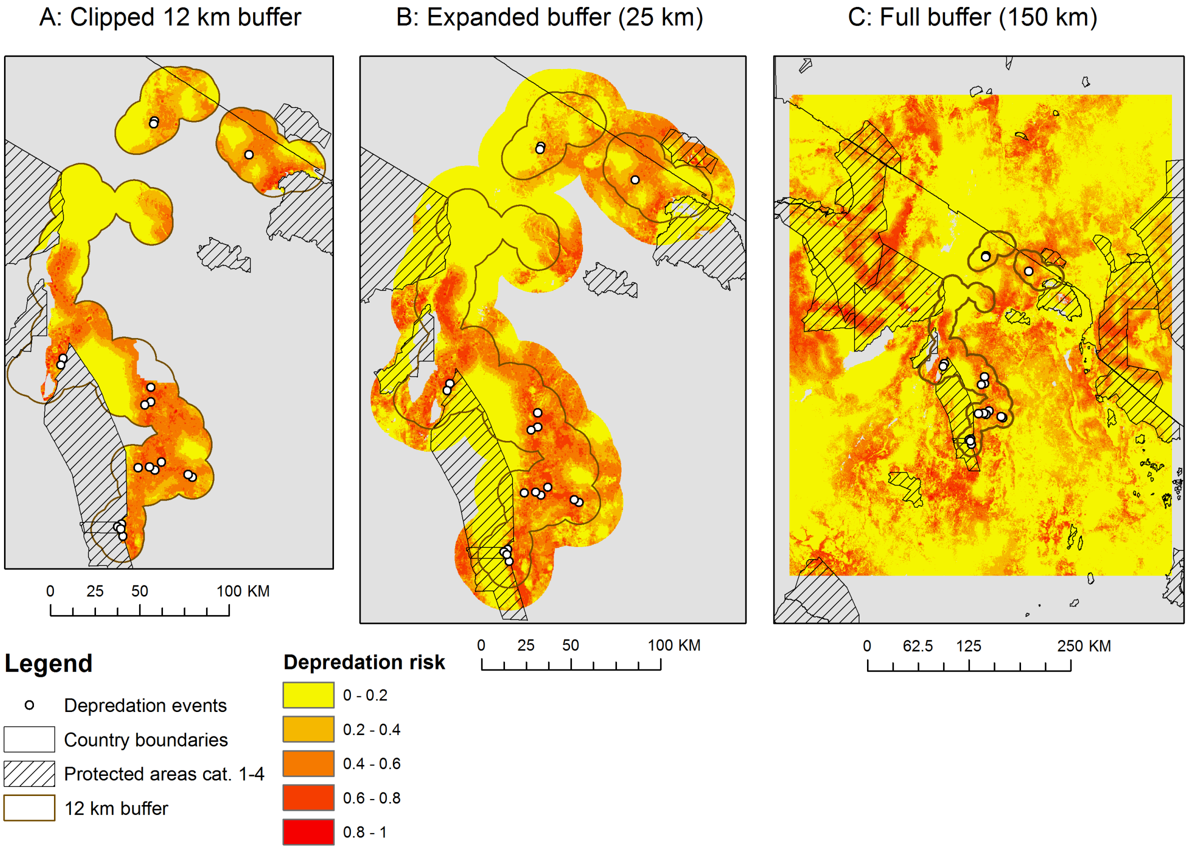


**Figure 6A.10** **Spatial predictions of depredation risk in Comparison 5: lion attacks on cattle at pasture in the wet season with different analysis extents.**

**Table 6A.2 Pearson correlation coefficients between data subsets and internal model consistency for each of the five comparisons. This table is an expanded version of the one in Chapter 6; it also indicates the results from the model before the data was subset. The values on the 1:1 line are the internal correlations of the spatial predictions among 100 runs. The values below the 1:1 line are the correlations in spatial predictions between the different models.**

| | | | | |
|---|---|---|---|---|
| **Comparison 1A: Location** | All attacks | All boma | All pasture | |
| All events | 0.99 ± 0.00 | | | |
| All boma events | 0.95 | 0.99 ± 0.00 | | |
| All pasture events | 0.86 | 0.69 | 0.99 ± 0.00 | |
| **Comparison 1B: Location** | … both locations | … at boma | … at pasture | |
| Hyena attacks in wet season on cattle... in both locations | 0.98 ± 0.01 | | | |
| … at boma | 0.96 | 0.98 ± 0.02 | | |
| … at pasture | 0.56 | 0.38 | 0.95 ± 0.05 | |
| **Comparison 2: Predator** | … all predators | Only hyena | Only leopard | Only lion |
| Attacks at pasture in wet season on all livestock by… all predators | 0.99 ± 0.00 | | | |
| … Only hyena | 0.93 | 0.98 ± 0.01 | | |
| … Only leopard | 0.75 | 0.7 | 0.99 ± 0.01 | |
| … Only lion | 0.83 | 0.79 | 0.68 | 0.98 ± 0.01 |
| **Comparison 3: Season** | ... both seasons | … in dry season | … in wet season | |
| Hyena attacks at pasture on shoats… … in both seasons | 0.99 ± 0.01 | | | |
| … in dry season | 0.86 | 0.98 ± 0.01 | | |
| … in wet season | 0.84 | 0.62 | 0.98 ± 0.01 | |
| **Comparison 4: Livestock** | ... all livestock | ... on cattle | … on shoats | |
| Hyena attacks at pasture in wet season... on all livestock | 0.98 ± 0.01 | | | |
| … on only cattle | 0.6 | 0.95 ± 0.05 | | |
| ... on only shoat | 0.93 | 0.42 | 0.98 ± 0.01 | |
| **Comparison 5: Background** | …, clipped | …, expanded | …, full extent | |
| Lion attacks at pasture in wet season on cattle… | | | | |
| … at clipped extent | 0.99 ± 0.01 | | | |
| … at expanded extent (25 km) | 0.96 | 0.98 ± 0.01 | | |
| ..., at full extent (150 km) | 0.86 | 0.89 | 0.99 ± 0.01 | |

## Part 2 Risky habitats

Predation theory suggests carnivores will select for areas with greater hunting success rather than prey abundance (Hopcraft, Sinclair, and Packer 2005; Balme, Hunter, and Slotow 2007). When analyzing hunting behavior, stages of a hunt can be decomposed and different predictors are important at the different stages (Hebblewhite, Merrill, and Mcdonald 2005). For instance, features at the landscape level may enhance predator/prey encounters while local scale terrain or vegetation may affect the chance of a successful attack. Local features that affect hunting success may be terrain (i.e. tributary junctions, gullies, steep slopes etc.), biotic (i.e. vegetation thickets, woody debris etc.), or human-caused (i.e. fences, roads etc.) (Ripple and Beschta 2004; Laundré, Hernández, and Ripple 2010). Thus, at the local scale, carnivores select for 'risky' habitats, areas where wild (and presumably domestic) prey is more vulnerable, and predation success is higher.

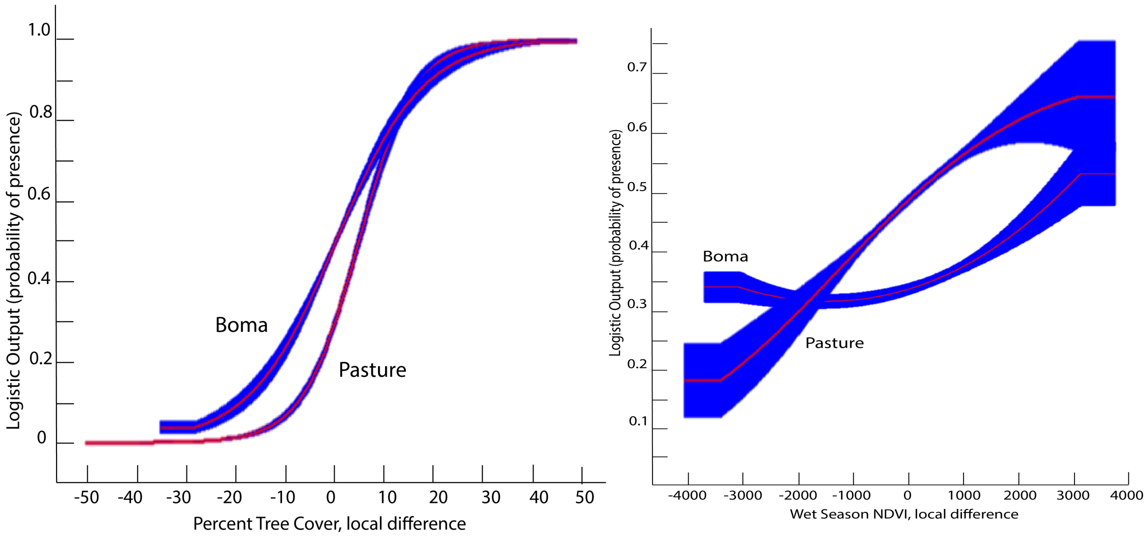
Indeed, a number of depredation risk studies have found that livestock are attacked in dense vegetation and in areas with short sightlines or other features that can enhance hunting success (Shrader et al. 2008; Kissling, Fernández, and Paruelo 2009; Davie et al. 2014; Soh et al. 2014; Miller et al. 2015b). However, local difference predictors (see section 6.2.3 for a description of how these were produced) allow a more nuanced approach to this than previous studies. To evaluate if depredation events are more likely to occur in risky habitats, prey vulnerability is expected to be higher in areas where the local vegetation is denser (for instance) than the surrounding area. The percent tree cover predictor relates to the overall percentage of cover, whereas the local difference of this predictor assesses if there is greater (or lesser) vegetation biomass in the focal cell compared to the local neighborhood. This is a closer geospatial representation of risky habitat (e.g. a locally dense thicket), than the simple unmodified variable. Therefore, response curves for local differences of wet season NDVI, and percent tree cover were reviewed. If attacks are more likely in risky habitat, then a higher probability of attack is expected at positive values (i.e. greater tree cover in that cell compared to the local neighborhood).

I investigated the response curves of Comparison 1A to illustrate the value of this approach. Probability of attack increased greatly as the local difference predictors for tree cover and NDVI increased (Figure 6A.12). As percent tree cover of the focal cell increased relative to the surrounding landscape, probability of attack rose from essentially zero to one. A similar but less dramatic increase is seen for wet season NDVI. This result is particularly interesting when compared to the response curve for percent tree cover overall (Figure 6A.4). For percent tree cover, the probability of attack drops quickly towards zero as tree cover approaches 20 - 30%. Thus, attacks are almost exclusively occurring in the more open landscapes of the analysis extent. However, within the relatively treeless plains, if there are patches of denser vegetation, then those areas are clearly very high-risk.

A corollary to this is that the local difference predictors should be more important for attacks at pasture than at boma. Features that lead to a successful attack at boma include the degree of human activity, and the condition and features of the boma itself (Ogada et al. 2003; Woodroffe et al. 2007). Alternatively, attacks at pasture likely rely to a greater extent simply on the biophysical predictors of the landscape. Indeed, the importance of the local difference of both variables was higher at pasture than at boma in Comparison 1A.

This is the first study using a distribution model to document that carnivore attacks on livestock are more likely in habitats with denser vegetation, i.e. risky habitats. Yet, this result is preliminary and requires more investigation. Jarnevich et al. (2015) cautions that conclusions from correlative distribution models should be treated as hypotheses and subject to additional scrutiny.

Figure 6A.12 Variation in response curves for local differences of A) percent tree cover, and B) wet season NDVI in Comparison 1A.



A.

B.

# Chapter 7

## General discussion

Low-lying clouds approach a small Maasai community in the Ngorongoro Conservation Area. © Andrew Jacobson

# Chapter 7: General Discussion

This thesis addresses some challenges and provides improved frameworks for commonly used methods to study wildlife distribution and threats. In particular, I consider issues associated with identification of human impacted areas and the use of species distribution modeling (SDM). In Chapter 3, I develop a new tool that enables land cover classification using Google Earth's high-resolution imagery data. Identification of human impacted areas in heterogeneous savannah habitat can be difficult with medium-resolution earth observation data and traditional classification methods. I apply this tool in East Africa to produce a new data set of human impacted areas, and compare it to existing land cover data sets. In Chapter 4, I review existing methods of generating pseudo-absence (PsA) data for use in discriminatory SDM approaches from an ecological standpoint and discuss how altering the generation strategy affects the modeling question. I show that not all current methods have ecological justification. I also develop several novel PsA selection strategies that use widely available data on species distributions; these pose different and potentially useful modeling questions. In Chapter 5, I apply the lessons from the previous chapter to model four large carnivore distributions in East Africa. I use one existing and two PsA selection strategies proposed in this thesis to build ecological inferences and support conclusions regarding several hypotheses. Results show that human land cover and human populations are both contributors to shrinking large carnivore range, but that expanding human impacted areas are a greater threat. In addition, results partially support the prediction that carnivores with larger home ranges are more sensitive to habitat loss than those with smaller home ranges. Modeling results also demonstrate that using PsA drawn from a species' extirpated range is particularly useful in identifying habitat suitability for locations with sparse data. Finally, I also apply modeled habitat suitability to expert-derived species range maps for conservation purposes. In Chapter 6, I highlight two important aspects from the larger SDM literature that are commonly overlooked when applying SDM to human-wildlife conflict (HWC) data to create spatial risk maps. I demonstrate how both issues can affect modeling results and conclusions. If properly considered in advance of modeling, these aspects can improve the identification of risk factors and produce more informative spatial risk models.

This thesis contributes to our understanding of the threats that large carnivores face, and to the approaches that may be most effective for their conservation, by exploring species distributions in relation to human populations and improved land cover data. All modeled large carnivore species showed rapidly decreasing habitat suitability as human population density (HPD) and the proportion of human land cover increased with human land cover a more important predictor. The implication of these results is that expanding agriculture in East Africa has the potential to greatly diminish remaining large carnivore habitat. Even low levels of land conversion (20-30%) were associated with reductions in carnivore presence probabilities below 50% (Figure 5.7). My results show that cheetah, lion, and wild dog are most strongly impacted (in order of most to least sensitive) unless conservation efforts are implemented. Leopards are not immune to land conversion or human population densities either, and are also vulnerable. This thesis highlights the importance of the PA network for large carnivore conservation (Appendix 5A) and identifies areas outside the network that act as refuges or range extension (Appendix 5B). In addition, I identify several priority areas where the probability of continuing large carnivore presence for all species is high, and other areas where populations are threatened. This can help prioritize where conservation action should occur.

The results presented here suggest that for large carnivore conservation, land use plans should be developed for East Africa that concentrate future agricultural and urban growth. Pastoralism, practiced with limited levels of cultivation, is compatible with large carnivore conservation particularly when practiced in conjunction with methods to minimize HWC. Pastoralism should be prioritized in buffers around and corridors between protected areas (PAs), as well as where farming is marginal. Policies that limit subdivision of land are also recommended. These conservation recommendations provide a way for East African communities to grow while limiting their impact on large carnivore persistence.

## 7.1 Existing challenges to common methods

This dissertation aims to contribute to conservation science by providing improved frameworks for common methods used to study wildlife distributions and threats, with examples from large carnivores in East Africa. Unsurprisingly, existing methods are imperfect and face many challenges. This dissertation specifically addresses three of these challenges:

- overcoming the difficulties of identifying human impacted areas from earth observation data is a critical first step in mapping habitat loss;

- determining an appropriate strategy for generating pseudo-absence (PsA) data. This is vital to effective discriminatory SDM but there is no clear consensus on how to achieve this nor on the ecological meaning of the various approaches in use;
- addressing the role of the background extent and standard practice in framing the modeling question of spatial risk using discriminatory distribution models, and how to implement these in practice.

Overall, these challenges constrain proper assessment of the threats species' face. Recognizing and overcoming these challenges will improve our ability to prioritize and address threats, develop effective solutions, and ultimately contribute to wildlife persistence.

## 7.2 Classification of croplands and urban areas in East Africa

Habitat loss is an important driver of the current extinction crisis (Pimm and Raven 2000) and identification of human impacted areas is a critical first step in mapping habitat loss. Yet, as I show in Chapter 3, existing land cover data sets for Africa variably and poorly classify croplands and urban areas. In part this is due to the heterogeneous nature of much of African croplands, combined with the inherent patchiness of savannah ecosystems. Additionally, traditional land cover classification techniques can be expensive, sometimes because the earth observation data have to be purchased and/or more likely because of specialized training and software requirements (Pettorelli, Safi, and Turner 2014). These elements can pose significant challenges to conservation practitioners. Thus, I developed a new tool, GE Grids, to address this shortcoming. This free tool overlays an interactive binary grid on Google Earth data and is the first to create raster data from this application that can be imported into GIS software. Google Earth provides free access to high-resolution (<10 m) imagery across the globe and is easy to use. An important difference of GE Grids from previous applications designed to access Google Earth data, is that this tool is the only one that allows the user to conduct their own land cover classification.

