PREDICTIVE MODELLING OF BENGAL TIGER DISTRIBUTION IN CORBETT TIGER RESERVE (CTR) AND MANAS NATIONAL PARK (MNP)

VANYA JHA
March, 2015

SUPERVISORS:
Dr. S. Nandy
Dr. ir. T. A. Groen
Dr. S. P. S. Kushwaha

PREDICTIVE MODELLING OF BENGAL TIGER DISTRIBUTION IN CORBETT TIGER RESERVE (CTR) AND MANAS NATIONAL PARK (MNP)

VANYA JHA
Enschede, The Netherlands, March, 2015

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: [Name course (e.g. Applied Earth Sciences)]

SUPERVISORS:
Dr. S. Nandy
Dr. ir. T. A. Groen
Dr. S. P. S. Kushwaha

THESIS ASSESSMENT BOARD:

Chairperson (ITC)	:Prof. Dr. Ir. M. G. Vosselman
ITC Professor	:Prof. Dr. V.G. (Victor) Jetten
External Examiner	:Dr. Gautam Talukdar
ITC Supervisor	:Dr.ir. T.A. Groen
IIRS Supervisor	:Dr. Subrata Nandy
	Dr. S. P. S. Kushwaha

ABSTRACT

Species Distribution models (SDMs) are increasingly used in order to define the ecological niche of a species. These models statistically relate the species occurrence data with environmental predictor variables to predict the distribution of a species in a geographical location. In the present study SDMs have been used to predict the distribution of tigers in relation to prey availability in two areas (1) Corbett Tiger Reserve (CTR) and Manas National Park (MNP). The biomod2 package in R including 8 different SDM techniques were employed in this study. Additionally, modelling was also performed within an ensemble forecast framework by using 7 ensemble modelling techniques. The effects of anthropogenic stress (settlements and roads) on the distribution of tigers were assessed in both study areas. Also the levels of anthropogenic disturbance were compared between the two study areas. Models calibrated and validated in CTR were transferred to MNP and vice-versa in order to assess the performance of transferred models. Ensemble models predicted the distribution of tigers in both study areas most accurately and outperformed all individual models. All individual models except SRE performed fairly well in predicting the distribution of tigers in both the study areas. Amongst individual models RF, GLM and ANN consistently performed well across all prey and predator species. Due to high levels of anthropogenic stress in CTR, there are high possibilities of human-tiger conflicts than in MNP.

Key-words: Species distribution models, biomod2, ensemble models, anthropogenic stress, model transferability, Bengal tiger, prey species

ACKNOWLEDGEMENTS

I would like to extend sincere thanks to the following persons and organisations for providing assistance and guidance during the journey of the research and writing of this thesis.

To Dr. S. Nandy, Dr.ir. T.A. Groen and Dr. S. P. S. Kushwaha my supervisors, for the total freedom they bestowed on me and for their timely guidance, support and encouragement throughout the research, I extend my gratitude. The discussions within the circle always remained thought provoking and enriched my thesis. Their constant motivation and ever-encouraging words made the wok possible.

Thanks to all my ITC and IIRS friends.

I sincerely thank Dr. S. Ghosh, for providing access to species data made it possible for me to keep moving ahead.

The most important acknowledgements are reserved for my parents, being teachers themselves, implanted in me a desire for learning from my childhood. This thesis would not have been written without their enduring love.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Understanding the spatio-temporal distribution of species and their interactions with the ecosystem components is a pivotal concept in disciplines like Ecology, Conservation biology and Evolution (Elith et al., 2006; Singh and Kushwaha, 2011). Knowledge about geographic distribution and patterns of biodiversity coupled with in depth information about the processes that drive it at different scales can aid in sustaining it (Skidmore et al., 2006). Such knowledgebase can facilitate decision makers, environmentalists and natural resource managers develop effective strategies for Biodiversity conservation (Niamir, 2009). In today's world, where there is an urgency to conserve endangered species under pressure from high anthropogenic stress, such information would assist in better integration of anthropogenic activities with natural processes (Skidmore et al., 2006; Singh and Kushwaha, 2011).

One such species that demands our immediate attention for conservation is the royal bengal tiger (*Panthera tigris tigris*), and therefore has been selected as the **Target species** for this research. The native population of tigers in India have seriously reduced in the last hundred years. Therefore, developing conservation strategies for this flagship species is of utmost importance. Numerous studies have been carried out on estimating tiger and co-predator population densities and evaluating average number of prey/prey biomass requirement for tigers using camera-trap and sign-based surveys (Karanth and Sunquist, 1995; Karanth et.al., 2004; Karanth et.al., 2006; Harihar et.al., 2009; Karanth and Chellam, 2009). Various studies have also been conducted to explore the Habitat suitability of tigers using Remote Sensing and GIS techniques at landscape level (Imam et.al., 2009; Karanth et.al., 2009; Jhala et.al., 2011; Ghosh, 2013). However, spatially explicit distribution modelling of tiger and its prey to understand the effects of anthropogenic stress on their distribution were minimally studied.

SDMs are now progressing towards inclusion of Biogeography and Phylogeography evolutionary theories together with empirical statistical modelling in order to achieve model realism. Recently, SDM's have witnessed such advancement that it is even possible to predict anthropogenic effects on patterns of biodiversity (Guisan and Thuiller, 2005). Till date such clear and reliable information has been used to explore and investigate complex ecological relationships. Predictive modelling and mapping based on these relationships are so efficient that they can provide alternative and cost effective ways to produce high quality habitat suitability maps that are indispensable for scientific management of wildlife reserves (Guisan and Zimmermann, 2000).

1.1. Research Problem

Traditionally, relationships between species occurrences and the physical environment they thrive in were observed by people and disseminated to the scientific community mostly in the form of qualitative studies (Grinnell, 1904). These studies were largely exploratory in nature and involved intensive fieldwork. Also, the maps produced by these studies are not necessarily useful enough for all applications. This is because the use of species distribution maps by various applications is largely dependent on the scale of the map. SDMs are numerical models that associate observations of species abundance or occurrence with environmental attributes to derive ecological insights on the habitat preference of the species and predict its distribution across landscapes (Elith and Leathwick, 2009). These models are cost effective solutions to derive distribution as well as abundance of species using presence/absence data. It must be noted here that absence data for most species are not available. However, species presence and absence data is a requisite for the most effective species distribution models (Barbet-Massin et al., 2012). Factors such as selection of most influential set of predictors, statistical model selection, considering appropriate scale, sampling methods used to obtain occurrence data, extrapolation extent and understanding the associations between geographic and environmental factors significantly influences the robustness and realism of a model (Elith and Leathwick, 2009). Having understood this aspect of SDMs it is important to realize that certain amount of uncertainty is always associated with each model.

Historically, almost entire Asia was dominated by tigers. Their home range stretched from Caucasus to Siberia and Indonesia. However, over the century tigers have vanished from southwest and central Asia leading to extinction of three tiger subspecies. 93% of the original historic range owned by the tigers is lost today (Sanderson et al., 2006). Over half of the world's tiger population resides in India with population estimates ranging from 1,571 to 1,875 individuals. In the past century there have been many detailed personal accounts of tiger biology and ecology in Indian context by Dunbar-Brander, 1923; Champion, 1929; Corbett, 1944; Schaller, 1967 and Thapar, 1989. Recent years have witnessed a decreasing trend in the population of tigers in India (Chundawat et al., 2011; Jhala et al., 2011). The prime reason that led to the decline of this once flourishing species in India was trophy hunting during the years 1875-1950. Tigers have received official protection since 1970 after years of persecution which drove them to the verge of extirpation (Karanth, 2003). Tigers are territorial by nature and can prosper in a wide variety of natural habitats. For breeding and dispersal tigers require large, intact and fairly connected landscapes with ample amount of large-sized prey (Karanth and Sunquist, 1995; Carbone et.al., 1999; Jhala et.al., 2011).

Most experts believe that tigers are extremely vulnerable to prey biomass depletion, habitat loss and poaching for illegal trade of animal parts (Dinerstein et al., 2007; Chapron et al., 2008). Tiger populations can withstand the stress of poaching and landscape alteration due to anthropogenic activities in areas with abundant prey base and still remain stable (Chapron et al., 2008). Therefore, it is all the more important to understand how the key prey species (mainly large ungulates) are distributed in varied landscapes and how alterations in land use land cover (natural or anthropogenic) may affect its distribution, abundance and diversity. This would then serve as an indicator as to how healthy the population of tigers (and other co-predators) actually is (Karanth, 2003).

.

1.1.1. Bengal Tiger

Ecologically, *Panthera tigris tigris* (Linnaeus, 1758) stands at the apex of food chain and its presence is an indicator of the health of the ecosystem. It belongs to the family 'Felidae' and the order 'Carnivora' and has been classified as 'Endangered' (EN) by the IUCN red list (Chundawat et al., 2011). On an average bengal tigers weigh between 130-230 kg and vary in length from 1.8 to 2.7 m. As discussed in section 1.1, tigers are highly territorial and require large, contiguous habitats with plentiful prey animals to ensure their long-term survival. Major wild prey of tigers in the Indian subcontinent include: sambar (*Cervus unicolor),* wild pig (*Sus scrofa)*, red deer (*Cervus elaphus),* barasingha (*Cervus duvaucelii),* hog deer (*Axis porcinus),* chital (*Axis axis),* nilgai (*Boselaphus tragocamelus),* muntjac (*Muntiacus muntjac),* blackbuck *Antilope cervicapra,* chinkara *Gazella bennettii,* wild water buffalo, *Bubalus arnee,* gaur *Bos gaurus,* goral *Naemorhedus goral,* takin *Budorcas taxicolor,* serow *Naemorhedus sumatraensis*. When preferred wild prey is scarce in a region, tigers are known to even kill calves of elephant *Elephas maximus* and rhino *Rhinoceros unicornis*. It may even prey upon smaller animals like langur, peacock, porcupine and even livestock when hungry. Tigers are also known to be opportunistic killers and occasionally prey upon other carnivore species such as leopard *Panthera pardus*, sloth bear *Melursus ursinus* and dhole *Cuon alpinus* (Karanth, 2003). On an average tigers require about 3,000 kg of prey or 50-60 prey animals annually which roughly translates into about 400-500 ungulates to sustain a single tiger throughout the year, if the natural reproductive rate of ungulates is considered (Schaller, 1967; Sunquist et al., 1999). Unlike lions, tigers are solitary animals and thus are more careful while hunting in order to avoid injuries. An injured tiger would eventually weaken and starve to death because of its inability to hunt.

Next to an abundant prey base tigers require large territories. The size of a territory required by a single tiger could be attributed to mainly three influences: a) Prey-base, b) availability of water; and c) mating opportunities. Therefore, in prey-rich habitats a breeding female tiger's home range may be as small as 10-20 km^2 whereas in areas with scarce prey-base, it may be as vast as 200-300 km^2. Home range of tigers may increase even more in case of poor quality habitats and in this scenario chances of human-tiger conflict increases manifold. The territory of a male tiger is larger which encompasses smaller home ranges of a few females. This ensures prolific breeding and maintains a stable tiger population in the area. In high prey density areas tigers can pack together but still fiercely defend their territories by leaving scent-marks (mostly urine) and scratches on trees. Females give birth to litters of 2-6 cubs nearly every two and half years. Therefore, in a breeding span of 10 years a tigress is capable of raising about 16 cubs. A tigress takes care of her cubs until they are two years of age. During this period she trains them to capture prey and kill it. After the cubs have fully grown, the tigress begins shunting them away. They have turned into transient tigers in search of their own territories. Weak, injured or ageing tigers are killed or driven out of their territories by young transient tigers and the cycle continues (Schaller, 1967; Sunquist et al., 1999; Karanth and Chellam, 2009).

Tigers have adapted to extreme conditions – on one hand they can be found in extremely hot regions where temperature rises up to 48°C (Rajasthan). On the other hand they are known to endure severe cold climate with temperature dropping as low as -35°C (Russia). However they cannot adapt in arid environments where there is scarcity of water. Also, they can be found at altitudes varying from sea level (Sunderban mangroves) to as high as 3,000 m (Bhutan). In India they can be found in almost all forest types – Evergreen/Semi-evergreen forests, Tropical Moist/Dry Deciduous forests, Riparian grasslands and forests of *terai*, mangrove forests and Sub-tropical mixed forests (Karanth, 2003).
From time to time there are cases of natural mortality in tigers due to poisonous snake bites, young cubs killed by wild dogs, leopards or forest fires and infanticide. But female tigers can make up for these losses by producing 4-6 cubs every two years. However, the stability of tiger population is significantly disturbed when humans interfere. Rampant hunting of prey species by humans renders the habitat unfit for tigers to live in due to lack of food and the population depletes rapidly. When a female tiger is killed by poachers, it is not only the death of one individual but elimination of breeding potential of that animal. In other words, an entire family is lost. Human encroachment on tiger homeland leads to severe habitat loss which in turn adversely affects the tiger population (Thapar, 1989; Karanth, 2003).

1.1.2. Selected Prey Species

Eight prey species were studied in this research. It is assumed that these prey species have a very high contribution in the diet of tigers as per the following published papers Karanth and Sunquist (1995) and Karanth (2003).