I used GE Grids in East Africa to identify human land cover. I compared the resulting data set with existing regional and global land cover data sets (Figure 3.5). I found ~30% of the region converted to human land cover, varying from ~18% in Kenya to ~85% in Burundi (Table 3.2). Land cover class agreement with my data set varied widely between products and even within products at the country level (Table 3.4). The highest resolution comparative data sets (30 m or less) had the best overlap with the GE Grids output. The usefulness of this approach is demonstrated by the quick take up of my data set: Masenga et al. (2016) use these new data on cropland and urban area extents to illustrate how these land cover types can constrain African wild dog dispersal; Broekhuis and Gopalaswamy (2016) use it in estimating cheetah densities in the Mara. I use this layer as an input into distribution modeling to ascertain the effects of human land cover and various human population densities on carnivore distribution.

The GE Grids tool has some limitations and other caveats. The process requires manual identification of land cover types and hence is time-consuming and best used at smaller extents (country or smaller). The manual process also limits easy replication. Finally, the tool can be substantially improved by integrating access to other high-resolution earth observation data sets (like Bing imagery), and by increasing the number of classification types that can be designated.

Yet GE Grids did enable the creation of a novel land cover data set for East Africa. Despite the manual identification of land cover, GE Grids had high reliability between users (83%) and from a single user (95%). Overall, the tool is useful to validate existing land cover classifications or to create unique classifications in situations that otherwise challenge traditional automated classification techniques. This approach has the potential for widespread adoption among conservation practitioners who desire accurate, and tailored information on specific land cover classes in their own study area.

## 7.3 Species distribution modeling and application to large carnivores in East Africa

SDM is an increasingly popular technique to study species' distributions, their drivers, and to explore ecological issues such as species response to novel conditions such as climate change (Ahmed et al. 2015; Guillera-Arroita et al. 2015). SDM can also be used more practically to assess habitat suitability (Abade, Macdonald, and Dickman 2014b), prioritize survey locations to discover unknown populations of rare species (Guisan et al. 2006), and develop conservation recommendations (Farhadinia et al. 2015). SDM algorithms are almost exclusively discriminatory, relying on contrast data to compare with species presence records. However, a key barrier to effective SDM is the ongoing uncertainty over the proper method of selecting PsA data when true absences are missing. Numerous PsA selection strategies have been proposed and compared (e.g.

Barbet-Massin et al. 2012), although these strategies should be viewed from an ecological as well as statistical context.

Previous research has established that how PsA data are selected affects distribution model output like spatial predictions and variable importance (Chefaoui and Lobo 2008; Stokland, Halvorsen, and Støa 2011; Merow, Smith, and Silander 2013). In Chapter 4, I use five PsA selection strategies to demonstrate how model output, but more importantly, ecological inference, differed in a case study examining cheetah range decline in East Africa. I proposed two novel PsA selection strategies (drawing PsA specifically from the resident or extirpated range of the species). I found that some proposed methods for PsA selection make little ecological sense and are therefore not recommended for this purpose. Comparing the method where PsA were drawn from the full extent of the study area (Full) and the species extirpated range (Extirpated) provided a new approach to explore the factors associated with cheetah presence and decline. In particular, comparison of the species response curves allowed for unique insight into the species response to changing conditions (Figure 4.5). Overall, researchers should be aware that the PsA selection process affects the modeling question, and hence that PsA should be selected in a way to ensure a match between the modeling and research questions, rather than maximizing evaluation metrics. Another important conclusion is that the new Extirpated method makes use of expert-based mapping processes and presence data, data widely available for many species, and is effective in mapping range decline.

In Chapter 5, I explored range decline for four large carnivores (cheetah, African wild dog, leopard, lion) in East Africa using two PsA selection strategies from the previous chapter that best matched this research question, and paired them with a new method. With these three PsA strategies, I tested three predictions: 1) increasing human populations densities and percentage of human land cover threaten large carnivore persistence, 2) carnivores with larger home ranges are more susceptible to habitat loss than those with smaller home ranges (specifically, cheetahs are most vulnerable, and leopards are least vulnerable) and 3) human land cover is a more important restriction on carnivore distributions than human population density.

Modeling results supported all three predictions. Increasing human populations and human land cover threaten large carnivore persistence as demonstrated by uniform and strongly negative species' response curves (Figure 5.6). In addition, human land cover restricted carnivore distributions to a greater degree than human population density, as it was almost universally the most important predictor for all species and PsA selection strategies (Table 5.8). Prediction 2 was largely supported as well: cheetahs, with the largest home range, and leopards, with the smallest home range, were the most and least susceptible to increases in human land cover, respectively (Figure 5.7). In some contradiction to the prediction, lions with smaller home ranges than either the wild dog or cheetah, had similar susceptibility to human land cover as the cheetah.

The novel PsA selection strategies allowed me to test several hypotheses, but they also have some practical limitations. In some cases, only one or a few PsA options may be appropriate, as they are dependent on the research aim. In addition, large carnivores distribution is relatively well known which enables PsA to be drawn from their historical distribution or parts thereof (such as former habitat) and for spatial predictions to be evaluated against existing range maps. However, for some species of concern, historic or current distributions are poorly known if at all.

The general recommendation for a standard SDM estimating a species distribution is to draw PsA from a tight buffer around the historical range of a species while including areas potentially reachable by dispersal (VanDerWal et al. 2009; Elith et al. 2011; Merow, Smith, and Silander 2013). Yet, selecting PsA only from the extirpated range of a species, as proposed in Chapter 4, means that a full response curve for biophysical variables is unlikely. Samples are only taken from within part of the species range, hence the full range of tolerated environmental conditions are not encountered (Austin 2007). However this is similarly true for studies conducted at smaller extents than the historical distribution of the species (i.e. an extent determined by a political boundary). The importance of this on model conclusions requires further study.

In addition, although comparing results from different PsA strategies can be useful, the outputs of distribution models using PsA data have low information content (Guillera-Arroita et al. 2015). Models using presence-absence records or occupancy data, neither of which rely on PsA data, enable distribution models to estimate true relative likelihood of occurrence, and probability of occurrence. Without absences, prevalence cannot be determined, and hence models using PsA do not predict probability of occurrence, but the relative likelihood of species observation. This means that the model does not give actual probability of occurrence, only the probability of occurrence relative to the survey effort. Thus, these approaches cannot differentiate between a rare species that is well surveyed or a common species that is under-surveyed. In some applications, such as estimation of species richness, higher information content outputs are necessary (Guillera-Arroita et al. 2015).

Overall, Chapters 4 and 5 together demonstrate the importance of considering PsA selection strategies from an ecological perspective, how different strategies affect the modeling question, and that the modeling and research question should be matched. This is true of any modeling yet has rarely been acknowledged in the SDM literature. The modeling approach developed here i.e. comparing alternative PsA selection strategies for the same species, is novel and provided some insight into applying SDMs more generally. The different strategies largely reaffirmed support for the ecological conclusions and contrasting response curves. Different strategies provided additional insight into species' response to changing conditions. Finally, of potential broad interest, I found that training models with PsA from former habitat led to more accurate predictions of current habitat. Researchers could apply these techniques more widely to other species of concern.

## 7.4 Spatial risk modeling case study with depredation records

Human actions can and do negatively impact wildlife. But human actions can also be neutral or positive. Real or perceived HWC can affect human attitudes and actions toward wildlife, potentially leading to direct persecution of wildlife (Hazzah et al. 2013; Dickman and Hazzah 2016). Direct take or killing of wildlife is a leading threat to carnivores (Ripple et al. 2014). Better understanding of the drivers of HWC, and amelioration of its impacts may engender more positive attitudes and actions towards wildlife (Dickman et al. 2014). In that context, researchers are increasingly applying discriminatory distribution models to a range of HWC data to create spatial risk models and identify drivers for things like carnivore attacks on livestock, or crop raiding.

Depredation risk models have the potential to ameliorate conflict if predictors of conflict are understood and spatial hotspots of conflict can be identified and avoided by livestock. Yet, two key aspects identified in the larger SDM literature have largely been neglected when applied to HWC data. These aspects are the initial consideration of biological and environmental factors that relate to the research aim, and the selection of the background extent. I discuss these issues in Chapter 6 in the context of providing an improved framework for conducting spatial risk modeling.

Conclusions from depredation risk models may be obfuscated or degraded when ecological factors that can impact the research aim are unaccounted for. Within the SDM literature, researchers have noted that characteristics such as gender (Conde et al. 2010), season (Takahata et al. 2014), or if the individual is resident or dispersing (Jackson et al.

2016) can influence habitat modeling outcomes. I demonstrate that these, and similar, issues can affect results of spatial risk modeling using a case study of livestock depredations in northern Tanzania. For instance, modeling outputs for location, season, and livestock were significantly and substantially different when accounting for these factors (e.g. Tables 6.3 and 6.4). Overall, I argue that before conducting a spatial risk model, researchers should consider a number of factors that may relate to the research aim (e.g. location, livestock species; Table 6.1).

Drawing on conclusions from Chapter 4, I also demonstrate the necessity of drawing a tight background extent to exclude areas where livestock and carnivores do not intersect. There are similar warnings regarding the analysis extent in the broader SDM literature (Phillips et al. 2009; VanDerWal et al. 2009), as well as in the field of crime science (Ratcliffe 2010). When PsA are drawn from areas where carnivores and livestock do not interact, this affects the contrast between areas where carnivores could, and areas where carnivores ultimately do, predate on livestock. A model with an extent much larger than the interaction zone between carnivores and livestock will emphasize broad-scale features and downplay fine-scale features that contribute to depredation risk. Ultimately, the research aim should be carefully matched to the modeling question, which in turn determines the choice of extent, whether any regions should be masked out, and how PsA are drawn.

Separately, I demonstrate that carnivores attack livestock in "risky" habitats (Appendix 6A). Many carnivores are known to select habitat for hunting according to prey accessibility rather than prey abundance (Hopcraft, Sinclair, and Packer 2005; Balme, Hunter, and Slotow 2007). Certain terrain, biotic, or human factors can influence prey accessibility; the presence of a ravine or dense vegetation may decrease a predator's chance of detection, and decrease the prey's ability to escape once the predator has been detected (Laundré, Hernández, and Ripple 2010). Although probability of attack decreased strongly with increasing tree cover overall (Figure 6.5), probability of attack increased strongly as the amount of local tree cover increased relative to the surrounding area (Figure 6.12). In addition, for carnivores to coexist with humans, they must have refuges where they can hide from humans and domestic animals (Oriol-Cotterill et al. 2015b). These refuges, in space and time, should be areas where humans infrequently visit or where chance of detection is lower. Indeed, some of the characteristics that make it a good refuge may also make prey more accessible. Thus, it's possible that during the day when humans and domestic animals are present on the landscape, that most carnivores retreat to refuges and that the same factors that make it a good refuge may also improve the chances of a successful hunt. Thus, these areas are doubly important to avoid as humans and livestock move and graze on the landscape.

## 7.5 Some common challenges

Various limitations and caveats have been discussed within each chapter. However, there are some common challenges and issues, particularly regarding SDM, that reoccur in several chapters and these are summarized here.

Data limitations for both distribution and conflict models are an obvious issue. For instance, accurate geospatial information on prey, competitor and livestock densities would have been useful. Although NDVI was used as a proxy for prey densities, substantial poaching occurs throughout East Africa reducing prey (and carnivore) densities below carrying capacity (Lindsey et al. 2013; Bauer et al. 2015; Rentsch and Packer 2015). More accurate estimates of prey densities, although rarely available, should improve habitat suitability modeling. Incorporating competitive effects into SDM is also recommended (Austin 2002; Araújo and Guisan 2006). I tested the inclusion of the lion distribution model into the wild dog model but the results did not agree with biological knowledge and I suspect the competitive effects are more dynamic and local than I could model. Finally, geospatial information on livestock densities could also enhance habitat suitability modeling as an indication of the intensity of pastoralist impacts. Of course, livestock densities and movement patterns, and the location of all livestock corrals, would also be helpful in spatial risk modeling.

Another consistent limitation was the lack of high-resolution environmental data for the distribution modeling chapters. Although many data sets were at 250 m resolution or better (e.g. MODIS, Worldpop, elevation), the lowest resolution data sets often set the grain of the analysis. These data include WorldClim (30 arc seconds) and ISRIC-WISE soil grids (1 km). Additionally, the GE Grids product was also at a coarser grain (0.01 degree), although I set this resolution as a balance between resolution and the extent of the analysis (~1.7 million km$^2$). I would have preferred to conduct GE Grids at around 500 m resolution as that would have better captured heterogeneous croplands. I did not do this as it would have increased the number of grids fourfold and greatly increased the time required. The spatial risk modeling was analyzed at a 250 m grain; a resolution closer to 100 m or smaller would likely improve results. Stalking distances of leopards (Bothma, van Rooyen, and le Riche 1997) and cheetah (Schaller 1968) are generally <100 m, and for some lions are only ~30 m (Stander 1992). Thus, geospatial information at these scales on vegetation and topography is important in studying hunting behavior.

More fine-scale information on the species/conflict data would have been useful too. For instance, biological knowledge like whether occurrence data represented transient or resident animals, adults or juveniles etc. Information on whether a sighting included dependent young would have enabled more confident identification of permanent carnivore range. Additionally, if a sighting represented a transient individual, then that data point could have been down-weighted or removed from the analysis to better capture permanent or resident range. If enough transient data had been gathered, I could have compared habitat selection between transient and resident individuals and consequently model habitat corridors for dispersal (Elliot et al. 2014b; Jackson et al. 2016). Similarly, if a transient (or dispersing) individual was responsible for a depredation event, this could have allowed analysis into the relative threats of resident vs. transient animals. Obtaining high spatial and temporal resolution movement data on carnivores and livestock within the same vicinity is likely to be substantially more difficult. However, if carnivores and livestock herds were tracked, this would provide insight into what turns interactions deadly (see Laporte et al. 2010 and Muhly et al. 2010).

Additional geospatial data on reproductive and mortality events could have also enhanced understanding of the predictors of carnivore persistence and decline. These data are more directly relevant to determining whether an area is a population source or sink, and ultimately what types of areas achieve stability. For instance, De Angelo et al. (2013) designed “two-dimensional” distribution models, one considering habitat suitability and the other survival (aka threats). They then spatially compared these models and divided the study area into four categories based on (Naves et al. 2003): attractive sinks, sinks, core areas, and refuges. This process enabled them to prioritize where interventions were necessary to prevent population declines.

These additional data would of course be useful in suitability modeling of the large carnivores, but in some sense, simulated data with a virtual species could have been more appropriate to demonstrate the importance of the study extent and method of PsA selection (such as Barbet-Massin et al. 2012). The species presence data suffered from sample selection bias and unknown prevalence. However, using a real species does not detract from the conceptual arguments about PsA generation strategies, and may have enabled greater insight, as it has a known ecology.

Overall, finer-scale data could lead to modeling that better identifies mechanisms, e.g. mechanisms behind depredation risk or habitat selection. Yet, modeling at larger extents can also be valuable in terms of being easier to apply and results that are more directly related to management. This suggests that the real need for distribution modeling at the extent of East Africa may be more accurate and relevant geospatial predictors rather than more resolved data. On the other hand, more fine-scale and accurate predictors would be valuable in understanding the mechanisms behind depredation risk in northern Tanzania. Further work using more geographically dispersed depredation records could provide more useful management recommendations regarding depredation risk at larger extents.

## 7.6 Next research steps

While overcoming some of the data limitations listed above may prompt a fresh look into the analysis, there are many other ways to move forward. Some are more immediate, and others broader and more long-term.

There are a number of potential next steps for the carnivore distribution modeling:

- Wrogemann (1975) compiled a list of cheetah sightings in East Africa from 1955-1964. Running a distribution model from that time period and comparing it to the contemporary model could provide more insight into how the cheetah niche has changed over time particularly in relationship to the expansion of human populations.
- For all carnivores, running a series of distribution models under future climate and land use scenarios would help estimate potential changes in habitat suitability and hence provide an understanding of persistence.
- For all carnivores, exploring the transferability of the models to other parts of their range, either in part or to all of Africa. There are few existing distribution models for these species in East Africa or other parts of their range. Transferring the existing model to new areas and evaluating its appropriateness could provide insight into how conditions may differ between regions and the generality of the model.
- Using new presence data (after 2015), or presence data from different sources (i.e. iNaturalist) as evaluation data.

- Creating a mechanistic niche model by incorporating physiological data (Kearney and Porter 2009). This may allow improved predictions into novel conditions, and provide better insight into predictors affecting persistence or decline.

For spatial risk modeling, potential next steps include:

- Gather fine-scale environmental layers and re-running the SDM analysis using a different and improved subset of predictors such as: land cover and vegetation, elevation and topography, livestock movement and/or densities, locations of all bomas/corrals, an improved roads data set, and prey densities and distribution. Or if some of these are not available, then evaluating and incorporating useful proxy data sets.
- Developing a collaborative initiative with other conservation organizations to establish a regional HWC database for East Africa and running a regional spatial risk model.
- Investigate incorporating carnivore habitat suitability output as an input.
- Comparing results of spatial association, interpolation, and correlative distribution modeling methods.
- Evaluating the results regarding attacks in risky habitats by splitting pasture conflicts by time of day, and checking the sensitivity of the results by comparing with different vegetation data sets and scales.
- Investigate the effectiveness of different methods used to prevent livestock depredation.