- *Cervus unicolor* (Kerr, 1792)
  Sambar is the largest deer species that resides in the forests of South-east and South Asia. It has a vast geographical distribution encompassing Sri Lanka, Nepal and India through southern China, Indochina and Burma to Indonesia, Malaysia and Philippines. Males may weigh up to 260 kg but females are relatively smaller. Sambar inhabits diverse habitats in India that range from scrub forests of Rajasthan and Gujarat, moist deciduous forests of Deccan peninsular, Oak and pine forests at the foothills of Himalaya to the semi-evergreen forests in the north-east. Sambar does not have any specialized food requirements and is known to even feed on browse, dry twigs and branches, fallen leaves, fruits and flowers. But, its preferred forage is young green grass to which it would immediately shift if available. It is due to this flexibility in the dietary preference of

Sambar that has enabled it to adapt to such varied forest types. Even though this species is found in widespread habitats, they are not abundant due to over hunting and poaching by humans. Due to its high biomass, this is one of the most preferred prey for tigers and other carnivores. It prefers hilly terrain and water is very important for its survival especially during summers. They are generally forest-dweller preferring high canopy cover. IUCN lists Sambar as vulnerable (VU) because of sustained decline in its population (Schaller, 1967; Bentley, 1978; Sankar and Goyal, 2004; Timmins et al., 2008a)

- Sus scrofa (Linnaeus, 1758)
  It is commonly known as **wild boar** and has been listed as least concern (LC) by IUCN Red list because of its widespread distribution. Today it is found in all continents except Antarctica and other few oceanic islands. Wild boar thrives in a wide variety of habitats from semi-desert to rain forests, grasslands, mangroves, reed jungles and temperate woodlands. It is also highly tolerant to habitat disturbances and therefore can be found in close proximity to agricultural areas for foraging. It prefers relatively dense forests but can be found in open areas like grasslands and scrublands. It is an omnivorous and its diet principally consist of fruits, roots, vegetables, seeds, crabs, earthworms, molluscs, fishes etc. They are gregarious and generally form herds of various sizes between 6-20 individuals. The major threat this specie faces is hunting by humans for food or retaliation for crop damage. It is also a preferred prey for tiger due to its size (Oliver and Leus, 2008).
- *Muntiacus vaginalis* (Boddaert, 1785)
  It is commonly known as Indian Muntjac or **Barking deer** and is commonly found in many parts of south and south-east Asia. It is highly adaptable and inhabits various forest types – evergreen and deciduous forests, dense and open forests, grassland, croplands, scrublands and even secondary forests. It can be found in plains as well as in rugged terrains. It mainly feeds on small seeds, buds, fruits, young grass and leaves. As per a finding by Kushwaha et al., (2004), muntjac preferred habitats with grass and herbs which is highly correlated to logged locations. This preference of logged areas may imply preference for lower altitudes as well. It is regarded as an important prey for tigers. IUCN lists this species as least concern (LC) but also indicates a decline in its population (Timmins et al., 2008b).
- *Bos gaurus* (Smith, 1827)
  It is commonly known as **gaur** or Indian bison. Historically, the home range of gaur was entire mainland of south and south-east Asia. Today, the habitat of this species has seriously contracted due to over hunting for medicinal products and food and therefore has been listed as vulnerable (VU) by IUCN. In India, it can be found in North-east, central India and western Ghats and a few pockets of Bihar and West Bengal. It prefers lower elevated areas (Choudhury, 2002). Schaller (1967) said that this species preferred semi-evergreen, moist and dry deciduous forests. They mostly eat green grasses but also twigs, barks, leaves, fruits, coarse dry grasses and bamboo and thus are both browsers and grazers. Gaur can live in disturbed habitats and are known to co-exist with small, sparsely populated villages. But over hunting has led to a sharp decline in its population (Duckworth et al., 2008).
- *Bubalus arnee* (Kerr, 1972)
  It is commonly known as **water buffalo** or wild buffalo. This species was widely spread throughout south Asia and Europe during the Pleistocene era. In, India they can be found in Madhya Pradesh, Arunachal Pradesh and Assam. The occurrence of water buffalo is strongly linked to water availability. They prefer low lying deciduous forests, alluvial grasslands and riparian forests. In dry season they are frequently found near marshes, water holes and perennial rivers. Very few pure-bred wild water buffalos exist today. IUCN lists water buffalo as endangered (EN) as it is highly vulnerable to diseases due to interbreeding with domestic and feral buffalos, habitat loss and fragmentation (Hedges et al., 2008).
- *Boselaphus tragocamelus* (Pallas, 1766)

It is commonly known as **nilgai** in India. It has been listed as least concern (LC) by IUCN because a population of above 100,000 has been estimated in India alone. They have a preference of agricultural areas, scrubs, dry deciduous forests and arid areas over deserts and dense forests. They mingle easily with domestic cattle and are often found in agricultural areas (Mallon, 2008)

– *Axis axis* (Erxleben, 1777)
  It is commonly known as **chital** or spotted deer. Though, chital flourishes in various habitats it avoids the two extremes – open deserts and dense moist forests. It is characterized as an edge species, because it prefers forest-grassland interfaces. Therefore, deciduous forests (dry/moist) adjoining scrubs or grasslands are ideal habitats for chital. During cold dry months, chital forages in woodlands where it can browse fallen leaf litter, fruits and twigs. But, as and when monsoon arrives, it prefers grasslands. Chital can easily inhabit areas where there is human presence. Because of its smaller size, it is prey to leopards and clouded leopards more than to tigers. IUCN classifies chital as least concern (LC) because its population is widespread (Duckworth et al., 2008)
– *Axis porcinus* (Zimmermann, 1780)
  It is commonly known as **hog deer.** Historically, the habitat of hog deer was vast, contiguous stretches of south-east Asia which have now contracted to isolated populations in these areas. IUCN has classified hog deer as endangered (EN) because vast populations have vanished over the years due to hunting for meat. In India, hog deer mainly inhabits the flood-plains of rivers Brahmaputra and Ganges and the terai grasslands. Hog deer usually prefers tall grasslands that can provide it enough coverage from predators. The highest densities are found in floodplain grasslands and avoids dense forests. It is more of a grazer than a browser and forages on young grasses, fruits, flowers and herbs. Similar to chital it is a small deer and therefore is less preferred by tigers (Timmins et al., 2012).

1.1.3. Selected Sympatric Carnivores

– *Panthera pardus fusca* (Meyer, 1794)
  It is commonly known as **leopard** and is highly adaptable to varied kinds of environment. IUCN lists leopard as least concern (LC) because it is widely distributed. There are a total of 9 recognized subspecies of leopards, some of which are on the brink of extinction. Leopards can thrive in the most diverse habitats including – marshes, deserts, scrublands, forests, grasslands and woodlands. In India, they are found in all forest types and can move up to an elevation of 5,200 m in the Himalayas. The dietary preference of leopard has a broad spectrum ranging from large mammals like deer, gazelle and boars to smaller animals like birds, monkeys, livestock and rodents. In order to protect their prey from other predators they generally drag their prey to branches of tall trees. They are known to coexist with tigers but generally avoid them (Henschel et al., 2008).
– *Neofelis nebulosa* (Griffith, 1821)
  It is commonly known as **clouded leopard** and is generally found in the foothills of Himalaya in Nepal and almost entire mainland Southeast Asia. This species is listed as vulnerable (VU) by IUCN because of the decreasing trend in its population due to illegal hunting for its skin. Clouded leopard habitats are strongly linked with dense tropical forests due to their arboreal characteristic. It is quite possible that due to sever fragmentation this species is driven to use even secondary, disturbed forests as their habitats. They are rarely recorded in scrublands, grasslands and dry deciduous forests. They can be found in higher elevated areas up to 3,000 m. It is smaller than a tiger or leopard but relatively larger than other small cats. It preys upon smaller animals like squirrel, porcupine, pangolin, hog deer etc. It is very hard to spot a clouded leopard but with the advent of camera traps, more occurrences have been documented (Sanderson et al., 2008).