A broader issue that warrants much greater attention is how carnivores adapt to human-dominated landscapes. Clearly both social (including legal and political) as well as environmental perspectives should be addressed. Many carnivores are habitat generalists and quite flexible in their habits and dietary requirements. Large carnivores are returning in many parts of Europe (Chapron et al. 2014). This suggests that adaptation to alternative habitats is possible and therefore human land cover may be a relatively less important threat for generalist carnivores in developed economies (Nowell and Jackson 1996). Mech (1995) captures this issue when he describes how gray wolves were initially characterized as creatures of the wilderness. He compiled multiple research articles that suggested wolves could only live in areas with road densities lower than 0.6 km per km$^2$. This supposed threshold was used for management purposes but was repeatedly revised and raised until wolves were found to venture within 30 miles of a major metropolitan area in the US. There appears to be a similar story for the Eurasian lynx, as they have increasingly tolerated and included human land cover into home ranges (Basille et al. 2009; Bouyer et al. 2014; Bouyer et al. 2015). While habitat loss is frequently cited as a threat to carnivore species, it appears what actually constitutes habitat loss is not very well understood. In light of that Bouyer et al. (2014) recommend trying to determine tolerance limits of carnivores to human impacts. But this must be carefully done so as to not repeat the problems characterized in Mech's wolf storyline.

A broader issue regarding SDM is understanding and quantifying the potential impact of violating the assumption that species are in equilibrium with their environment. This is one of the fundamental assumptions of SDM (Franklin 2009; Elith and Leathwick 2009). This is mostly considered an issue when using models for extrapolation purposes, e.g. species invasion (Elith and Leathwick 2009). Yet, some have suggested we are in the 6th mass extinction event (Ceballos et al. 2015) and a new geologic epoch, the Anthropocene (Crutzen 2002), due to pervasive human influences across the globe. Very few species' distributions have not already been constrained or undergone recent changes due to human activity. Thus, current presence data are already biased by human impacts on species and habitats and I believe this is not simply an issue when using SDM for extrapolation as suggested in the literature (Elith, Kearney, and Phillips 2010). So if we have already violated the equilibrium assumption, to what extent does that affect the meaning of the SDM output? For instance, applying a distribution model using contemporary locations of mature redwood trees is not likely to be very appropriate (or useful), as conditions have changed substantially since they sprouted a thousand or more years ago. Although that may be an extreme case, the broader issue is to what extent this may be true for all (or nearly all) species. A potential solution would be to only model areas where reproduction is occurring as a sign of current suitability. However, even in this case, reproduction may occur in areas that are "attractive sinks" (Delibes, Gaona, and Ferreras 2001). Thus predictions of suitability must be interpreted carefully. Similarly, researchers often apply distribution models at an extent smaller than the full historical extent plus areas reachable by dispersal, such as e.g. a political unit. In these cases, the models are fundamentally not investigating the distribution of the species, but rather something akin to occupancy or habitat suitability. If it is occupancy that is being modeled, then this suggests that only particular data and methods are appropriate for modeling occupancy rather than the realized niche.

## 7.7 The conservation context & recommendations

This research has generated a number of results and insights that can be directly linked to conservation. While some of these conservation recommendations were not explicitly tested or evaluated, they complement the results of this dissertation. The audience for these recommendations includes NGOs, governmental entities, and private citizens/communities. Initially, I highlight a number of recommendations before returning to a more nuanced description of them. The recommendations fall under three broad categories, land use planning, management & interventions, and biological surveys:

- Land use planning:
    - Set up land use plans at the regional or national level to allow for growing human populations and expanding anthropogenic land uses while minimizing isolation of protected areas and wildlife populations.
    - Enforce land use plans at the village level to designate conservation areas and grazing/forest banks.
    - Assist in obtaining legal land rights, with a focus on communal ownership in pastoral systems.
    - Disincentive sedentarization within pastoral landscapes.
    - Immediately move to identify and protect remaining wildlife corridors from land use conversion.
    - Consider establishment of additional protected areas where carnivore suitability is high and current levels of land use conversion/natural resource extraction is negligible.
- Management and interventions:
    - Promote traditional forms of pastoralism that allow seasonal movement of livestock, healthy grazing lands, and maintain prey populations and areas of dense vegetation.
    - Seek to better understand depredation risks and implement effective HWC mitigation efforts at larger extents.
    - Secure protected areas from illegal encroachment and poaching.
- Biological surveys:
    - Survey areas of high carnivore suitability outside areas of known range to document if populations exist and ultimately better protect them.
    - Survey areas within known carnivore range that are identified as under severe threat to check their status, and move to protect them if feasible.

Africa’s human population was less than a quarter of a billion people in 1950 (229 million), but has increased rapidly. By 2013, Africa’s human population was 1.1 billion, and by 2050 is expected to rise to between 2.1 and 2.7 billion (UN 2013). That is a near 10-fold expansion in 100 years. In the coming decades, East Africa will change dramatically with the populations of Burundi, Tanzania and Uganda, projected to quintuple between 2000 and 2100 (UN 2013). Pressure on natural resources, including wildlife, will be immense. Already nearly 30% of the terrestrial land surface in East Africa has been converted to human-dominated land cover, and this is approaching 90% in Burundi and Rwanda (Table 3.2).

In contrast, East Africa contains over 100 terrestrial PAs (excluding Forest Reserves and Wildlife Management Areas which have some of the least protection), which cover 14.4% of the terrestrial land surface (Jason Riggio et al. unpublished data). In some countries, like Tanzania, the percentage is higher at just over 20%. This nearly meets, and in some cases exceeds the Aichi Biodiversity Target 11 of protecting 17% of land by 2020. While PAs could be better placed to protect overall biodiversity (i.e. small-ranged or endemic species), the PA system is extensive, includes significant populations of threatened and endangered species (ibid.), and substantial areas of important carnivore habitat. To this point the system has largely been effective in preventing encroachment, although it’s been less effective in protecting wildlife populations. Therefore, as human pressures increase in the coming decades, maintaining the integrity and increasing the resilience of the PA system, in large part by preventing poaching and illegal encroachment, will be critical.

The effectiveness of PAs for protecting species and habitat can be influenced by activities outside of them. Edge effects via interaction with people and domestic animals can threaten wildlife populations and thereby undermine the effectiveness of PAs. In many cases, human-caused mortality is the single largest cause of mortality for carnivore populations in PAs (Woodroffe and Ginsberg 1998). An ongoing challenge for PAs is to help redistribute their economic benefits to local communities to offset costs of restricted access to natural resources (Naughton-Treves, Holland, and Brandon 2005). But strong economic performance can in turn cause higher rates of immigration than in the surrounding areas (Wittemyer et al. 2008). Increased population and development can cause the complete isolation of PAs. Increasing ‘hard’ edges between PAs and human-dominated lands can turn historical wildlife movement patterns into incursions of crop and village land, and ultimately precipitate proposals to fence the PA. Fencing can solve some issues (such as wildlife wandering into crop or pasture lands) while causing a potentially larger suite of issues (Woodroffe, Hedges, and Durant 2014). Maintaining corridors to preserve genetic connectivity of PAs is essential and preserving existing ones is easier than attempting to restore connections after they have been lost. An additional challenge posed by the insularization of PAs is that for some wide-ranging species such as cheetah and wild dog, even the largest PAs in the region cannot contain viable populations. Predator-proof fencing can cause particular challenges for species like these (Durant et al. 2015).

Yet, in some sense, wide-ranging carnivores and herbivores that may be most affected by fencing are also some of the most 'deserving'. These mobile, conflict-causing wildlife species are more likely to go extinct than those with smaller home ranges due to increased interaction with humans (Woodroffe and Ginsberg 1998). Thus, in the absence of fencing, preventing and mitigating HWC will be essential in buffer areas and habitat corridors. Innovative solutions are needed to reduce the impacts of crop raiding, carnivore depredations and other negative effects from wildlife (but see Dickman, Macdonald, and Macdonald 2011). A suite of activities are often used to change opinion and ultimately behavior toward wildlife (see Tanzanian and Kenyan NGOs like the Tanzania People and Wildlife Fund, Ruaha Carnivore Program, and Ewaso Lions). Effective organizations blend education and raising awareness, with activities that increase benefits (economic or otherwise), and decrease costs of living with wildlife, to spur attitude and behavior change. Ultimately the communities living with wildlife must see the value in tolerating wildlife so that every encounter with an animal does not turn into "conflict." Mitigating HWC will be essential for species like cheetah because PAs are too small to maintain viable populations. Even compared to other wide-ranging carnivores like the lion, cheetah and wild dog tend to live at densities ~ 1/10th that of lions, hence requiring much larger areas to achieve similar population sizes (Durant 2007).

In reality, all of the large carnivores discussed here may be termed protection-reliant (Durant et al. 2016). Protection-reliant species are already threatened but face additional threats because a substantial portion of their population persists outside PAs where threat is higher. Thus, while PAs may hold substantial populations, conservation for protection-reliant species should emphasize improving growth rates in unprotected landscapes, e.g. minimizing edge effects. This entails promoting coexistence, and managing human development and activities in an efficient manner. Zoning can be effective such that particular activities can be included/excluded from an area, and that residents have realistic expectations about interactions with wildlife. Buffer zones and corridors around PAs could be zoned for pastoralism and programs alleviating HWC for instance can be concentrated to have the greatest impact. Broadly speaking, programs that maintain traditional pastoral systems, and prevent sedentarization and subdivision of grazing lands, are also critical. With the astounding population projections, planning for growth will be foundational to preserving PA integrity, maintaining wildlife populations, and preventing their isolation.

# References cited

Aarts, G., Fieberg, J. & Matthiopoulos, J., 2012. Comparative interpretation of count, presence-absence and point methods for species distribution models. *Methods in Ecology and Evolution*, 3(1), pp.177–187.

Abade, L., Macdonald, D.W. & Dickman, A.J., 2014a. Using Landscape and Bioclimatic Features to Predict the Distribution of Lions, Leopards and Spotted Hyaenas in Tanzania's Ruaha Landscape. *PloS one*, 9(5), p.e96261.

Abade, L., Macdonald, D.W. & Dickman, A. J., 2014b. Assessing the relative importance of landscape and husbandry factors in determining large carnivore depredation risk in Tanzania's Ruaha landscape. *Biological Conservation*, 180, pp.241–248.

Acevedo, P. et al., 2012. Delimiting the geographical background in species distribution modelling. *Journal of Biogeography*, 39(8), pp.1383–1390.

Aguirre-Gutiérrez, J. et al., 2013. Fit-for-Purpose: Species Distribution Model Performance Depends on Evaluation Criteria - Dutch Hoverflies as a Case Study. *PLoS ONE*, 8(5).

Ahlers, T. et al. eds., 2014. *Africa 2050: Realizing the Continent's Full Potential*, Oxford University Press.

Ahmed, S.E. et al., 2015. Scientists and software - surveying the species distribution modelling community J. Elith, ed. *Diversity and Distributions*.

Albert, C.H. et al., 2010. Sampling in ecology and evolution - bridging the gap between theory and practice. *Ecography*, 33(6), pp.1028–1037.

Alinovi, L., Di Gregorio, A. & Latham, J., 2000. The FAO AFRICOVER-Eastern Africa Project. Land cover assessment based on remote sensing. General description.

Altmann, A. et al., 2010. Permutation importance: a corrected feature importance measure. *Bioinformatics*, 26(10), pp.1340–1347.

Anderson, R.P. & Gonzalez Jr., I., 2011. Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with Maxent. *Ecological Modelling*, 222(15), pp.2796–2811. Available at: http://dx.doi.org/10.1016/j.ecolmodel.2011.04.011.

Anderson, R.P. & Raza, A., 2010. The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: preliminary tests with montane rodents (genus Nephelomys) in Venezuela. *Journal of Biogeography*, 37(7), pp.1378–1393.

De Angelo, C. et al., 2013. Understanding species persistence for defining conservation actions: A management landscape for jaguars in the Atlantic Forest. *Biological Conservation*, 159, pp.422–433.

De Angelo, C., Paviolo, A. & Di Bitetti, M., 2011. Differential impact of landscape transformation on pumas (*Puma concolor*) and jaguars (*Panthera onca*) in the Upper Paraná Atlantic Forest. *Diversity and Distributions*, 17(3), pp.422–436.

Anwar, M.B. et al., 2011. Food habits of the snow leopard *Panthera uncia* (Schreber, 1775) in Baltistan, Northern Pakistan. *European Journal of Wildlife Research*, 57(5), pp.1077–1083.

Araújo, M.B. & Guisan, A., 2006. Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, 33(10), pp.1677–1688.

Arnold, J.E.M., Köhlin, G. & Persson, R., 2006. Woodfuels, livelihoods, and policy interventions: Changing Perspectives. *World Development*, 34(3), pp.596–611.

Athreya, V. et al., 2014. A cat among the dogs: leopard *Panthera pardus* diet in a human-dominated landscape in western Maharashtra, India. *Oryx*.

Athreya, V. et al., 2013. Big cats in our backyards: persistence of large carnivores in a human dominated landscape in India. *PloS one*, 8(3), p.e57872.

Athreya, V. et al., 2015. Spotted in the News: Using Media Reports to Examine Leopard Distribution, Depredation, and Management Practices outside Protected Areas in Southern India. *Plos One*, 10(11), p.e0142647.

Austin, M., 2007. Species distribution models and ecological theory: A critical assessment and some possible new approaches. *Ecological Modelling*, 200(1–2), pp.1–19.

Austin, M.P., 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling*, 157(2–3), pp.101–118.

Baguette, M. et al., 2013. Individual dispersal, landscape connectivity and ecological networks. *Biological Reviews*, 88(2), pp.310–326.

Bailey, T.N., 1993. *The African leopard: ecology and behavior of a solitary felid*, New York, USA: Columbia University Press.

Baldus, R.D. & Cauldwell, A.E., 2004. *Tourist Hunting and it's Role in Development of Wildlife Management Areas in Tanzania*, Dar es Salaam, TZ.

Balme, G.A., Slotow, R. & Hunter, L.T.B., 2009. Impact of conservation interventions on the dynamics and persistence of a persecuted leopard (Panthera pardus) population. *Biological Conservation*, 142(11), pp.2681–2690.

Balme, G., Hunter, L. & Slotow, R., 2007. Feeding habitat selection by hunting leopards Panthera pardus in a woodland savanna: prey catchability versus abundance. *Animal Behaviour*, 74(3), pp.589–598.

Barbet-Massin, M. et al., 2012. Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution*, 3(2), pp.327–338.

Barbosa, A.M. et al., 2013. New measures for assessing model equilibrium and prediction mismatch in species distribution models. *Diversity and Distributions*, 19(10), pp.1333–1338.

Barua, M., Bhagwat, S. A. & Jadhav, S., 2013. The hidden dimensions of human–wildlife conflict: Health impacts, opportunity and transaction costs. *Biological Conservation*, 157, pp.309–316.

Barve, N. et al., 2011. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, 222(11), pp.1810–1819.

Basille, M. et al., 2009. What shapes Eurasian lynx distribution in human dominated landscapes: selecting prey or avoiding people? *Ecography*, 32(4), pp.683–691.

Bateman, P.W. & Fleming, P.A., 2012. Big city life: Carnivores in urban environments. *Journal of Zoology*, 287(1), pp.1–23.

Bauer, H., Chapron, G., et al., 2015. Lion (*Panthera leo*) populations are declining rapidly across Africa, except in intensively managed areas. *Proceedings of the National Academy of Sciences*, p.201500664.

Bauer, H., Packer, C., et al., 2015. Panthera leo. The IUCN Red List of Threatened Species 2015: e.T15951A79929984.

Bauer, H. & Iongh, H.H. De, 2005. Lion (*Panthera leo*) home ranges and livestock conflicts in Waza National Park , Cameroon. pp.208–214.

Becker, M. et al., 2013a. Evaluating wire-snare poaching trends and the impacts of by-catch on elephants and large carnivores. *Biological Conservation*, 158, pp.26–36.

Becker, M.S. et al., 2013b. Estimating past and future male loss in three Zambian lion populations. *The Journal of Wildlife Management*, 77(1), pp.128–142.

Bekure, S. et al. eds., 1991. *Maasai herding: An analysis of the livestock production system of Maasai pastoralists in eastern Kajiado District, Kenya*, Addis Ababa, Ethiopia: International Livestock Centre for Africa.

Belbachir, F. et al., 2015. Monitoring rarity: the critically endangered saharan cheetah as a flagship species for a threatened ecosystem. *PloS one*, 10(1), p.e0115136.

Benjaminsen, T.A. & Bryceson, I., 2012. Conservation, green/blue grabbing and accumulation by dispossession in Tanzania. *Journal of Peasant Studies*, 39, pp.335–355.

Benton, T.G., Vickery, J.A. & Wilson, J.D., 2003. Farmland biodiversity: Is habitat heterogeneity the key? *Trends in Ecology and Evolution*, 18(4), pp.182–188.

Berger, K.M. & Gese, E.M., 2007. Does interference competition with wolves limit the distribution and abundance of coyotes? *The Journal of animal ecology*, 76(6), pp.1075–85.

Berger, K.M., Gese, E.M. & Berger, J., 2008. Indirect effects and traditional trophic cascades: a test involving wolves, coyotes, and pronghorn. *Ecology*, 89(3), pp.818–28.

Beschta, R.L. & Ripple, W.J., 2009. Large predators and trophic cascades in terrestrial ecosystems of the western United States. *Biological Conservation*, 142(11), pp.2401–2414.

Bininda-Emonds, O.R., Gittleman, J.L. & Purvis, a, 1999. Building large trees by combining phylogenetic information: a complete phylogeny of the extant Carnivora (Mammalia). *Biological reviews of the Cambridge Philosophical Society*, 74(2), pp.143–75.

Bissett, C. & Bernard, R.T.F., 2007. Habitat selection and feeding ecology of the cheetah (*Acinonyx jubatus*) in thicket vegetation: is the cheetah a savanna specialist? *Journal of Zoology*, 271, p.060818015547003.

Blaschke, T., 2010. Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(1), pp.2–16.

de Boer, W.F. et al., 2010. Spatial distribution of lion kills determined by the water dependency of prey species. *Journal of Mammalogy*, 91(5), pp.1280–1286.

Bohm, T. & Honer, O.R., 2015. Crocuta crocuta. The IUCN Red List of Threatened Species 2015: e.T5674A45194782. Available at: www.iucnredlist.org/details/5674/0.

Van Bommel, L. et al., 2007. Factors affecting livestock predation by lions in Cameroon. *African Journal of Ecology*, 45(4), pp.490–498.

Bongaarts, J., 2009. Human population growth and the demographic transition. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 364(1532), pp.2985–90.

Bontemps, S. et al., 2011. *GLOBCOVER 2009 - product description and validation report*.

Boria, R.A. et al., 2014. Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. *Ecological Modelling*, 275, pp.73–77.

Bothma, J. du P., van Rooyen, N. & le Riche, E.A.N., 1997. Multivariate analysis of the hunting tactics of Kalahari leopards. *Koedoe*, 40(1), pp.41–56.

Bouché, P. et al., 2012. Game over! Wildlife collapse in northern Central African Republic. *Environmental Monitoring and Assessment*, 184(11), pp.7001–7011.

Bouyer, Y. et al., 2015. Eurasian lynx habitat selection in human-modified landscape in Norway: Effects of different human habitat modifications and behavioral states. *Biological Conservation*, 191, pp.291–299.

Bouyer, Y. et al., 2014. Tolerance to anthropogenic disturbance by a large carnivore: the case of Eurasian lynx in south-eastern Norway. *Animal Conservation*.

Boyce, M.S. et al., 2002. Evaluating resource selection functions. *Ecological Modelling*, 157, pp.281–300.

Boyce, M.S., 2006. Scale for resource selection functions. *Diversity and Distributions*, 12(3), pp.269–276.

Brady, N.C. & Weil, R.R., 1996. *The nature and properties of soil*, Prentice-Hall Inc.

Brashares, J.S. et al., 2011. Economic and geographic drivers of wildlife consumption in rural Africa. *Proceedings of the National Academy of Sciences*, 108(34), pp.13931–13936.

Breiman, L. et al., 1984. *Classification and regression trees*, CRC Press.

Breiman, L., 2001. Random Forests. *Machine Learning*, 45(1), pp.5–32.

van Breugel, P. et al., 2012. Potential natural vegetation map of Eastern Africa: an interactive vegetation v 1.1. Available at: http://vegetationmap4africa.org.

Bright, E.A. et al., 2015. Landscan.

Brockington, D. & Igoe, J., 2006. Eviction for Conservation: A Global Overview. *Conservation and Society*, 4(3), pp.424–470.

Broekhuis, F. et al., 2013. Risk avoidance in sympatric large carnivores: reactive or predictive? *The Journal of Animal Ecology*.

Broekhuis, F. & Gopalaswamy, A.M., 2016. Counting Cats : Spatially Explicit Population Estimates of Cheetah (Acinonyx jubatus) Using Unstructured Sampling Data, (May), pp.1–15.

Broennimann, O. et al., 2015. package ecospat.

Brown, J.L., 2014. SDMtoolbox: a python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses B. Anderson, ed. *Methods in Ecology and Evolution*.

Bucklin, D.N. et al., 2014. Comparing species distribution models constructed with different subsets of environmental predictors. *Diversity and Distributions*, pp.1–13.

Burdett, C.L. et al., 2010. Interfacing models of wildlife habitat and human development to predict the future distribution of puma habitat. *Ecosphere*, 1(1).

Burton, A.C. et al., 2011. Evaluating persistence and its predictors in a West African carnivore community. *Biological Conservation*, 144(9), pp.2344–2353.

Butchart, S.H.M. et al., 2010. Global Biodiversity: Indicators of Recent Declines. *Science*, 328, pp.1164–1169.

Butt, B., 2010. Seasonal space-time dynamics of cattle behavior and mobility among Maasai pastoralists in semi-arid Kenya. *Journal of Arid Environments*, 74(3), pp.403–413.

Carbone, C. & Gittleman, J.L., 2002. A common rule for the scaling of carnivore density.

*Science*, 295, pp.2273–2276.

Cardillo, M. et al., 2004. Human population density and extinction risk in the world's carnivores. *PLoS biology*, 2(7), p.E197.

Cardillo, M. et al., 2005. Multiple Causes of High Extinction Risk in Large Mammal Species. *Science*, 309(5738), pp.1239–1241.

Cardillo, M. et al., 2008. The predictability of extinction: biological and external correlates of decline in mammals. *Proceedings of the Royal Society B*, 275(1641), pp.1441–1448.

Caro, T., 1994. *Cheetahs of the Serengeti Plains: group living in an asocial species*, Chicago and London: University of Chicago Press.

Caro, T. & Davenport, T.R.B., 2015. Wildlife and wildlife management in Tanzania. *Conservation Biology*, 30(4), pp.716–723.

Caro, T.M., 2003. Umbrella species: critique and lessons from East Africa. *Animal Conservation*, 6(2), pp.171–181.

Caro, T.M. & Stoner, C.J., 2003. The potential for interspecific competition among African carnivores. *Biological Conservation*, 110(1), pp.67–75.

Carroll, C. & Miquelle, D.G., 2006. Spatial viability analysis of Amur tiger Panthera tigris altaica in the Russian Far East: the role of protected areas and landscape matrix in population persistence. *Journal of Applied Ecology*, 43(6), pp.1056–1068.

Carvell, C. et al., 1998. How Diana Climbed the Ratings at the Zoo. *Nature*, 395(6699), pp.213–213.

Ceballos, G. et al., 2015. Accelerated modern human – induced species losses: Entering the sixth mass extinction. *Science Advances*, 1.

Celesia, G.G. et al., 2009. Climate and landscape correlates of African lion (*Panthera leo*) demography. *African Journal of Ecology*, 48, pp.58–71.

Celik, E., 2015. package vita. Available at: https://cran.r-project.org/web/packages/vita.

Chamard, S., 2006. The History of Crime Mapping and Its Use by American Police Departments. *Alaska Justice Forum*, 23(3), pp.4–8.

Chapron, G. et al., 2014. Recovery of large carnivores in Europe's modern human-dominated landscapes. *Science*, 346, pp.1517–1519.

Chardonnet, P. et al., 2010. *Managing the conflicts between people and lion: Review and insights from the literature and field experience*, Rome.

Chase, M.J. et al., 2016. Continent-wide survey reveals massive decline in African savannah elephants. *PeerJ*, 4, p.e2354.

Chefaoui, R.M. & Lobo, J.M., 2008. Assessing the effects of pseudo-absences on predictive distribution model performance. *Ecological Modelling*, 210, pp.478–486.

Childes, S.., 1988. The past history, present status and distribution of the hunting dog

Lycaon pictus in Zimbabwe. *Biological Conservation*, 44(4), pp.301–316.

CIA, 2016. The World Factbook 2015-2016.

CIESIN, 2013. Global roads open access data set, Version 1, gROADS v1.

Coe, M.J., Cumming, D.H. & Phillipson, J., 1976. Biomass and production of large African herbivores in relation to rainfall and primary production. *Oecologia*, 22(4), pp.341–354.

Conde, D.A. et al., 2010. Sex matters: Modeling male and female habitat differences for jaguar conservation. *Biological Conservation*, 143(9), pp.1980–1988.

Consorte-McCrea, A., 2013. Relationships between the maned wolf and people. In *Ecology and Conservation of the Maned Wolf: Multidisciplinary Perspectives*. pp. 35–52.

Cooper, A.B., Pettorelli, N. & Durant, S.M., 2007. Large carnivore menus: factors affecting hunting decisions by cheetahs in the Serengeti. *Animal Behaviour*, 73(4), pp.651–659.

Costelloe, B. et al., 2016. Global Biodiversity Indicators Reflect the Modeled Impacts of Protected Area Policy Change. *Conservation Letters*, 9(1), pp.14–20.

Courchamp, F., Clutton-Brock, T. & Grenfell, B., 2000. Multipack dynamics and the Allee effect in the African wild dog, Lycaon pictus. *Animal Conservation*, 3(4), pp.277–285.

Cozzi, G. et al., 2012. Fear of the dark or dinner by moonlight? Reduced temporal partitioning among Africa's large carnivores. *Ecology*, 93(12), pp.2590–2599.

Craigie, I.D. et al., 2010. Large mammal population declines in Africa's protected areas. *Biological Conservation*, 143(9), pp.2221–2228.

Creel, S. & Creel, N.M., 1996. Limitation of African Wild Dogs by Competition with Larger Carnivores. *Conservation Biology*, 10(2), pp.526–538.

Creel, S. & Creel, N.M., 2002. *The African wild dog: behavior, ecology and conservation*,

Creel, S., Spong, G. & Creel, N., 2001. Interspecific competition and the population biology of extinction-prone carnivores. In J. L. Gittleman et al., eds. *Carnivore Conservation*. Cambridge, Uk: Cambridge University Press, p. 675.

Crooks, K.R., 2002. Relative Sensitivities of Mammalian Carnivores to Habitat Fragmentation. *Conservation Biology*, 16(2), pp.488–502.

Crooks, K.R. & Soulé, M.E., 1999. Mesopredator release and avifaunal extinctions in a fragmented system. *Nature*, 400, pp.563–566.

Crutzen, P.J., 2002. Geology of mankind. *Nature*, 415(23).

Cushman, S.A. et al., 2015. A multi-scale assessment of population connectivity in African lions (Panthera leo) in response to landscape change. *Landscape Ecology*.

Cutler, D.R. et al., 2007. Random forests for classification in ecology. *Ecology*, 88(11), pp.2783–2792.

Dalerum, F. et al., 2009. Diversity and depletions in continental carnivore guilds: implications for prioritizing global carnivore conservation. *Biology letters*, 5, pp.35–38.

Dalerum, F. et al., 2008. The potential for large carnivores to act as biodiversity surrogates in southern Africa. *Biodiversity and Conservation*, 17(12), pp.2939–2949.

Danko, D.M., 1992. Digital Chart of the World Project. *Photogrammetric Engineering and Remote Sensing*, 58(8), pp.1125–1128.

Darkoh, M.B.., 2003. Regional perspectives on agriculture and biodiversity in the drylands of Africa. *Journal of Arid Environments*, 54(2), pp.261–279.

Davidson, A.D. et al., 2009. Multiple ecological pathways to extinction in mammals. *PNAS*, 106(26), pp.10702–10705.

Davidson, Z. et al., 2013. Seasonal diet and prey preference of the African lion in a waterhole-driven semi-arid savanna. *PloS one*, 8(2), p.e55182.

Defourny, P. et al., 2008. *Accuracy Assessment of a 300 m Global Land Cover Map: The GlobCover Experience*,

DEFRA, 2014. *The Strategy for Achieveing Officially Tuberculosis Free Status for England*,

Delibes, M., Gaona, P. & Ferreras, P., 2001. Effects of an attractive sink leading to maladaptive habitat selection. *The American Naturalist*, 158(3).

Dickman, A., 2008. *Key Determinants of Conflict Between People and Wildlife, Particularly Large Carnivores, Around Ruaha National Park, Tanzania*. University College London.

Dickman, A.J. et al., 2014. Carnivores, culture and “contagious conflict”: Multiple factors influence perceived problems with carnivores in Tanzania’s Ruaha landscape. *Biological Conservation*, 178, pp.19–27.

Dickman, A.J., 2010. Complexities of conflict: the importance of considering social factors for effectively resolving human-wildlife conflict. *Animal Conservation*, 13(5), pp.458–466.

Dickman, A.J. & Hazzah, L., 2016. Money, Myths and Man-Eaters: Complexities of Human-Wildlife Conflict. In *Problematic Wildlife*. pp. 339–356.

Dickman, A.J., Macdonald, E. a & Macdonald, D.W., 2011. A review of financial instruments to pay for predator conservation and encourage human-carnivore coexistence. *Proceedings of the National Academy of Sciences of the United States of America*, 108(34), pp.13937–44.

DiMiceli, C.M. et al., 2011. Annual global automated MODIS vegetation continuous fields (MOD44B) at 250m spatial resolution for data years beginning day 65, 2000 - 2010, collection 5 percent tree cover.

Dolrenry, S. et al., 2014. A Metapopulation Approach to African Lion (*Panthera leo*)

Conservation A. L. Roca, ed. *PLoS ONE*, 9(2), p.e88081.

Dormann, C. et al., 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, 36(1), pp.27–46.

Dormann, C. et al., 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, 30(5), pp.609–628.

Douglas, L.R. & Alie, K., 2014. High-value natural resources: Linking wildlife conservation to international conflict, insecurity, and development concerns. *Biological Conservation*, 171, pp.270–277.

Durant, S. et al., 2015. Acinonyx jubatus. *The IUCN Red List of Threatened Species. e.T219A50649567*.

Durant, S., 2007. Range-wide Conservation Planning for Cheetah and Wild Dog. *Cat News*, 46, p.13.

Durant, S.M., 1998. Competition refuges and coexistence: an example from Serengeti carnivores. *Journal of Animal Ecology*, 67(3), pp.370–386.

Durant, S.M., Craft, M.E., et al., 2010a. Does size matter? An investigation of habitat use across a carnivore assemblage in the Serengeti, Tanzania. *Journal of Animal Ecology*, 79(5), pp.1012–22.

Durant, S.M. et al., 2014. Fiddling in biodiversity hotspots while deserts burn? Collapse of the Sahara's megafauna. *Diversity and Distributions*, 20(1), pp.114–122.

Durant, S.M., 2000. Living with the enemy: avoidance of hyenas and lions by cheetahs in the Serengeti. *Behavioral Ecology*, 11(6), pp.624–632.

Durant, S.M., Dickman, A.J., et al., 2010b. Past, present, and future of cheetahs in Tanzania: their behavioural ecology and conservation. In D. W. Macdonald & A. J. Loveridge, eds. *Biology and Conservation of Wild Felids*. Oxford University Press.

Durant, S.M. et al., 2016. The global decline of cheetah *Acinonyx jubatus* and what it means for conservation. *Proceedings of the National Academy of Sciences*.

van Dyk, G. & Slotow, R., 2003. The effects of fences and lions on the ecology of African wild dogs reintroduced to Pilanesberg National Park, South Africa. *African Zoology*, 38(1), pp.79–94.

East, R., 1999. *African Antelope Database 1998*, Gland, Switzerland and Cambridge, UK: IUCN.

East, R., 1984. Rainfall, soil nutrient status and biomass of large African savanna mammals. *African Journal of Ecology*, 22, pp.245–270.

Eck, J.E. et al., 2005. *Mapping Crime: Understanding Hot Spots*, Washington D.C.

Ehrlich, P.R. & Holdren, J.P., 1971. Impact of Population Growth. *Science*, 171(3977), pp.1212–1217.

Elith, J. et al., 2011. A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, 17(1), pp.43–57.

Elith, J. et al., 2006. Novel methods improve prediction of species ' distributions from occurrence data. *Ecography*, 2, pp.129–151.

Elith, J., Kearney, M. & Phillips, S., 2010. The art of modelling range-shifting species. *Methods in Ecology and Evolution*, 1(4), pp.330–342.

Elith, J. & Leathwick, J., 2007. Predicting species distributions from museum and herbarium records using multiresponse models fitted with multivariate adaptive regression splines. *Diversity and Distributions*, 13(3), pp.265–275.

Elith, J. & Leathwick, J., 2009. Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, …*, 40, pp.677–697.

Elliot, N.B., Cushman, S.A., Loveridge, A.J., et al., 2014a. Movements vary according to dispersal stage, group size, and rainfall: The case of the African lion. *Ecology*, 95(10), pp.2860–2869.

Elliot, N.B., Cushman, S.A., Macdonald, D.W., et al., 2014b. The devil is in the dispersers: Predictions of landscape connectivity change with demography. *Journal of Applied Ecology*, 51(5), pp.1169–1178.