For the sake of convenience, only common names of all the species discussed in the above sections (highlighted in **bold**) will be referred hereafter in this thesis.

1.2. Research Objectives

1.2.1. General Objectives

The main objective of this research is to understand the distribution of bengal tigers in relation to prey availability and to investigate whether there is a difference in the distribution pattern of tigers in MNP and CTR. It also aims at evaluating & testing the performance of the model when it is trained and validated in MNP and projected in CTR and vice-versa. In order to achieve the general objective following specific objectives have been formulated.

1.2.2. Specific Objectives

- To determine the spatial distribution of bengal tigers in MNP and CTR.
- To identify the most important set of predictor variables that influences the distribution of bengal tigers in MNP and CTR.
- To check whether presence of anthropogenic stress is a more important factor in explaining the distribution of tigers in CTR than in MNP.
- To evaluate the performance of the models when they are transferred from CTR to MNP and vice-versa.

1.3. Research Questions

- Do the anticipated explanatory variables (Distance from Anthropogenic stress, Distance from water, Prey availability, NDVI, altitude, slope) significantly contribute in the predictive model for distribution of tigers?
- Is the presence of anthropogenic stress a more important factor in explaining the distribution of tigers in CTR than MNP?

1.4. Research Hypothesis

Hypothesis 1:

H1: The anticipated explanatory variables (Distance from Anthropogenic stress, Distance from water, Prey availability, NDVI, altitude, slope) **significantly contribute** in the predictive model for tiger distribution.

Hypothesis 2:

H1: The presence of anthropogenic stress is a **more important** factor in explaining the distribution of tigers in CTR than in MNP.

1.5. Research Outputs

- Spatial distribution maps for eight prey species in MNP and CTR.
- Combined prey distribution map for MNP and CTR.
- Spatial distribution maps for both co-predators in MNP and CTR.
- Spatial distribution maps for tiger in MNP and CTR.
- Spatial distribution models for all of the above.
- Spatial distribution maps for transferred models in MNP and CTR.

2. STUDY LOCATION, MATERIALS AND METHODS

In this chapter, the materials and methods used have been discussed in detail. The chapter has been structured as follows:

- Study location
- Research workflow
- Species occurrence data
- Explanatory variables
- Predictive statistical modelling

2.1. Study location

Two different study locations were selected for this research: 1) Corbett Tiger Reserve (CTR), Uttarakhand and 2) Manas National Park (MNP), Assam. Both regions form an integral corridor for the Terai-Arc Landscape (TAL) as shown in Figure 2-1 (a). *'Terai'* (meaning 'lowlands' in Sanskrit) is one of the most unique landscapes in the world characterized by wetlands, tall grasslands and Sal (*Shorea robusta*) dominated mixed deciduous forests. TAL inhabits one of the highest density of tigers in the world and therefore holds high conservation significance. Today, this region supports a population of about 3 million people, half of whom live below the poverty line. In order to support their livelihood they exploit the natural resources available in the landscape leading to severe fragmentation and degradation of forests (Chanchani et al., 2014).

Two areas were set for this study in order to compare the distribution pattern of tigers and also to test Model transferability.

2.1.1. Corbett Tiger Reserve (CTR)

Corbett tiger Reserve is located in the foothills of Himalaya in the state of Uttarakhand, India. Historically, this area is uniquely significant because it was declared as the first protected area of the country. In 1972, it was also the first protected area to be declared as a Tiger Reserve encompassing Corbett National Park and Sonanadi Wildlife Sanctuary. It extends from 29º25'-29º40' N to 78º50'-79º50' E with a total geographical area of 1287.64 km^2. Figure 2-1 (b) shows the geographical location of CTR. This region experiences semi-tropical climate which is characterized by hot and humid weather with abundant rainfall (typically monsoonal). Summers last for three months (mid-March to mid-June) and are generally hot with maximum temperature ranging from 37º C to 44º C. Winters start in mid-November and extend up to mid-march with very low temperatures occasionally reaching 0º C. The annual precipitation varies from 1250 mm to 1400mm. The terrain of this area is hilly and undulating and consists of several valleys and ridges. The elevation ranges from 400 m in the south to about 1205 m in north-east of CTR. River Ramganga is the major perennial source of water in CTR along with its tributaries Palain, Mandal and Sonanadi. Also, in 1974 the Ramganga reservoir came into existence with an area of about 82 km^2. It serves as an important source of water for many animals. A total of 140 forest villages (known as *Khattas*) and *Gujjar dera* (local pastoral community) are found in and around the buffer zone of CTR. Only one *Gujjar dera* could be located inside the core zone of CTR. Mixed moist deciduous forests dominated by Sal is the major forest type of CTR with intermittent grasslands. 110 different tree species are found in this area. Also, CTR also harbors a lavish faunal diversity with about 685 birds, 49 mammals and 39 reptiles (Kushwaha et al., 2008).

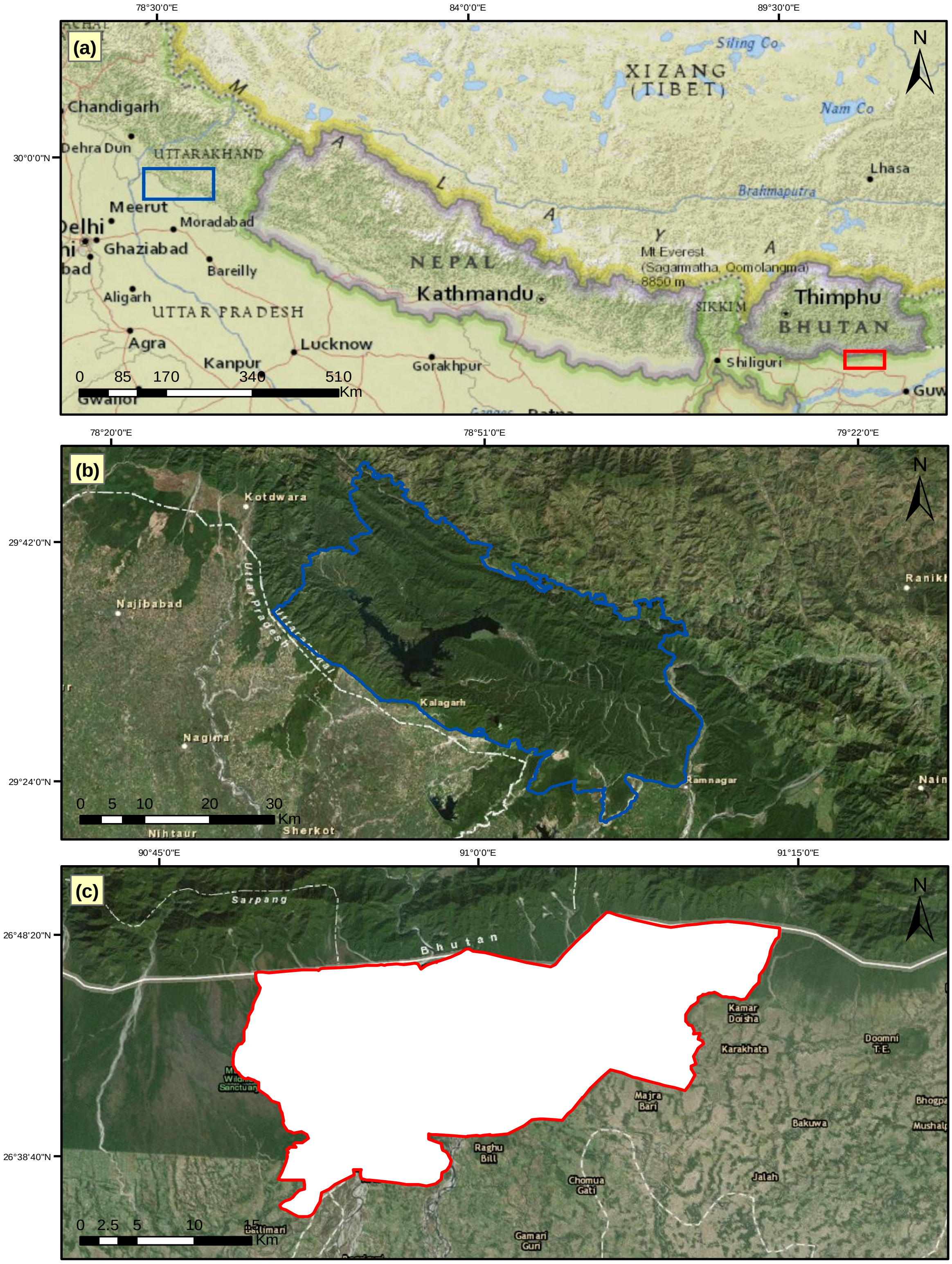


Figure 2-1: Study Area - (a) Terai-Arc Landscape; (b) CTR; (c) MNP

2.1.2. Manas National Park (MNP)

Manas National Park is situated in the state of Assam, India bordered by the International boundary between India and Bhutan in the north, densely populated villages in the south and by reserved forests to the east and west. In 1985, MNP was declared as a World heritage site by UNESCO and it is also known as the core area for Manas Tiger Reserve (which encompasses 2837 km^2). It lies between 26°35'-26°50' N and 90°45'-91°15' E covering a total geographical area of about 500 km^2. Figure 2-1 (c) shows the geographical location of MNP. The prevailing climate of this region is semi-tropical with summers that are hot and humid and mild winters. The temperature rises up to maximum 37° C during summers and to a minimum of 11° C during winters. Therefore, the variation in temperature is minimal across the seasons. Monsoons are relatively longer in Manas and as a result most of the area is flooded during that time which promotes the growth of tall grasses. Altitude ranges from 50 m on the southern boundary to about 200 m in the north which suggests that the terrain is flat and gently sloping to the south. River manas is the lifeline of MNP and it serves as perennial source of water for all wildlife. Other important river systems include *Jongrong, Mora-Manas, Gyati, Rabang* and *Chorpuli Garuchara.* There are numerous seasonal rivers and rivulets that drain from north to south reshaping the terrain. There are thickly populated villages in the periphery of MNP and a substantial part of manas (southern part) is entirely agriculture with pockets of villages. But no villages are found in the core area of MNP. There are three predominant vegetation types in Manas – a) alluvial grasslands, b) mixed moist deciduous forest and c) semi-evergreen forest. MNP inhabits around 476 birds, 60 mammals, 42 reptiles and 7 amphibians. It is also home to many endemic species like hispid hare, golden langur and pygmy hog (Borah et al., 2012).