Elton, C.S., 1927. *Animal Ecology*, New York, NY: The MacMillan Company.

Engler, R., Guisan, A. & Rechsteiner, L., 2004. An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *Journal of Applied Ecology*, 41(2), pp.263–274.

ESRI, 2014. ArcGIS Desktop: Release 10.2.1.

Estes, A.B. et al., 2012. Land-cover change and human population trends in the greater Serengeti ecosystem from 1984-2003. *Biological Conservation*, 147(1), pp.255–263.

Evans, J.S. et al., 2011. Modeling Species Distribution and Change Using Random Forest. In *Predictive Species and Habitat Modeling in Landscape Ecology: Concepts and Applications*. pp. 139–159.

Evans, J.S. & Murphy, M.A., 2015. package "rfUtilities."

Fahrig, L., 2002. Effect of habitat fragmentation on the extinction threshold: a synthesis. *Ecological Applications*, 12(2), pp.346–353.

Fahrig, L., 2003. Effects of habitat fragmentation on biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, 34, pp.487–515.

Fahrig, L., 2001. How much habitat is enough? *Biological Conservation*, 100, pp.65–74.

Fahrig, L., 1997. Relative effects of habitat loss and fragmentation on population extinction. *The Journal of Wildlife Management*, 61(3), pp.603–610.

Fanshawe, J.H., Frame, L.H. & Ginsberg, J.R., 1991. The wild dog—Africa's vanishing carnivore. *Oryx*, 25(3), pp.137–146.

Farhadinia, M.S. et al., 2015. Leveraging trans-boundary conservation partnerships: Persistence of Persian leopard (Panthera pardus saxicolor) in the Iranian Caucasus. *Biological Conservation*, 191, pp.770–778.

Farris, Z.J. et al., 2016. Patterns of spatial co-occurrence among native and exotic carnivores in north-eastern Madagascar. *Animal Conservation*, 19(2), pp.189–198.

Ficetola, G. et al., 2014. Predicting wild boar damages to croplands in a mosaic of agricultural and natural areas. *Current Zoology*, 60(2), pp.170–179.

Foley, J.A. et al., 2005. Global consequences of land use. *Science (New York, N.Y.)*, 309(5734), pp.570–4.

Fontúrbel, F.E. & Simonetti, J.A., 2011. Translocations and human-carnivore conflicts: problem solving or problem creating? *Wildlife Biology*, 17(2), pp.217–224.

Ford, A.T. et al., 2014. Large carnivores make savanna tree communities less thorny. *Science*, 346(6207), pp.346–349.

Forester, B.R., Dechaine, E.G. & Bunn, A.G., 2013. Integrating ensemble species distribution modelling and statistical phylogeography to inform projections of climate change impacts on species distributions. *Diversity and Distributions*, 19, pp.1480–1495.

Forrest, J.L. et al., 2012. Conservation and climate change: Assessing the vulnerability of snow leopard habitat to treeline shift in the Himalaya. *Biological Conservation*, 150(1), pp.129–135.

Fourcade, Y. et al., 2014. Mapping species distributions with MAXENT using a geographically biased sample of presence data: A performance assessment of methods for correcting sampling bias. *PLoS ONE*, 9(5), pp.1–13.

Franklin, J., 2009. *Mapping Species Distributions: Spatial Inference and Prediction*,

Fratkin, E., 2001. East African pastoralism in transition: Maasai, Boran, and Rendille cases. *African Studies Review*, 44(3), pp.1–25.

Freeman, E. a. & Moisen, G.G., 2008. A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. *Ecological Modelling*, 217, pp.48–58.

Friedl, M.A. et al., 2010. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, 114, pp.168–182.

Fritz, S. et al., 2011. Highlighting continued uncertainty in global land cover maps for the user community. *Environmental Research Letters*, 6(4), p.44005.

Fritz, S., See, L. & Rembold, F., 2010. Comparison of global and regional land cover maps with statistical information for the agricultural domain in Africa. *International Journal of Remote Sensing*, 31(9), pp.2237–2256.

Fuller, T.K. et al., 1992. Long distance dispersal by African wild dogs in East and South Africa. *African Zoology*, 106, pp.535–537.

Fuller, T.K. & Sievert, P.R., 2001. Carnivore demography and the consequences of changes in prey availability. In *Carnivore Conservation*. Cambridge, UK: Cambridge University Press.

Fund, G.C., 2013. *Africa's giraffe: a conservation guide*, Surrey, UK.

Fynn, R.W.S. et al., 2016. Strategic management of livestock to improve biodiversity conservation in African savannahs: A conceptual basis for wildlife-livestock coexistence. *Journal of Applied Ecology*, 53(2), pp.388–397.

García-Rangel, S. & Pettorelli, N., 2013. Thinking spatially: The importance of geospatial techniques for carnivore conservation. *Ecological Informatics*, 14, pp.84–89.

Gasaway, W.C. et al., 1992. The role of predation in limiting moose at low density in Alaska and Yukon and implications for conservation. *Wildlife Monographs*, 120, pp.3–59.

Gauch Jr., H.G., 2003. *Scientific Method in Practice*, Cambridge, UK: Cambridge University Press.

GDAL, 2013. GDAL - Geospatial Data Abstraction Library: Version 1.10.1.

Gehrt, S.D., Anchor, C. & White, L. a., 2009. Home Range and Landscape Use of Coyotes in a Metropolitan Landscape: Conflict or Coexistence? *Journal of Mammalogy*, 90(5), pp.1045–1057.

Gelorini, V. & Verschuren, D., 2012. Historical climate-human-ecosystem interaction in East Africa: a review. *African Journal of Ecology*, 51, pp.409–421.

Gilbert, S.L. et al., 2016. Socioeconomic Benefits of Large Carnivore Recolonization Through Reduced Wildlife-Vehicle Collisions. *Conservation Letters*, 0(July), pp.1–9.

Gittleman, J.L. et al. eds., 2001. *Carnivore Conservation*, Cambridge, UK: Cambridge University Press.

Goldman, M.J., 2011. Strangers in their own land: Maasai and wildlife conservation in Northern Tanzania. *Conservation and Society*, 9(1), pp.65–79.

Goldman, M.J., de Pinho, J.R. & Perry, J., 2013. Beyond ritual and economics: Maasai lion hunting and conservation politics. *Oryx*, pp.1–11.

Gong, P. et al., 2013. Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data. *International Journal of Remote Sensing*, 34(7), pp.2607–2654.

Google, 2014. Google Earth Features.

Gorman, M.L. et al., 1998. High hunting costs make African wild dogs vulnerable to kleptoparasitism by hyaenas. *Nature*, 391, pp.479–481.

Graham, C.H. et al., 2008. The influence of spatial errors in species occurrence data used in distribution models. *Journal of Applied Ecology*, 45(2007), pp.239–247.

Green, J.M.H. et al., 2013. Deforestation in an African biodiversity hotspot: Extent, variation and the effectiveness of protected areas. *Biological Conservation*, 164, pp.62–72.

Gros, P.M., Kelly, M.J. & Caro, T.M., 1996. Estimating carnivore densities for conservation purposes: indirect methods compared to baseline demographic data. *Oikos*, 77(2), pp.197–206.

Group, W.B., 2016. *Analysis of International Funding to Tackle Illegal Wildlife Trade*, Washington D.C.

Grünewald, C., Schleuning, M. & Böhning-Gaese, K., 2016. Biodiversity, scenery and infrastructure: Factors driving wildlife tourism in an African savannah national park. *Biological Conservation*, 201, pp.60–68.

Gubbi, S., 2012. Patterns and correlates of human–elephant conflict around a south Indian reserve. *Biological Conservation*, 148(1), pp.88–95.

Guillera-Arroita, G. et al., 2015. Is my species distribution model fit for purpose? Matching data and models to applications. *Global Ecology and Biogeography*, 24(3), pp.276–292.

Guisan, A. et al., 2006. Using niche-based models to improve the sampling of rare species. *Conservation Biology*, 20(2), pp.501–511.

Guisan, A. & Thuiller, W., 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, 8(9), pp.993–1009.

Gusset, M. et al., 2009. Human–wildlife conflict in northern Botswana: livestock predation by Endangered African wild dog *Lycaon pictus* and other carnivores. *Oryx*, 43(1), p.67.

Haddad, N.M. et al., 2015. Habitat fragmentation and its lasting impact on Earth's ecosystems. *Science Advances*, 1, pp.1–9.

Haken, J., 2011. *Transnational Crime in the Developing World*.

Hannerz, F. & Lotsch,  a., 2008. Assessment of remotely sensed and statistical inventories of African agricultural fields. *International Journal of Remote Sensing*, 29(13), pp.3787–3804.

Hansen, M.C. et al., 2000. Global land cover classification at 1 km spatial resolution using a classification tree approach. *International Journal of Remote Sensing*, 21(6–7), pp.1331–1364.

Hanski, I., 1998. Metapopulation dynamics. *Nature*, 396(6706), pp.41–49.

Hanski, I. & Ovaskainen, O., 2000. The metapopulation capacity of a fragmented landscape. *Nature*, 404(6779), pp.755–8.

Harcourt, A.H., Parks, S.A. & Woodroffe, R., 2001. Human density as an influence on species/area relationships: double jeopardy for small African reserves? *Biodiversity and Conservation*, 10, pp.1011–1026.

Hayward, M., Obrien, J. & Kerley, G., 2007. Carrying capacity of large African predators: Predictions and tests. *Biological Conservation*, 139(1–2), pp.219–229.

Hayward, M.W. et al., 2006c. Prey preferences of the African wild dog *Lycaon pictus*: ecological requirements for conservation. *Journal of Mammalogy*, 87(6), pp.1122–1131.

Hayward, M.W., Hofmeyr, M., et al., 2006b. Prey preferences of the cheetah (*Acinonyx jubatus*) (Felidae: Carnivora): morphological limitations or the need to capture rapidly consumable prey before kleptoparasites arrive? *Journal of Zoology*, 270(4), pp.615–627.

Hayward, M.W., Henschel, P., et al., 2006a. Prey preferences of the leopard (*Panthera pardus*). *Journal of Zoology*, 270, pp.298–313.

Hayward, M.W. & Kerley, G.I.H., 2008. Prey preferences and dietary overlap amongst Africa's large predators. *South African Journal of Wildlife Research*, 38(2), pp.93–108.

Hayward, M.W. & Kerley, G.I.H., 2005. Prey preferences of the lion (*Panthera leo*). *Journal of Zoology*, 267(3), p.309.

Hazelton, P. & Murphy, B., 2007. *Soil test results: What do all the numbers mean?* 2nd ed., Collingwood: CSIRO Publishing.

Hazzah, L. et al., 2017. From Attitudes to Actions: Predictors of Lion Killing by Maasai Warriors. *Plos One*, 12(1), p.e0170796.

Hazzah, L. et al., 2013. The influence of park access during drought on attitudes toward wildlife and lion killing behaviour in Maasailand, Kenya. *Environmental Conservation*, 40(3), pp.266–276.

Hebblewhite, M. et al., 2011. Predicting potential habitat and population size for reintroduction of the Far Eastern leopards in the Russian Far East. *Biological Conservation*, 144(10), pp.2403–2413.

Hengl, T. et al., 2014. SoilGrids1km--global soil information based on automated mapping. *PloS one*, 9(8), p.e105992.

Hengl, T. & Reuter, H.I., 2009. Soil Mapping Applications. In *Geomorphometry: Concepts, Software, Applications*. Elsevier, pp. 461–480.

Henschel, P., 2008. *The conservation biology of the leopard Panthera pardus in Gabon:*

*Status, threats and strategies for conservation*. University of Gottingen.

Henschel, P. et al., 2014. The Lion in West Africa Is Critically Endangered M. Hayward, ed. *PLoS ONE*, 9(1), p.e83500.

Hertzog, L.R., Besnard, A. & Jay-Robert, P., 2014. Field validation shows bias-corrected pseudo-absence selection is the best method for predictive species-distribution modelling. *Diversity and Distributions*, 20(12), pp.1403–1413.

Hijmans, R.J., 2012. Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model. *Ecology*, 93(3), pp.679–688.

Hijmans, R.J. et al., 2016. package dismo.

Hijmans, R.J., 2016. package Raster.

Hijmans, R.J. et al., 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15), pp.1965–1978.

Hilborn, A. et al., 2012. Stalk and chase: how hunt stages affect hunting success in Serengeti cheetah. *Animal Behaviour*, 84(3), pp.701–706.

Hirzel, A.H. et al., 2006. Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling*, 199(2), pp.142–152.

Hirzel, A.H., Helfer, V. & Metral, F., 2001. Assessing habitat-suitability models with a virtual species. *Ecological Modelling*, 145, pp.111–121.

Hirzel, A.H. & Le Lay, G., 2008. Habitat suitability modelling and niche theory. *Journal of Applied Ecology*, 45(5), pp.1372–1381.

Holt, R.D., 1977. Predation, apparent competition, and the structure of prey communities. *Theoretical Population Biology*, 12(2), pp.197–229.

Holt, R.D. & Polis, G.A., 1997. A Theoretical Framework for Intraguild Predation and competition. *The American Naturalist*, 149(4), pp.745–764.

Homewood, K.M. & Rodgers, W.A., 1991. *Maasailand Ecology: Pastoralist Development and Conservation in Ngorongoro, Tanzania*, Cambridge, Uk: Cambridge University Press.

Homewood, K.M., Trench, P. & Brockington, D., 2012. Pastoralist livelihoods and wildlife revenues in East Africa: a case for coexistence? *Pastoralism: Research, Policy and Practice*, 2(1), p.19.

Hopcraft, J.G.C., Sinclair, a. R.E. & Packer, C., 2005. Planning for success: Serengeti lions seek prey accessibility rather than abundance. *Journal of Animal Ecology*, 74(3), pp.559–566.

Hubel, T.Y. et al., 2016b. Additive opportunistic capture explains group hunting benefits in African wild dogs. *Nature Communications*, 7, pp.1–11.

Hubel, T.Y. et al., 2016a. Energy cost and return for hunting in African wild dogs and cheetahs. *Nature Communications*, 7, p.11034.

Hunter, J.S., Durant, S.M. & Caro, T.M., 2007. Patterns of scavenger arrival at cheetah kills in Serengeti National Park Tanzania. *African Journal of Ecology*, 45(3), pp.275–281.

Hunter, L. et al., 2012. Urgent and comprehensive reform of trophy hunting of lions is a better option than an endangered listing: A science-based approach.

Huygens, O.C. & Hayashi, H., 1999. Using electric fences to reduce Asiatic black bear depredation in Nagano prefecture, central Japan. *Wildlife Society Bulletin*, 27(4), pp.959–964.

Ikanda, D. & Packer, C., 2008. Ritual vs. retaliatory killing of African lions in the Ngorongoro Conservation Area, Tanzania. *Endangered Species Research*, 6(December), pp.67–74.

Inskip, C. & Zimmermann, A., 2009. Human-felid conflict: a review of patterns and priorities worldwide. *Oryx*, 43(1), p.18.

IUCN/SSC, 2007a. *Regional conservation strategy for the cheetah and African wild dog in Eastern Africa*, Gland, Switzerland.

IUCN/SSC, 2007b. *Regional conservation strategy for the cheetah and African wild dog in Southern Africa*, Gland, Switzerland.

IUCN/SSC, 2012. *Regional Conservation Strategy for the Cheetah and African Wild Dog in Western, Central and Northern Africa*, Gland, Switzerland.

Jackson, C.R. et al., 2016. Evaluating habitat connectivity methodologies: a case study with endangered African wild dogs in South Africa. *Landscape Ecology*.

Jackson, C.R. et al., 2014. Heading for the hills: risk avoidance drives den site selection in african wild dogs. *PloS one*, 9(6), p.e99686.

Jackson, R.M. et al., 1996. Reducing livestock depredation losses in the Nepalese Himalaya. In *Proceedings of the 17th Vertebrate Pest Conference*. pp. 241–247.

Jacobson, A. et al., 2015. A novel approach to mapping land conversion using Google Earth with an application to East Africa. *Environmental Modelling & Software*, 72, pp.1–9.

Jacobson, A.P. et al., 2013. Distribution and abundance of lions in northwest Tete Province, Mozambique. *Tropical Consevation Science*, 6(1), pp.87–107.

Jacobson, A.P. et al., 2016. Leopard (*Panthera pardus*) status, distribution, and the research efforts across its range. *PeerJ*, 4, p.e1974.

Janecka, J.E. et al., 2016. Genetic differences in the response to landscape fragmentation by a habitat generalist, the bobcat, and a habitat specialist, the ocelot. *Conservation Genetics*.

Jarnevich, C.S. et al., 2015. Caveats for correlative species distribution modeling. *Ecological Informatics*, 29, pp.6–15.

Jarvis, A. et al., 2008. Hole-filled SRTM for the globe Version 4.

Jenkins, C.N., Pimm, S.L. & Joppa, L.N., 2013. Global patterns of terrestrial vertebrate diversity and conservation. *Proceedings of the National Academy of Sciences of the United States of America*, 110(28), pp.E2602-10.