2.2. Research workflow

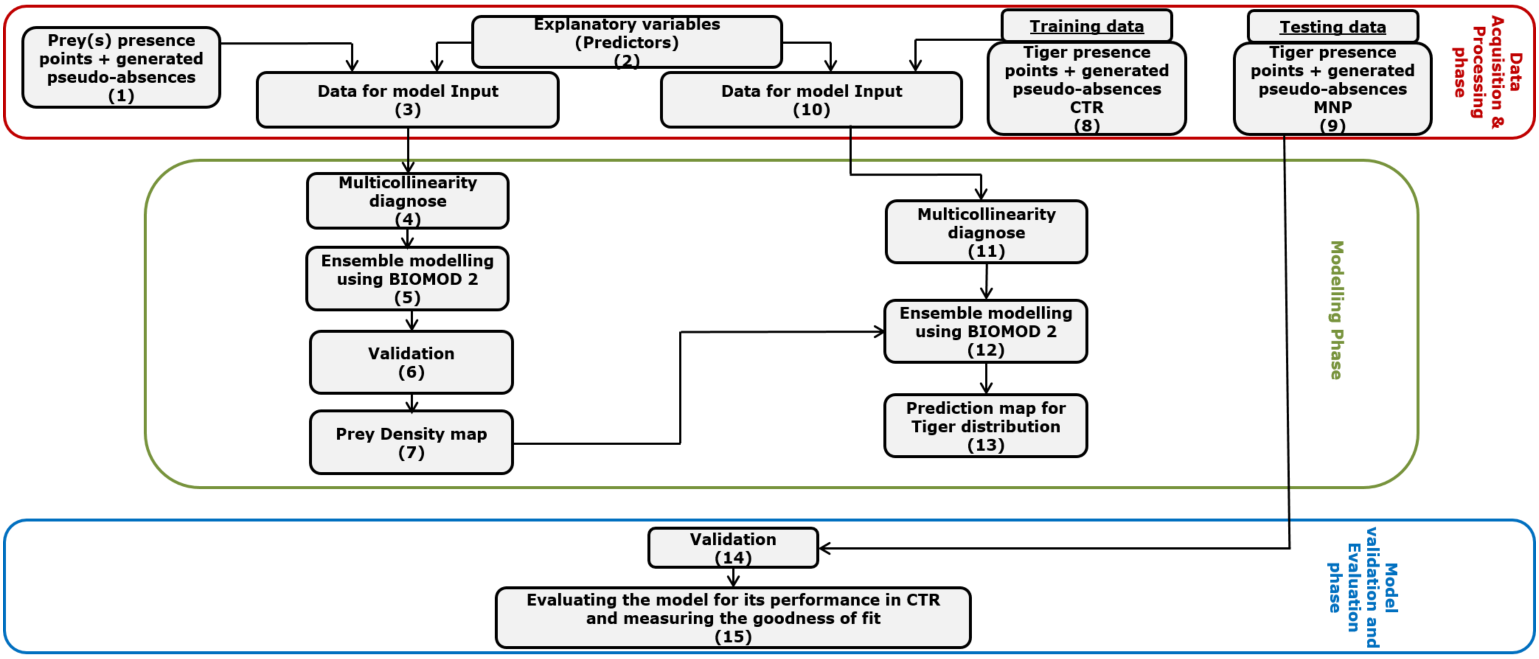


Figure 2-2: Reesaerch workflow and steps

In order to achieve the objectives and sufficiently answer the research questions, the methodology shown in Figure 2-2 was followed. There were 14 steps that were implemented. (1) Preparation of prey presence layer and generation of appropriate pseudo-absences for all prey species, (2) Preparation of required explanatory layers, (3) Preparation of all layers in model compatible format, (4) multicollinearity diagnoses, (5) Performing species distribution modelling (8 models) on all prey species using biomod2 (Thuiller et al., 2009), (6) Validating the results obtained in step 5 on 30% dataset (data separated for validation earlier), (7) Combining all the prey distribution maps to prepare a single prey density map to be inputted for modelling the distribution of tigers, (8) and (9) Preparing the presence + pseudo-absence layer for tigers in both MNP and CTR, (10) Preparing all the layers in model compatible format, (11) multicollinearity diagnoses , (12) Performing ensemble modelling to determine the probability distribution of tigers using biomod2 (Thuiller et al., 2009), (13) Validating the habitat suitability models obtained in step 12 and (14) Evaluating the predictive power of all models and measuring the goodness of fit.

2.3. Species Distribution data

In this section, the specifications of presence/occurrence data for tigers, co-predators and each prey species have described in detail for both CTR and MNP.

2.3.1. Tiger and co-predator occurrence data

– Tiger and leopard presence points in CTR

For CTR, tiger presence data was provided by Dr. Afifullah Khan, Aligarh Muslim University (AMU). The data had been collected in field surveys conducted during October-November, 2003. 431 random plots were laid across the study area which followed stratified random sampling design. Direct/Indirect signs (sightings, pug marks or scats) for tigers were recorded as presence points using a hand-held GPS within a 50 m radius circle around each plot. In total 25 presence points were recorded in this survey. The findings of this research could be found in the following published document Kushwaha et al (2008).
In order to increase the number of presence points, more occurrence data was sought after from different sources. As per another survey conducted during 2005-2007, 614 presence points were collected. The data

was provided by Dr. Afifullah Khan, AMU. No fixed sampling strategy was adopted in this scenario and occurrence data was mostly collected in the buffer zone of CTR. This data was also collected on the basis of direct/indirect signs and expert knowledge on tigers using a hand-held GPS. The habitat type in which the signs were found were also recorded which gave a fairly good idea about the land use preference of tigers in CTR. During this survey, 9 presence points for leopards had also been recorded. These points were also used in this research to study the distribution of leopards in CTR along with tigers.

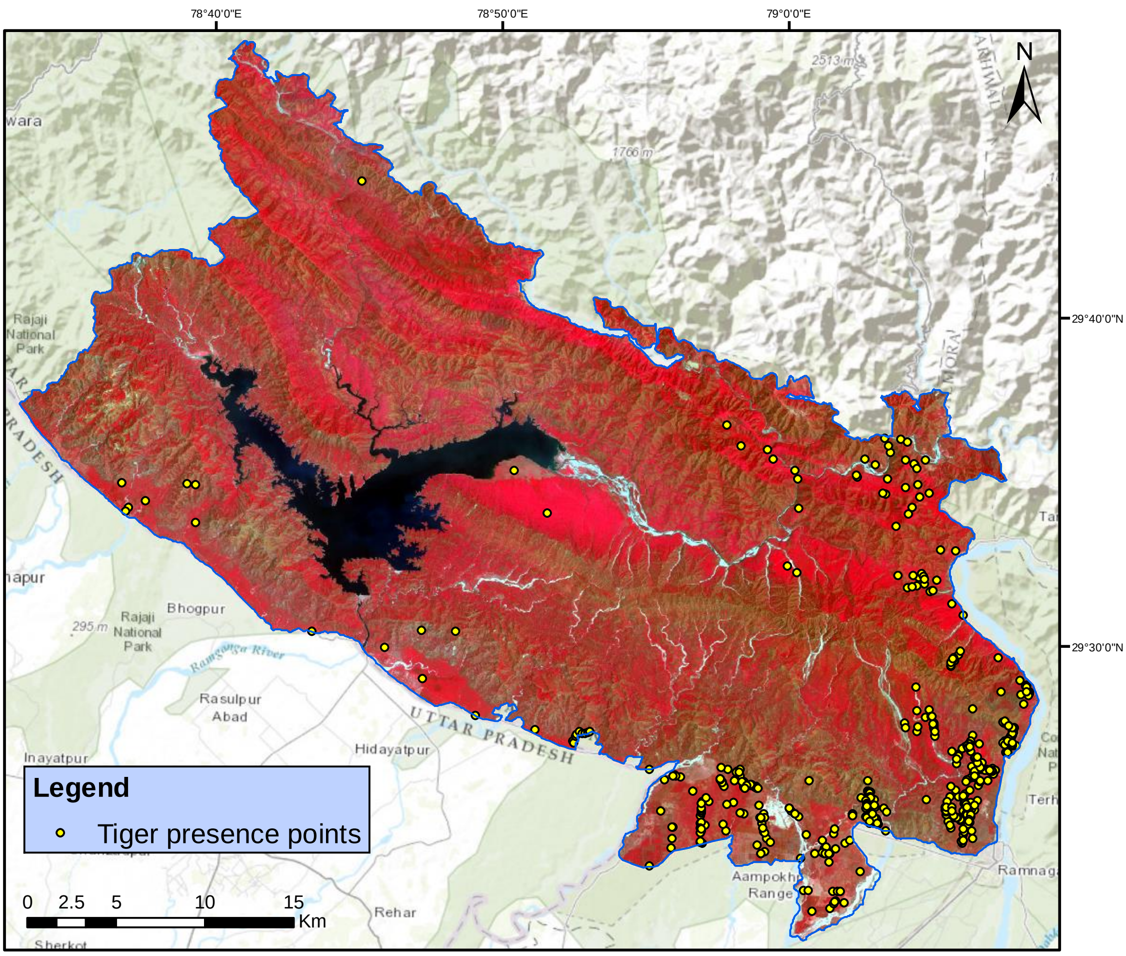
It was observed that around 120 tiger presence points were falling outside the scope of present study area. These points were found to be in the periphery of CTR where numerous villages are present which suggests human-tiger conflicts. After visual interpretation, 126 points were removed because either they were not within the CTR boundary or the points were falling in most unlikely places (For example: a few points were falling inside the reservoir). Finally **513** presence points for **tiger** and **8** presence points for **leopard** were used for modelling the potential distribution of these species. GPS points of tiger and leopard occurrences were requested not to be published. Figure 2-3 shows the distribution of tiger presence locations in CTR. The geographical distribution of leopard presence points have been shown in Figure 3-10 in the results section.

- Tiger, leopard and clouded leopard presence points in MNP

The occurrence data for tigers and its co-predators (leopard and clouded leopard) were provided by Dr. Sonali Ghosh, Wildlife Institute of India (WII). This data was collected using remotely triggered camera traps using capture-recapture framework. This technique of data collection is widely accepted by the scientific community (Karanth et al., 2004). 67 and 78 camera traps were strategically placed across the MNP landscape from November 2011 to February 2011 and November 2012 to February 2013 respectively. Appendix 1 shows the geographical position of the camera traps placed in MNP. Out of 145 camera traps laid in MNP, **tigers** were captured in **77** locations, **leopards** in **31** locations and **clouded leopard** in **16** locations. These were used as presence points for each species in this study. Figure 2-4 shows the distribution of tiger presence locations in MNP. The geographical distribution of leopard and clouded leopard occurrences have been shown in Figure 3-11 in the results section. This technique also take into account the unique stripe patterns of tigers and is able to distinguish individual tigers. This information was not used in the present study. A detailed published report on the camera trapping details in MNP and further analysis on abundance and density estimation of tigers and other species can be found in Borah et al (2012). GPS points of tiger, leopard and clouded leopard occurrences were requested not to be published.

2.3.2. Prey species occurrence data

- <u>Prey presence points in CTR</u>



For CTR, prey presence data was provided by Dr. Afifullah Khan, AMU. The data had been collected in

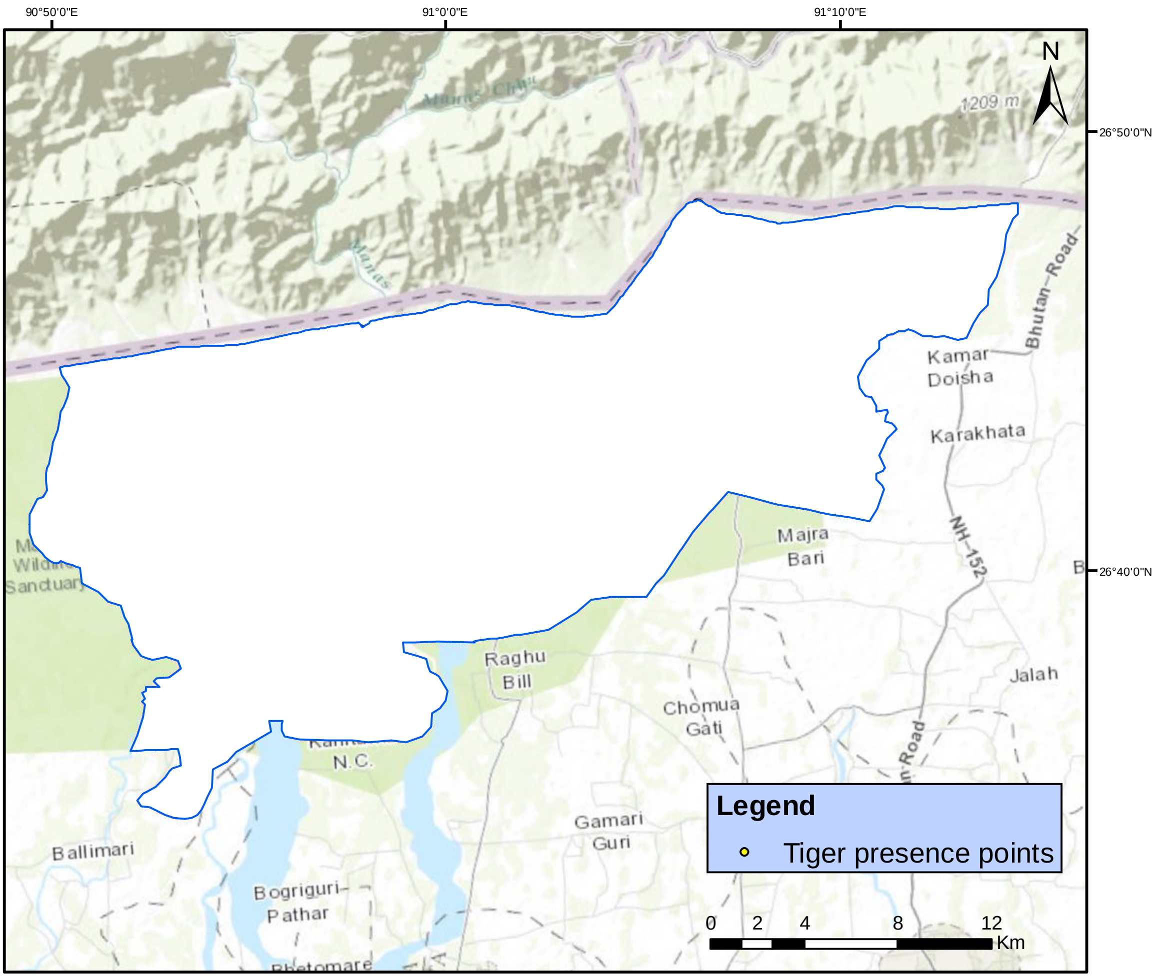


Figure 2-2-4: Distribution of tiger presence locations in MNP

field surveys conducted during October-November, 2003. 431 random plots were laid across the study area which followed stratified random sampling design. Direct/Indirect signs (sightings, pug marks, pellets and dung) for various prey species were recorded as presence points using a hand-held GPS within a 50 m radius circle around each plot. In total 25 presence points were recorded in this survey. The findings of this research could be found in the following published document Kushwaha et al (2008).

Occurrence data for 5 prey species were collected in CTR taking into account specific prey preference of tigers by incorporating expert knowledge and consulting following published papers Karanth and Sunquist (1995) and Karanth (2003). The number of presence points collected in case of each prey are as follows: **nilgai** – **7** presence points; **sambar** - **25** presence points; **wild boar** – **10** presence points; **chital** – **58** presence points and **barking deer** – **40** presence points. GPS points of prey presence occurrences were requested not to be published. The geographical distribution of prey presence points is shown in Figure 3-1

– <u>Prey presence points in MNP</u>

For MNP, prey presence data was provided by Dr. Sonali Ghosh, WII. A total of 22 transects were laid in MNP covering the entire landscape. These transects were traversed multiple times at different dates by forest guards and the GPS location, animal bearing and distance at which the species was spotted was noted down. This information was provided by the Department of forest, Assam and was used to extract transect lines and geographical location of the species recorded using "Bearing Distance to line" tool in ArcGIS 10.1.