Jensen, M.E., Simonson, G.H. & Dosskey, M., 1990. Correlation between soils and sagebrush-dominated plant communities of Northeastern Nevada. *Soil Science Society of America*, 54(3), pp.902–910.

Jhala, Y. V. & Sharma, D.K., 1997. Childlifting by wolves in eastern uttar Pradesh. *Journal of Wildlife Research*, 2, pp.94–101.

Jiménez-Valverde, A., Lobo, J.M. & Hortal, J., 2008. Not as good as they seem: The importance of concepts in species distribution modelling. *Diversity and Distributions*, 14(6), pp.885–890.

Johnson, D.H., 1980. The Comparison of Usage and Availability Measurements for Evaluating Resource Preference. *Ecology*, 61(1), pp.65–71.

Jongeling, T.B. & Koetsier, T., 2014. The predicament of the African wild dog, *Lycaon pictus*, is less precarious than claimed. *African Journal of Ecology*, 52(4), pp.466–470.

Kanagaraj, R. et al., 2013. Modelling species distributions to map the road towards carnivore conservation in the tropics. *Raffles Bulletin of Zoology*, 28, pp.85–107.

Kanyamibwa, S., 1998. Impact of war on conservation: Rwandan environment and wildlife in agony. *Biodiversity and Conservation*, 7(11), pp.1399–1406.

Karanth, K.U. et al., 2004. Tigers and their prey: Predicting carnivore densities from prey abundance. *Proceedings of the National Academy of Sciences of the United States of America*, 101(14), pp.4854–4858.

Kaswamila, A., 2009. Human-wildlife conflicts in Monduli District, Tanzania. *International Journal of Biodiversity Science, Ecosystem Services & Management*, 5(4), pp.199–207.

Kearney, M. & Porter, W., 2009. Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology letters*, 12(4), pp.334–50.

Keinath, D.A. et al., 2016. A global analysis of traits predicting species sensitivity to habitat fragmentation. *Global Ecology and Biogeography*, pp.1–13.

Kellert, S.R. et al., 1996. Human Culture and Large Carnivore Conservation in North America. *Conservation Biology*, 10(4), pp.977–990.

Kelly, M.J. et al., 1998. Demography of the Serengeti cheetah (*Acinonyx jubatus*) population: the first 25 years. *Journal of Zoology*, 244(4), pp.473–488.

Kerr, J.T. & Ostrovsky, M., 2003. From space to species: ecological applications for remote sensing. *Trends in Ecology & Evolution*, 18(6), pp.299–305.

Khorozyan, I. et al., 2015. Big cats kill more livestock when wild prey reaches a minimum threshold. *Biological Conservation*, 192, pp.268–275.

Kiffner, C. et al., 2014. From savannah to farmland: effects of land-use on mammal communities in the Tarangire-Manyara ecosystem, Tanzania. *African Journal of Ecology*.

Kingdon, J., 1977. *East African Mammals: Volume IIIA, Carnivores* Edition 19., Chicago: University of Chicago Press.

Kissui, B.M., 2008. Livestock predation by lions, leopards, spotted hyenas, and their vulnerability to retaliatory killing in the Maasai steppe, Tanzania. *Animal Conservation*, 11(5), pp.422–432.

Kock, R.A., 1995. Wildlife utilization: use it or lose it - a Kenyan perspective. *Biodiversity and Conservation*, 4(3), pp.241–256.

Kolowski, J.M. & Holekamp, K.E., 2006. Spatial, temporal, and physical characteristics of livestock depredations by large carnivores along a Kenyan reserve border. *Biological Conservation*, 128(4), pp.529–541.

Kramer-Schadt, S. et al., 2013. The importance of correcting for sampling bias in MaxEnt species distribution models. *Diversity and Distributions*, 19(11), pp.1366–1379.

Kruuk, H., 2002. *Hunter and Hunted: Relationships between carnivores and people*, Cambridge, UK: Cambridge University Press.

Kruuk, H., 1972. Surplus killing by carnivores. *Journal of Zoology*, 166, pp.233–244.

Kuijper, D.P.J. et al., 2016. Paws without claws? Ecological effects of large carnivores in anthropogenic landscapes. *Proceedings of the Royal Society B: Biological Sciences*, 283(1841), p.20161625.

Kushnir, H. et al., 2010. Human and Ecological Risk Factors for Unprovoked Lion Attacks on Humans in Southeastern Tanzania. *Human Dimensions of Wildlife*, 15(5), pp.315–331.

Kushnir, H. et al., 2014. Using landscape characteristics to predict risk of lion attacks on humans in south-eastern Tanzania. *African Journal of Ecology*, pp.1–9.

Landa, A. et al., 1997. Wolverines and their prey in southern Norway. *Canadian Journal of Zoology*, 75, pp.1292–1299.

Laporte, I. et al., 2010. Effects of wolves on elk and cattle behaviors: Implications for livestock production and wolf conservation. *PLoS ONE*, 5(8).

Latham, J. et al., 2014. FAO Global Land Cover (GLC-SHARE) Beta-Release 1.0 Database.

Laundré, J.W., Hernández, L. & Altendorf, K.B., 2001. Wolves, elk, and bison: reestablishing the "landscape of fear" in Yellowstone National Park, U.S.A. *Canadian Journal of Zoology*, 79(8), pp.1401–1409.

Laundré, J.W., Hernández, L. & Ripple, W.J., 2010. The Landscape of Fear: Ecological Implications of Being Afraid. *The Open Ecology Journal*, 3, pp.1–7.

Laurance, W.F. et al., 2015. Estimating the Environmental Costs of Africa's Massive "development Corridors." *Current Biology*, 25(24), pp.3202–3208.

Laurenson, M.K., 1994. High juvenile mortality in cheetahs (*Acinonyx jubatus*) and its consequences for maternal care. *Journal of Zoology*, 234, pp.387–408.

Laurenson, M.K., 1995. Implications of high offspring mortality for cheetah population dynamics. In A. R. E. Sinclair & P. Arcese, eds. *Serengeti II: Dynamics, Management, and Conservaiton of an Ecosystem*. Chicago: University of Chicago Press, pp. 385–399.

Lehman, N. et al., 1991. Introgression of Coyote Mitochondrial DNA Into Sympatric North American Gray Wolf Populations. *Evolution*, 45(1), pp.104–119.

Lehner, B. & Döll, P., 2004. Development and validation of a global database of lakes, reservoirs and wetlands. *Journal of Hydrology*, 296, pp.1–22.

Lehner, B., Verdin, K. & Jarvis, A., 2006. *HydroSHEDS technical documentation*, Washington D.C.

Leigh, K. A. et al., 2012. Loss of genetic diversity in an outbreeding species: small population effects in the African wild dog (*Lycaon pictus*). *Conservation Genetics*, 13(3), pp.767–777.

Lescureux, N. & Linnell, J.D.C., 2014. Warring brothers: The complex interactions between wolves (*Canis lupus*) and dogs (*Canis familiaris*) in a conservation context. *Biological Conservation*, 171, pp.232–245.

Lewis, D. & Jackson, J., 2005. Safari hunting and conservation on communal land in southern Africa. In R. Woodroffe, S. Thirgood, & A. Rabinowitz, eds. *People and Wildlife: Conflict or Coexistence*. Cambridge University Press, pp. 239–251.

Lewis, J.S. et al., 2015. The effects of urbanization on population density, occupancy, and detection probability of wild felids. *Ecological Applications*, 25, pp.1880–1895.

Li, J. et al., 2013. Human-snow leopard conflicts in the Sanjiangyuan Region of the Tibetan Plateau. *Biological Conservation*, 166, pp.118–123.

Liaw, A. & Wiener, M., 2002. Classification and Regression by randomForest. , 2(December), pp.18–22.

Lichtenfeld, L., 2005. *Our Shared Kingdom At Risk : Human – Lion Relationships in the 21st Century*.

Lichtenfeld, L.L., Trout, C. & Kisimir, E.L., 2014. Evidence-based conservation: predator-proof bomas protect livestock and lions. *Biodiversity and Conservation*.

Linard, C. et al., 2012. Population distribution, settlement patterns and accessibility across Africa in 2010. *PloS one*, 7(2), p.e31743.

Lindenmayer, D.B. & Fischer, J., 2013. *Habitat Fragmentation and Landscape Change: An Ecological and Conservation Synthesis*, Island Press.

Lindsey, P.A. et al., 2016. Life after Cecil: Channelling global outrage into funding for conservation in Africa. *Conservation Letters*, 0(0), pp.1–6.

Lindsey, P.A., Balme, G., et al., 2013b. The bushmeat trade in African savannas: Impacts, drivers, and possible solutions. *Biological Conservation*, 160, pp.80–96.

Lindsey, P.A. et al., 2012. The significance of African lions for the financial viability of trophy hunting and the maintenance of wild land. *PloS one*, 7(1), p.e29332.

Lindsey, P.A., Balme, G.A., et al., 2013a. The Trophy Hunting of African Lions: Scale, Current Management Practices and Factors Undermining Sustainability M. Hayward, ed. *PLoS ONE*, 8(9), p.e73808.

Lindsey, P.A. et al., 2007. Wildlife Viewing Preferences of Visitors to Protected Areas in South Africa: Implications for the Role of Ecotourism in Conservation. *Journal of Ecotourism*, 6(1), pp.19–33.

Lindsey, P. A. et al., 2014. Underperformance of african protected area networks and the case for new conservation models: insights from zambia. *PloS one*, 9(5), p.e94109.

Lindsey, P. A. et al., 2011. Ecological and financial impacts of illegal bushmeat trade in Zimbabwe. *Oryx*, 45(1), pp.96–111.

Lindsey, P. a. et al., 2006. Potential of trophy hunting to create incentives for wildlife conservation in Africa where alternative wildlife-based land uses may not be viable. *Animal Conservation*, 9(3), pp.283–291.

Lindsey, P., Roulet, P. & Romanach, S., 2007. Economic and conservation significance of the trophy hunting industry in sub-Saharan Africa. *Biological Conservation*, 134(4), pp.455–469.

Lindsey, P. et al., 2017. The performance of African protected areas for lions and their prey. *Biological Conservation*, 209, pp.137-149.

Linkie, M. et al., 2006. Assessing the viability of tiger subpopulations in a fragmented landscape. *Journal of Applied Ecology*, 43(3), pp.576–586.

Linnell, J.D.C. et al., 1999. Camivores and that kill livestock: do “problem individuals” really exist? *Wildlife Society Bulletin*, 27(3), pp.698–705.

Linnell, J.D.C. et al., 2002. *The fear of wolves: a review of wolf attacks on humans*,

Linnell, J.D.C. et al., 2005. Zoning as a means of mitigating conflicts with large carnivores: principles and reality. In R. Woodroffe, S. Thirgood, & A. Rabinowitz, eds. *People and Wildlife: Conflict or Coexistence*. Cambridge, UK: Cambridge University Press, pp. 162–175.

Linnell, J.D.C. & Strand, O., 2000. Interference interactions, co-existence and conservation of mammalian carnivores. *Diversity and Distributions*, 6(4), pp.169–176.

Linnell, J.D.C., Swenson, J.E. & Andersen, R., 2001. Predators and people: conservation of large carnivores is possible at high human densities if management policy is favourable. *Animal Conservation*, 4, pp.345–349.

Loarie, S.R., Tambling, C.J. & Asner, G.P., 2013. Lion hunting behaviour and vegetation structure in an African savanna. *Animal Behaviour*, 85(5), pp.899–906.

Lobo, J.M., Jiménez-Valverde, A. & Real, R., 2008. AUC: A misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, 17, pp.145–151.

Lobo, J.M. & Tognelli, M.F., 2011. Exploring the effects of quantity and location of pseudo-absences and sampling biases on the performance of distribution models with limited point occurrence data. *Journal for Nature Conservation*, 19(1), pp.1–7.

Löe, J., 2002. *Large Carnivore Related Deaths: A Conservation Issue*. Norwegian University of Science and Technology.

Löe, J. & Röskaft, E., 2004. Large carnivores and human safety: a review. *Ambio*, 33(6), pp.283–8.

Lopez-Bao, J.V. et al., 2015. Carnivore coexistence: Wilderness not required. *Science*, 348(6237), pp.871–872.

Loucks, C. et al., 2010. Sea level rise and tigers: predicted impacts to Bangladesh's Sundarbans mangroves. *Climatic Change*, 98(1–2), pp.291–298.

Loveridge, A.J. & Canney, S., 2009. *African Lion Distribution Modelling Project*, Horsham, UK.

Lyamuya, R.D. et al., 2014. Human–carnivore conflict over livestock in the eastern part of the Serengeti ecosystem, with a particular focus on the African wild dog *Lycaon pictus*. *Oryx*, pp.1–7.

Macdonald, D.W., 1980. *Rabies and Wildlife*, Oxford, UK: Oxford University Press.

Macdonald, D.W. & Loveridge, A.J. eds., 2010. *Biology and conservation of wild felids*, Oxford, UK: Oxford University Press.

Macdonald, D.W. & Sillero-Zubiri, C. eds., 2004. *Biology and Conservation of Wild Canids*, Oxford: Oxford University Press.

MacKenzie, D.I. et al., 2006. *Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence*, Oxford, UK: Elsevier Inc.

Madden, F., 2004. Creating Coexistence between Humans and Wildlife: Global Perspectives on Local Efforts to Address Human–Wildlife Conflict. *Human Dimensions of Wildlife*, 9(4), pp.247–257.

Maddox, T.M., 2003. *The ecology of cheetahs and other large carnivores in a pastoralist-dominated buffer zone*. University College London.

Manly, B.F.J. et al., 2002. *Resource Selection by Animals: Statistical Design and Analysis for*

*Field Studies* 2nd ed., Dordrecht: Kluwer Academic.

Di Marco, M. et al., 2014. Drivers of extinction risk in African mammals: the interplay of distribution state, human pressure, conservation response and species biology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369.

Marcstrom, V., Kenward, R. & Engren, E., 1988. The impact of predation on boreal tetraonids during vole cycles. *Journal of Animal Ecology*, 57, pp.859–872.

Marker, L.L., 2002. *Aspects of cheetah biology, ecology and conservation strategies on Namibian Farmlands*. University of Oxford.

Márquez, C. et al., 2013. Risk mapping of illegal poisoning of avian and mammalian predators. *The Journal of Wildlife Management*, 77(1), pp.75–83.

Mascia, M.B. & Pailler, S., 2011. Protected area downgrading, downsizing, and degazettement (PADDD) and its conservation implications. *Conservation Letters*, 4, pp.9–20.

Masenga, E.H. et al., 2016. Insights into long-distance dispersal by African wild dogs in East Africa. *African Journal of Ecology*, 54(1), pp.95–98.

Mashintonio, A.F. et al., 2014. Data-driven discovery of the spatial scales of habitat choice by elephants. *PeerJ*, 2, p.e504.

Mateo-Tomás, P. et al., 2012. Alleviating human-wildlife conflicts: identifying the causes and mapping the risk of illegal poisoning of wild fauna. *Journal of Applied Ecology*, 49(2), pp.376–385.

Matiko, D., 2014. Wildlife Conservation Leases are Considerable Conservation Options outside Protected Areas: The Kitengela - Nairobi National Park Wildlife Conservation Lease Program. *Journal of Ecosystem & Ecography*, 4(2), pp.2–9.

McCallum, I. et al., 2006. A spatial comparison of four satellite derived 1km global land cover datasets. *International Journal of Applied Earth Observation and Geoinformation*, 8(4), pp.246–255.

McNutt, J.W., 1996. Sex-biased dispersal in African wild dogs, *Lycaon pictus*. *Animal Behaviour*, 52(1996), pp.1067–1077.

McPherson, J.M., Jetz, W. & Rogers, D.J., 2004. The effects of species' range sizes on the accuracy of distribution models: Ecological phenomenon or statistical artefact? *Journal of Applied Ecology*, 41(5), pp.811–823.

Mech, L.D. et al., 2000. Assessing Minnesota factors farms to predispose depredations wolf cattle. *Wildlife Society Bulletin*, 28(3), pp.623–629.

Mech, L.D., 1995. The Challenge and Opportunity of Recovering Wolf Populations. *Conservation Biology*, 9(2), pp.270–278.

van der Meer, E. et al., 2013. Ecological trap in the buffer zone of a protected area: effects of indirect anthropogenic mortality on the African wild dog Lycaon pictus. *Oryx*, (October 2012), pp.1–9.

Merow, C. et al., 2014. What do we gain from simplicity versus complexity in species distribution models? *Ecography*, 37, pp.1267–1281.

Merow, C., Smith, M.J. & Silander, J. a., 2013. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography*, 36(10), pp.1058–1069.

Miller, J.R.B., Jhala, Y. V., Jena, J., et al., 2015. Landscape-scale accessibility of livestock to tigers: Implications of spatial grain for modeling predation risk to mitigate human-carnivore conflict. *Ecology and Evolution*, 5(6), pp.1354–1367.

Miller, J.R.B., 2015. Mapping attack hotspots to mitigate human-carnivore conflict: approaches and applications of spatial predation risk modeling. *BioScience*.

Miller, J.R.B., Jhala, Y. V. & Jena, J., 2015. Livestock losses and hotspots of attack from tigers and leopards in Kanha Tiger Reserve, Central India. *Regional Environmental Change*.

Mills, G. & Hofer, H., 1998. *Hyaenas: Status Survey and Conservation Action Plan*, IUCN, Gland, Switzerland and Cambridge, UK: IUCN/SSC Hyaena Specialist Group.

Mills, G. & Mills, M., 2014. Cheetah cub survival revisited: a re-evaluation of the role of predation, especially by lions, and implications for conservation. *Journal of Zoology*, 2.

MNRERF & WWF, 2014. *Strategy for Conservation of the Far-Eastern Leopard in the Russian Federation*, Moscow.

Mponzi, B.P., Lepczyk, C.A. & Kissui, B.M., 2014. Characteristics and distribution of live-stock losses caused by wild carnivores in Maasai Steppe of northern Tanzania. *Human-Wildlife Interactions*, 8(2), pp.218–227.

Msoffe, F.U. et al., 2011. Spatial correlates of land-use changes in the Maasai- Steppe of Tanzania : Implications for conservation and environmental planning. , 3(July), pp.280–290.

Msuha, M.J. et al., 2012. Conserving biodiversity in a changing world: land use change and species richness in northern Tanzania. *Biodiversity and Conservation*, 21(11), pp.2747–2759.

Muhly, T.B. et al., 2010. Differential risk effects of wolves on wild versus domestic prey have consequences for conservation. *Oikos*, 119(8), pp.1243–1254.

Mwakatobe, A., Nyahongo, J. & Røskaft, E., 2013. Livestock Depredation by Carnivores in the Serengeti Ecosystem, Tanzania. *Environment and Natural Resources Research*, 3(4).

Myers, N. et al., 2000. Biodiversity hotspots for conservation priorities. *Nature*, 403(6772), pp.853–858.

Myers, N., 1975. The cheetah in Africa - Report of a survey in Africa from the Sahara southwards.

Myers, N., 1976. *The Leopard Panthera pardus in Africa*, Morges, Switzerland.

NASA LP DAAC, 2012. MODIS MCD12.

National Geomatics Center of China, 2014. 30 m Global Land Cover Data Product (GlobeLand30).

Naughton-Treves, L., 1997. Farming the forest edge: vulnerable places and people around Kibale National Park, Uganda. *The Geographical Review*, 87, pp.27–46.

Naughton-Treves, L., Holland, M.B. & Brandon, K., 2005. The Role of Protected Areas in Conserving Biodiversity and Sustaining Local Livelihoods. *Annual Review of Environment and Resources*, 30(1), pp.219–252.

Naves, J. et al., 2003. Endangered Species Constrained by Natural and Human Factors: the Case of Brown Bears in Northern Spain. *Conservation Biology*, 17(5), pp.1276–1289.

Nelson, F., 2007. Emergent or illusory? Community wildlife management in Tanzania. , (146), pp.1–32.

Nelson, F., 2009. *Natural Conservationists? Evaluating the Impact of Pastoralist Land Use Practices on Tanzania's Wildlife Economy*, Arusha, Tanzania.

Nelson, Nshala, R. & Rodgers, W., 2007. The Evolution and Reform of Tanzanian Wildlife Management. *Conservation and Society*, 5(2), p.232.

Neumann, R.P., 2000. Land, Justice, and the Politics of Conservation in Tanzania. In C. Zerner, ed. *People, Plants, and Justice: The Politics of Nature Conservation*. New York, USA: Columbia University Press, pp. 117–133.

Newmark, W.D., 1996. Insularization of Tanzanian Parks and the Local Extinction of Large Mammals. *Conservation Biology*, 10(6).

Newsome, T.M. et al., 2014. The ecological effects of providing resource subsidies to predators. *Global Ecology and Biogeography*.

Northrup, J.M. et al., 2013. Practical guidance on characterizing availability in resource selection functions under a use-availabilty design. *Ecology*, 94(7), pp.1456–1463.

Norton-Griffiths, M. & Southey, C., 1995. The opportunity costs of biodiversity conservation: a case study of Kenya. *Ecological Economics*, 12, pp.125–139.

Nowell, K. & Jackson, P. eds., 1996. *Wild cats: status survey and conservation action plan*, Gland, Switzerland: IUCN/SSC Cat Specialist Group.

Odden, J. et al., 2008. Vulnerability of domestic sheep to lynx depredation in relation to roe deer density. *Journal of Wildlife Management*, 72(1), pp.276–282.

Oehler, J.D. & Litvaitis, J.A., 1996. The role of spatial scale in understanding responses of medium-sized carnivores to forest fragmentation. *Canadian Journal of Zoology*, 74, pp.2070–2079.

Ogada, M.O. et al., 2003. Limiting Depredation by African Carnivores: the Role of Livestock Husbandry. *Conservation Biology*, 17(6), pp.1521–1530.

Ogutu, J.O. et al., 2016. Extreme Wildlife Declines and Concurrent Increase in Livestock Numbers in Kenya: What Are the Causes? *Plos One*, 11(9), p.e0163249.

Okello, M.M., 2005. Land Use Changes and Human–Wildlife Conflicts in the Amboseli Area, Kenya. *Human Dimensions of Wildlife*, 10(1), pp.19–28.

Okello, M.M. & Kiringe, J.W., 2004. Threats to Biodiversity and their Implications in Protected and Adjacent Dispersal Areas of Kenya. *Journal of Sustainable Tourism*, 12(1), pp.55–69.

Okello, M.M., Kiringe, J.W. & Warinwa, F., 2014. Human-Carnivore Conflicts in Private Conservancy Lands of Elerai and Oltiyiani in Amboseli Area, Kenya. *Natural Resources*, 5(8), pp.375–391.

Okello, M.M., Manka, S.G. & D'Amour, D.E., 2008. The relative importance of large mammal species for tourism in Amboseli National Park, Kenya. *Tourism Management*, 29(4), pp.751–760.

Oli, M.K., Taylor, I.R. & Rogers, M.E., 1994. Snow leopard (*Panthera uncia*) predation of livestock: An assessment of local perceptions in the Annapurna Conservation Area, Nepal. *Biological Conservation*, 68(1), pp.63–68.

Olson, D.M. et al., 2001. Terrestrial Ecoregions of the World: A New Map of Life on Earth. *BioScience*, 51(11), p.933.

Oriol-Cotterill, A. et al., 2015b. Landscapes of Coexistence for terrestrial carnivores: The ecological consequences of being downgraded from ultimate to penultimate predator by humans. *Oikos*, 124(10), pp.1263–1273.

Oriol-Cotterill, A. et al., 2015a. Spatiotemporal patterns of lion space use in a human-dominated landscape. *Animal Behaviour*, 101, pp.27–39.

van Orsdol, K.G., Hanby, J.P. & Bygott, J.D., 1985. Ecological correlates of lion social organization. , pp.97–112.

Packer, C. et al., 2013. Conserving large carnivores: dollars and fence. *Ecology letters*, 16(5), pp.635–41.

Packer, C. et al., 2010. Effects of trophy hunting on lion and leopard populations in Tanzania. *Conservation Biology*, 25, pp.142–53.

Packer, C. et al., 2009. Sport hunting, predator control and conservation of large carnivores. *PloS one*, 4(6), p.e5941.

Packer, C. et al., 1999. Viruses of the Serengeti: patterns of infection and mortality in

African lions. *Journal of Animal Ecology*, 68(6), pp.1161–1178.

Palomares, F. & Caro, T.M., 1999. Interspecific Killing among Mammalian Carnivores. *The American Naturalist*, 153(5), pp.492–508.

Patterson, B.D. et al., 2004. Livestock predation by lions (*Panthera leo*) and other carnivores on ranches neighboring Tsavo National ParkS, Kenya. *Biological Conservation*, 119(4), pp.507–516.

Pearce, J. & Ferrier, S., 2000. Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling*, 133(3), pp.225–245.

Pearce, J.L. & Boyce, M.S., 2006. Modelling distribution and abundance with presence-only data. *Journal of Applied Ecology*, 43(3), pp.405–412.

Peterson, A.T., 2001. Predicting species' geographic distributions based on ecological niche modeling. *Condor*, 103, pp.599–605.

Peterson, A.T., Ball, L.G. & Cohoon, K.C., 2002. Predicting distributions of Mexican birds using ecological niche modelling methods. *Ibis*, 144, pp.e27–e32.

Pettorelli, N. et al., 2010. Carnivore biodiversity in Tanzania: revealing the distribution patterns of secretive mammals using camera traps. *Animal Conservation*, 13(2), pp.131–139.

Pettorelli, N. et al., 2009. Exploring habitat use by cheetahs using ecological niche factor analysis. *Journal of Zoology*, 277(2), pp.141–148.

Pettorelli, N. et al., 2011. Predation, individual variability and vertebrate population dynamics. *Oecologia*, 167(2), pp.305–314.

Pettorelli, N., 2013. *The Normalized Difference Vegetation Index*, Oxford, UK: Oxford University Press.

Pettorelli, N. et al., 2005. Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in ecology & evolution*, 20(9), pp.503–10.

Pettorelli, N., Safi, K. & Turner, W., 2014. Satellite remote sensing, biodiversity research and conservation of the future. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(20130190), pp.20130190–20130190.

Pfeifer, M., Platts, P.J., et al., 2012a. Land use change and carbon fluxes in East Africa quantified using earth observation data and field measurements. *Environmental Conservation*, 40(3), pp.241–252.

Pfeifer, M., Burgess, N.D., et al., 2012b. Protected Areas: Mixed Success in Conserving East Africa's Evergreen Forests. *Plos ONE*, 7(6), pp.1–10.

Pfeifer, M., Disney, M., et al., 2012c. Terrestrial ecosystems from space: a review of earth observation products for macroecology applications. *Global Ecology and Biogeography*, 21(6), pp.603–624.

Phillips, S.J. et al., 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications*, 19(1), pp.181–197.

Phillips, S.J., Anderson, R.P. & Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190(3–4), pp.231–259.

Phillips, S.J. & Dudik, M., 2008. Modeling of species distributions with Maxent : new extensions and a comprehensive evaluation (December), pp.161–175.

Phillips, S.J. & Elith, J., 2013. On estimating probability of presence from use – availability or presence – background data. *Ecology*, 94(6), pp.1409–1419.

Pimm, S.L. et al., 2014. The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, 344(6187), pp.1246752–1246752.

Pimm, S.L. & Raven, P., 2000. Biodiversity. Extinction by numbers. *Nature*, 403(6772), pp.843–845.

Piquero, A.R. & Weisburd, D. eds., 2010. *Handbook of Quantitative Criminology*, New York, USA: Springer-Verlag.

Poessel, S.A., Gese, E.M. & Young, J.K., 2017. Environmental factors influencing the occurrence of coyotes and conflicts in urban areas. *Landscape and Urban Planning*, 157, pp.259–269.

Polis, G.A., Myers, C.A. & Holt, R.D., 1989. The ecology and evolution of intraguild predation: potential competitors that eat each other. *Annual Review of Ecology, Evolution, and Systematics*, 20, pp.297–330.

Pomilia, M.A., Mcnutt, J.W. & Jordan, N.R., 2015. Ecological predictors of African wild dog ranging patterns in Northern Botswana. *Journal of Mammalogy*, 96(6), pp.1214–1223.

Potere, D., 2008. Horizontal Positional Accuracy of Google Earth's High-Resolution Imagery Archive. *Sensors*, 8, pp.7973–7981.

Potere, D. & Schneider, A., 2007. A critical look at representations of urban areas in global maps. *GeoJournal*, 69(1–2), pp.55–80.

Pratt, D.J. & Gwynne, M.D. eds., 1977. *Rangeland management and ecology in East Africa*, Huntington, N.Y.: Krieger Publishing Co.

Prins, H.H.T., 1987. Nature conservation as an integral part of optimal land use in East Africa: The case of the Masai ecosystem of Northern Tanzania. *Biological Conservation*, 40(2), pp.141–161.

Purvis, A. et al., 2000. Predicting extinction risk in declining species. *Proceedings. Biological sciences / The Royal Society*, 267(1456), pp.1947–52.

Quigley, H.B. & Herrero, S., 2005. Characterization and prevention of attacks on humans.

In *People and Wildlife: Conflict or Coexistence*. Cambridge, UK: Cambridge University Press.

R, 2016. R: A language and environment for statistical computing.

Rabinowitz, A. & Zeller, K. a., 2010. A range-wide model of landscape connectivity and conservation for the jaguar, Panthera onca. *Biological Conservation*, 143(4), pp.939–945.

Ratcliffe, J., 2010. Crime mapping: spatial and temporal challenges. In A. R. Piquero & D. Weisburd, eds. *Handbook of Quantitative Criminology*. pp. 5–24.

Ray, J.C. et al. eds., 2005. *Large Carnivores and the Conservation of Biodiversity*, Washington D.C.: Island Press.

Ray, J.C., Hunter, L. & Zigouris, J., 2005. *Setting conservation and research priorities for larger African carnivores. WCS Working Paper No. 24*, New York, USA.

Reid, R.S., Thornton, P.K. & Kruska, R.L., 2004. Loss and fragmentation of habitat for pastoral people and wildlife in east Africa: concepts and issues. *African Journal of Range & Forage Science*, 21(3), pp.171–181.

Rentsch, D. & Packer, C., 2015. The effect of bushmeat consumption on migratory wildlife in the Serengeti ecosystem, Tanzania. *Oryx*, 49(2), pp.287–294.

Riggio, J. et al., 2013. The size of savannah Africa: a lion's (*Panthera leo*) view. *Biodiversity and Conservation*, 22, pp.17–35.

Riley, S.J., DeGloria, S.D. & Elliott, R., 1999. A terrain ruggedness index that quantifies topographic heterogeneity. *Intermountain Journal of Science*, 5, pp.23–27.

Ripple, W. et al., 2014. Status and Ecological Effects of the World's Largest Carnivores. *Science*, 343.

Ripple, W.J. et al., 2016. Bushmeat hunting and extinction risk to the world's mammals. *Royal Society Open Science*, 3(10), p.160498.

Ripple, W.J. et al., 2015. Collapse of the world's largest herbivores, (May).

Ripple, W.J. & Beschta, R.L., 2004. Wolves, elk, willows, and trophic cascades in the upper Gallatin Range of Southwestern Montana, USA. *Forest Ecology and Management*, 200, pp.161–181.

Roelke-Parker, M.E. et al., 1996. A canine distemper virus epidemic in Serengeti lions (*Panthera leo*). *Nature*, 379(6564), pp.441–445.

Roever, C.L. et al., 2013. The pitfalls of ignoring behaviour when quantifying habitat selection N. Roura-Pascual, ed. *Diversity and Distributions*.

Rogala, J.K. et al., 2011. Human Activity Differentially Redistributes Large Mammals in the Candaian Rockies Nation Parks. *Ecology and Society*, 16(3), p.16.

Rondinini, C. & Boitani, L., 2012. Mind the map: trips and pitfalls in making and reading maps of carnivore distribution. In L. Boitani & R. A. Powell, eds. *Carnivore Ecology and Conservation*. Oxford, UK: Oxford University Press, pp. 31–46.

Rondinini, C., Stuart, S. & Boitani, L., 2005. Habitat suitability models and the shortfall in conservation planning for African vertebrates. *Conservation Biology*, 19(5), pp.1488–1497.

Rostro-García, S. et al., 2016. Scale dependence of felid predation risk: identifying predictors of livestock kills by tiger and leopard in Bhutan. *Landscape Ecology*, 31(6), pp.1277–1298.

Rostro-García, S., Kamler, J.F. & Hunter, L.T.B., 2015. To kill, stay or flee: the effects of lions and landscape factors on habitat and kill site selection of cheetahs in South Africa. *PloS one*, 10(2), p.e0117743.

Sachedina, H. & Nelson, F., 2010. Protected areas and community incentives in savannah ecosystems: a case study of Tanzania's Maasai Steppe. *Oryx*, 44(3), pp.390–398.

Sanderson, E.W. et al., 2002. Planning to Save a Species: the Jaguar as a Model. *Conservation Biology*, 16(1), pp.58–72.

Saupe, E.E. et al., 2012. Variation in niche and distribution model performance: The need for a priori assessment of key causal factors. *Ecological Modelling*, 237–238, pp.11–22.

Scantlebury, D.M. et al., 2014. Flexible energetics of cheetah hunting strategies provide resistance against kleptoparasitism. *Science*, 346(6205), pp.79–81.

Schaller, G.B., 1968. Hunting behaviour of the cheetah in the Serengeti National Park, Tanzania. *East African Wildlife Journal*, 6, pp.95–100.

Schaller, G.B., 1972. *The Serengeti Lion*, Chicago and London: University of Chicago Press.

Schneider, A., Friedl, M.A. & Potere, D., 2009. A new map of global urban extent from MODIS satellite data. *Environmental Research Letters*, 4(4), p.44003.

Schuette, P., Creel, S. & Christianson, D., 2013. Coexistence of African lions, livestock, and people in a landscape with variable human land use and seasonal movements. *Biological Conservation*, 157, pp.148–154.

Sedano, F., Gong, P. & Ferrao, M., 2005. Land cover assessment with MODIS imagery in southern African Miombo ecosystems. *Remote Sensing of Environment*, 98(4), pp.429–441.

See, L., Comber, A., et al., 2013. Comparing the quality of crowdsourced data contributed by expert and non-experts. *PloS one*, 8(7), p.e69958.

See, L., McCallum, I. & Fritz, S., 2013. Mapping cropland in Ethiopia using crowdsourcing. *International Journal of Geosciences*, 4(August), pp.6–13.

Senay, S.D., Worner, S.P. & Ikeda, T., 2013. Novel Three-Step Pseudo-Absence Selection

Technique for Improved Species Distribution Modelling. *PLoS ONE*, 8(8).

Sharp, N.C.C., 1997. Timed running speed of a cheetah (*Acinonyx jubatus*). *Journal of Zoology*, 241, pp.493–494.

Shenk, T.M. & Franklin, A.B. eds., 2001. *Modeling in Natural Resource Management: Development, Interpretation, and Application*, Washington D.C.: Island Press.

Shrader, A.M. et al., 2008. Do free-ranging domestic goats show "landscapes of fear"? Patch use in response to habitat features and predator cues. *Journal of Arid Environments*, 72(10), pp.1811–1819.

Sih, A. et al., 1985. Predation, competition, and prey communities: a review of field experiments. *Annual Review of Ecology and Systematics*, 16, pp.269–311.

Sih, A., Englund, G. & Wooster, D., 1998. Emergent impacts of multiple predators on prey. *Trends in Ecology & Evolution*, 13(9), pp.350–355.

Sillero-Zubiri, C. et al., 2015. Saharan cheetah *Acinonyx jubatus hecki*, a ghostly dweller on Niger's Termit massif. *Oryx*, pp.1–4.

Sillero-Zubiri, C. & Laurenson, M.K., 2001. Interactions between carnivores and local communities: conflict or co-existence? In J. L. Gittleman et al., eds. *Carnivore Conservation*. Cambridge, UK: Cambridge University Press, pp. 282–312.

Sillero-Zubiri, C., Sukumar, R. & Treves, A., 2007. Living with wildlife: the roots of conflict and the solutions. In D. Macdonald & K. Service, eds. *Key Topics in Conservation Biology*. Malden, MA: Blackwell Publishing, pp. 255–272.

Soberón, J. & Peterson, A.T., 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics*, 2, pp.1–10. Available at: http://wber.oxfordjournals.org/cgi/doi/10.1093/wber/lhm022.

Stander, P.E., 1992. Cooperative hunting in lions: The role of the individual. *Behavioral Ecology and Sociobiology*, 29(6), pp.445–454.

Stander, P.E. et al., 1997. The ecology of asociality in Namibian leopards. *Journal of Zoology*, 242(2), pp.343–364.

Stein, A. et al., 2015. Panthera pardus. *IUCN Red List of Threatened Species*.

Stensgaard, A.-S. et al., 2009. Virtual globes and geospatial health: the potential of new tools in the management and control of vector-borne diseases. *Geospatial health*, 3(2), pp.127–41.

Stevens, F.R. et al., 2015. Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data. *PloS one*, 10(2), pp.1–22.

Stewart, F.E.C. et al., 2016. Wolverine behavior varies spatially with anthropogenic footprint: implications for conservation and inferences about declines. *Ecology and Evolution*.

Stockwell, D. & Peters, D., 1999. The GARP modelling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science*, 13(2), pp.143–158.

Stokland, J.N., Halvorsen, R. & Støa, B., 2011. Species distribution modelling-Effect of design and sample size of pseudo-absence observations. *Ecological Modelling*, 222(11), pp.1800–1809.

Stoner, C., Caro, T., Mduma, S., Mlingwa, C., Sabuni, G. & Borner, M., 2007a. Assessment of effectiveness of protection strategies in Tanzania based on a decade of survey data for large herbivores. *Conservation Biology*, 21(3), pp.635–46.

Stoner, C., Caro, T., Mduma, S., Mlingwa, C., Sabuni, G., Borner, M., et al., 2007b. Changes in large herbivore populations across large areas of Tanzania. *African Journal of Ecology*, 45(2), pp.202–215.

Strobl, C. et al., 2008. Conditional variable importance for random forests. *BMC Bioinformatics*, 9, p.307.

Sunquist, M.E. & Sunquist, F., 2001. Changing landscapes: consequences for carnivores. In J. L. Gittleman et al., eds. *Carnivore Conservation*. Cambridge, UK: Cambridge University Press, pp. 399–418.

Sunquist, M.E. & Sunquist, F., 2002. *Wild Cats of the World*, Chicago: University of Chicago Press.

Swanepoel, L.H. et al., 2013. Extent and fragmentation of suitable leopard habitat in South Africa N. Pettorelli & V. Penteriani, eds. *Animal Conservation*, 16(1), pp.41–50.

Swanepoel, L.H. et al., 2015. Survival rates and causes of mortality of leopards *Panthera pardus* in southern Africa. *Oryx*, 49, pp.595–603.

Swanson, A. et al., 2014. Cheetahs and wild dogs show contrasting patterns of suppression by lions. *Journal of Animal Ecology*, 83, pp.1418–1427.

Takahata, C. et al., 2014. Habitat Selection of a Large Carnivore along Human-Wildlife Boundaries in a Highly Modified Landscape. *PLoS ONE*, 9(1).

Tapper, S.C., Potts, G.R. & Brockless, M.H., 1996. The effect of an experimental reduction in predation pressure on the breeding success and population density of grey partridges. *Journal of Applied Ecology*, 33, pp.965–978.

Tchuenté, A.T.K., Roujean, J.-L. & De Jong, S.M., 2011. Comparison and relative quality assessment of the GLC2000, GLOBCOVER, MODIS and ECOCLIMAP land cover data sets at the African continental scale. *International Journal of Applied Earth Observation and Geoinformation*, 13(2), pp.207–219.

Terborgh, J. et al., 2001. Ecological meltdown in predator-free forest fragments. *Science*, 294, pp.1923–1926.

Thirgood, S., Woodroffe, R. & Rabinowitz, A., 2005. The impact of human–wildlife conflict on human lives and livelihoods. In R. Woodroffe, S. Thirgood, & A. Rabinowitz, eds. *People and Wildlife*. Cambridge, UK: Cambridge University Press, pp. 13–26.

Thomas, C.D. et al., 2004. Extinction risk from climate change. *Nature*, 427(6970), pp.145–148. Available at: http://www.nature.com/doifinder/10.1038/nature02121.

Thorn, M. et al., 2013. Characteristics and determinants of human-carnivore conflict in South African farmland. *Biodiversity and Conservation*, 22(8), pp.1715–1730.

Thuiller et al., 2004. Effects of restricting environmental renge of data to project current and future species distributions. *Ecography*, 27, pp.165–172.

Thuiller, W., 2004. Patterns and uncertainties of species' range shifts under climate change. *Global Change Biology*, 10(12), pp.2020–2027.

Thuiller, W., Georges, D. & Engler, R., 2014. biomod2: Ensemble platform for species distribution modelling.

Toni, P. & Lodé, T., 2013. Fragmented Populations of Leopards in West-Central Africa: Facing an Uncertain Future. *African Zoology*, 48(2), pp.374–387.

Townshend, J.R.G. et al., 1991. Global land cover classification by remote sensing: Present capabilities and future possibilities. *Remote Sensing of Environment*, 35, pp.243–255.

Townshend, J.R.G. et al., 2011. Vegetation continuous fields MOD44b, 2010 Percent Tree Cover Collection 5.

Treves, A. et al., 2004. Predicting Human-Carnivore Conflict: a Spatial Model Derived from 25 Years of Data on Wolf Predation on Livestock. *Conservation Biology*, 18(1), pp.114–125.

Treves, A. & Karanth, K.U., 2003. Human-Carnivore Conflict and Perspectives on Carnivore Management Worldwide. *Conservation Biology*, 17(6), pp.1491–1499.

Treves, A. & Naughton-Treves, L., 1999. Risk and opportunity for humans coexisting with large carnivores. *Journal of Human Evolution*, 36, pp.275–282.

Tuqa, J.H. et al., 2014. Impact of severe climate variability on lion home range and movement patterns in the Amboseli ecosystem, Kenya. *Global Ecology and Conservation*, 2, pp.1–10.

UN, 2013. *World Population Prospects: The 2012 Revision Volume I: Comprehensive Tables. ST/ESA/SER.A/336*,

UNEP & IUCN, 2015. The World Database on Protected Areas (WDPA). Available at: www.protectedplanet.net.

USAID, 2011. *Uganda environmental threats and opportunities assessment*, Washington D.C.

UWA, 2010. *Strategic Action Plan for Large Carnivore Conservation in Uganda*, Valeix, M. et al., 2012. Behavioural adjustments of a large carnivore to access secondary prey in a human-dominated landscape. *Journal of Applied Ecology*, 49(1), pp.73–81.

Valeix, M. et al., 2009. How key habitat features influence large terrestrial carnivore movements: waterholes and African lions in a semi-arid savanna of north-western Zimbabwe. *Landscape Ecology*, 25(3), pp.337–351.

Van Valkenburgh, B., 2007. Deja vu: the evolution of feeding morphologies in the Carnivora. *Integrative and Comparative Biology*, 47(1), pp.147–163.

Vanak, A.T. et al., 2013. Moving to stay in place: behavioral mechanisms for coexistence of African large carnivores. *Ecology*, 94(11), pp.2619–31.

Vancutsem, C. et al., 2012. Harmonizing and Combining Existing Land Cover/Land Use Datasets for Cropland Area Monitoring at the African Continental Scale. *Remote Sensing*, 5(1), pp.19–41.

VanDerWal, J. et al., 2009. Selecting pseudo-absence data for presence-only distribution modeling: How far should you stray from what you know? *Ecological Modelling*, 220(4), pp.589–594.

Velasquez-Manoff, M., 2016. A National Cure for Lyme Disease. *New York Times Sunday Review*.

Visconti, P. et al., 2011. Future hotspots of terrestrial mammal loss. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 366(1578), pp.2693–702.

Vitousek, P.M. et al., 1997. Human Domination of Earth's Ecosystems. *Science*, 277, pp.494–499.

Van Vliet, N., Nasi, R. & Taber, A., 2011. From the Forest to the Stomach: Bushmeat Consumption from Rural to Urban Settings in Central Africa. In S. Shackleton, ed. *Non-Timber Forest Products in the Global Context*. Springer-Verlag, pp. 209–224.

Warren, D.L., Glor, R.E. & Turelli, M., 2010. ENMTools: A toolbox for comparative studies of environmental niche models. *Ecography*, 33(3), pp.607–611.

Warren, D.L. & Seifert, S.N., 2011. Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications*, 21(2), pp.335–342.

Watling, J.I. et al., 2012. Do bioclimate variables improve performance of climate envelope models? *Ecological Modelling*, 246, pp.79–85.

Watson, F.G.R. et al., 2015. Human encroachment into protected area networks in Zambia: implications for large carnivore conservation. *Regional Environmental Change*, 15, pp.415–429.

Webster, H., McNutt, J.W. & McComb, K., 2012. African Wild Dogs as a Fugitive Species:

Playback Experiments Investigate How Wild Dogs Respond to their Major Competitors. *Ethology*, 118(2), pp.147–156.

Wegmann, M. et al., 2014. Role of African protected areas in maintaining connectivity for large mammals.

Weise, F.J. et al., 2015. A home away from home: insights from successful leopard (*Panthera pardus*) translocations. *Biodiversity and Conservation*, 24(7), pp.1755–1774.

Western, D. & Dunne, T., 1979. Environmental aspects of settlement site decisions among pastoral Maasai. *Human Ecology*, 7(1), pp.75–98.

Western, D., Groom, R. & Worden, J., 2009. The impact of subdivision and sedentarization of pastoral lands on wildlife in an African savanna ecosystem. *Biological Conservation*, 142(11), pp.2538–2546.

Western, D., Russell, S. & Cuthill, I., 2009. The status of wildlife in protected areas compared to non-protected areas of Kenya. *PloS one*, 4(7), p.e6140.

Whittington-Jones, B.M. et al., 2011. Habitat selection by transient African wild dogs (*Lycaon pictus*) in northern KwaZulu-Natal, South Africa: implications for range expansion. *South African Journal of Wildlife Research*, 44(2), pp.135–147.

WHO, 1998. *World Survey of Rabies*, Geneva, Switzerland.

Wilmers, C.C. et al., 2013. Scale dependent behavioral responses to human development by a large predator, the puma. *PloS one*, 8(4), p.e60590.

Wilson, A.M. et al., 2013. Locomotion dynamics of hunting in wild cheetahs. *Nature*, 498(7453), pp.185–9.

Wilson, K. et al., 2005. Measuring and incorporating vulnerability into conservation planning. *Environmental management*, 35(5), pp.527–43.

Wilson, S.M. et al., 2006. Landscape conditions predisposing grizzly bears to conflicts on private agricultural lands in the western USA. *Biological Conservation*, 130(1), pp.47–59.

Wisz, M.S. & Guisan, A., 2009. Do pseudo-absence selection strategies influence species distribution models and their predictions? An information-theoretic approach based on simulated data. *BMC Ecol*, 9, p.8.

Wittemyer, G. et al., 2008. Accelerated human population growth at protected area edges. *Science (New York, N.Y.)*, 321(5885), pp.123–6.

Wolf, C. & Ripple, W.J., 2016. Prey depletion as a threat to the world's large carnivores. *Royal Society Open Science*, 3.

Woodroffe, R., Lindsey, P.A., et al., 2007c. African wild dogs can subsist on small prey: implications for conservation. *Journal of Mammalogy*, 88(1), pp.181–193.

Woodroffe, R., Frank, L.G., et al., 2007b. Livestock husbandry as a tool for carnivore conservation in Africa's community rangelands: a case–control study. *Biodiversity and Conservation*, 16(4), pp.1245–1260.

Woodroffe, R. et al., 2005. Livestock predation by endangered African wild dogs (*lyca*) in northern Kenya. *Biological Conservation*, 124(2), pp.225–234.

Woodroffe, R., 2000. Predators and people: using human densities to interpret declines of large carnivores. *Animal Conservation*, 3(2), pp.165–173.

Woodroffe, R., 2011. Ranging behaviour of African wild dog packs in a human-dominated landscape. *Journal of Zoology*, 283(2), pp.88–97.

Woodroffe, R., Davies-Mostert, H., et al., 2007a. Rates and causes of mortality in Endangered African wild dogs *Lycaon pictus*: lessons for management and monitoring. *Oryx*, 41(2), p.215.

Woodroffe, R. & Frank, L.G., 2005. Lethal control of African lions (*Panthera leo*): local and regional population impacts. *Animal Conservation*, 8(1), pp.91–98.

Woodroffe, R. & Ginsberg, J.R., 1999. Conserving the African wild dog *Lycaon pictus*. I. Diagnosing and treating causes of decline. *Oryx*, 33(2), pp.132–142.

Woodroffe, R. & Ginsberg, J.R., 1998. Edge Effects and the Extinction of Populations Inside Protected Areas. *Science*, 280(2126), pp.2126–2128.

Woodroffe, R., Hedges, S. & Durant, S.M., 2014. To fence or not to fence. *Science*, 344(6179), pp.46–48.

Woodroffe, R. & Sillero-Zubiri, C., 2012. Lycaon pictus. *IUCN 2013. The IUCN Red List of Threatened Species. Version 2013.2*.

Worden, J., Reid, R. & Gichohi, H., 2003. *Land-Use Impacts on Large Wildlife and Livestock in the Swamps of the Greater Amboseli Ecosystem , Kajiado District , Kenya*, Nairobi, Kenya.

Wrogemann, N., 1975. *Cheetah under the sun*, Johannesburg: McGraw-Hill.

Yackulic, C.B. et al., 2013. Presence-only modelling using MAXENT: When can we trust the inferences? *Methods in Ecology and Evolution*, 4(3), pp.236–243.

Yarnell, R.W. et al., 2013. The influence of large predators on the feeding ecology of two African mesocarnivores: the black-backed jackal and the brown hyaena. *South African Journal of Wildlife Research*, 43(2), pp.155–166.

Yu, L. & Gong, P., 2012. Google Earth as a virtual globe tool for Earth science applications at the global scale: progress and perspectives. *International Journal of Remote Sensing*, 33(12), pp.3966–3986.

Zaniewski, A.E., Lehmann, A. & Overton, J.M.., 2002. Predicting species spatial distributions using presence-only data: a case study of native New Zealand ferns.

*Ecological Modelling*, 157, pp.261–280.

Zulu, L.C. & Richardson, R.B., 2013. Charcoal, livelihoods, and poverty reduction: Evidence from sub-Saharan Africa. *Energy for Sustainable Development*, 17(2), pp.127–137.