Occurrence data for 6 prey species out of 14 species (total number of species spotted in the transect survey) were selected in MNP taking into account specific prey preference of tigers by incorporating expert knowledge and consulting following published papers Karanth and Sunquist (1995) and Karanth (2003). The number of presence points collected in case of each prey are as follows: **water buffalo** – **7** presence points; **sambar** - **6** presence points; **wild boar** – **22** presence points; **gaur** – **21** presence points, **barking deer** – **13** presence points and **hog deer** – **24** presence points. GPS points of prey presence occurrences were requested not to be published.

All the presence point layers were defined in *WGS 1984 UTM zone 44N* and *WGS 1984 UTM zone 46N* projection system for CTR and MNP respectively.

2.4. Explanatory variables

11 predictors were identified as significant in explaining the distribution of tigers taking aid from expert knowledge. These explanatory variables could be categorized into three groups – 1) Biological variables, 2) Anthropogenic variables and 3) Topographic variables. All the layers were defined in *WGS 1984 UTM zone 44N* and *WGS 1984 UTM zone 46N* projection system for CTR and MNP respectively.

2.4.1. Data pre-processing

LISS 3 orthorectified images were used in this research to derive a few explanatory layers. In order to use the image for preparing the layers, radiometric distortions had to be corrected. Therefore radiometric correction was performed by converting the DN image to radiance. The radiance image was then converted to Top of Atmospheric (TOA) reflectance.

2.4.2. Biological variables

Biological variables used in this research pertain to NDVI, vegetation density and prey density. NDVI and vegetation density layers were prepared from the LISS 3 image for the month of November in both CTR and MNP. It is a measure of vegetation density and condition. The distribution of herbivores are highly affected by the type and density of vegetation. Some like sambar prefer dense lush forests whereas species like wild boar prefer scrublands where the vegetation density is relatively lower. Therefore, both these layers are crucial in explaining the distribution of prey species (herbivores). Even tigers are known to prefer vegetation with medium densities where it can have sufficient camouflage and can hunt its prey with ease. The layer prey density is very crucial as it may help in explaining the distribution of tiger the most. Prey density has been derived by combining the probability distribution map of each prey species. There are 3 predictor variables in this group.

2.4.3. Anthropogenic variables an land cover

Anthropogenic variables in this research are limited to Distance to roads and Distance to settlements. Both these variables are extremely important in analyzing and exploring the effects of anthropogenic stress on the distribution of tigers which is also a major research question in this study. Approximate distance to roads (Highways and non-metallic roads within the study area) and settlements were prepared by using the Euclidean distance tool in ArcGIS. Roads and even small villages were extracted by manual digitization using toposheets, high resolution image LISS 4 (5.8 m spatial resolution) and taking cues from Google earth. A few locations of small villages (*Gujjar*) in CTR were also provided by the department of forest, Uttarakhand. In order to maintain consistency in the classification scheme of land cover map used in both study areas, the land cover map prepared by Bhuvan was used in this study. Bhuvan prepares land cover maps for entire India by visual interpretation using LISS 3 images at a sale of 1:50,000 every five years. The land cover map was classified into 2 suitable classes for each species based on expert knowledge. The land cover preferences of each species along with a description of land cover classes is given in Appendix 2. Land use map for the year 2005-2006 has been used for CTR whereas for MNP the map prepared in 2011-2012 has been used. There are 3 predictor variables in this group.

2.4.4. Topographic variables

Topographic variables in this study were derived from SRTM (30 m) data which is freely available online. There are 5 predictor variables in this group – elevation, slope, southness, westness and distance to rivers. Elevation is an important factor that may affect the distribution of tigers. Generally, tigers prefer lower elevated areas because it is easier for them to spot their prey and hunt it. Also, most bulky herbivores (with more meat) prefer lower elevated areas because it is easier for them to move in these regions. These herbivores like nilgai or wild buffalo are ideal prey for the tiger. On the other hand Sambar is also one of the most preferred prey for tiger and it prefers high elevated areas. Therefore, this suggests that although tiger prefers lower elevated areas, in situations of prey scarcity, it may move up to higher elevations in search for food. Slope is another important variable that may affect the distribution of tigers. Most grazing herbivores prefer gentle slope (Schaller, 1967) and the tiger prefers these herbivores as its prey. Tigers are also known to prefer southern aspect particularly during winters (which are quite severe in corbett). Aspect of slope was converted to "Southness" and "Westness values" ranging from 0 to 180 and 1 to 180 respectively. Higher the value more south or west the slope is facing at. Also water is extremely important for tigers to survive, therefore distance to water is another important predictor variable.

2.5. Predictive Statistical Modelling

Guisan and Zimmermann (2000) have defined Species Distribution Models **(SDMs)** as – "*empirical models relating field observations to environmental predictor variables, based on statistically or theoretically derived response surfaces"*. Development of SDMs are intertwined with parallel developments in statistics and computer science supported by sound theoretical background of predictive ecology (Guisan and Thuiller, 2005). Integration of these three disciplines have embarked on improved understanding of Species Distribution modelling. The earliest use of modelling can be traced back to Johnston (1924), who predicted the invasive spread of a cactus species in Australia. This was followed by Hittinka (1963), who assessed the effect of climatic determinants on the distribution of several European species (Pearson and Dawson, 2003). This marked the first stage for SDMs which was based on non-spatial statistical quantification of species-environment relationship based on empirical data (Guisan and Thuiller, 2005). With the advent of technology such as GPS by which one could collect spatial data, species distribution modelling took a giant leap. Space was added as a new dimension to Species distribution models. Such developments lead to spatially explicit statistical and empirical modelling of species distribution.

In the past few years, the use of SDMs to model biodiversity has increased exponentially (Guisan and Thuiller, 2005). Rapid advancement in data mining techniques have impacted SDMs to a great extent. Implementation of relatively new algorithms such as machine learning on SDMs has resulted in improved predictions. There SDMs are now progressing towards inclusion of Biogeography and Phylogeography evolutionary theories together with empirical statistical modelling in order to achieve model realism.

2.5.1. Multicollinearity diagnoses

Strong correlation between multiple predictor variables suggests high multicollinearity which results in high uncertainties in regression models. Multicollinearity induces large standard deviation in the coefficients of regression leading to Type II error. In order to detect collinearity pairwise pearson correlations (r) between each combination of predictor variables is calculated. This gives us highly correlated pairs of variables (ITC_Handouts). Variance Inflation factor (VIF) can be used to detect multicollinearity (Montgomery et al., 1982). It is calculated by the mathematical equation given in the equation below. VIF larger than 10 indicates collinearity may be a problem and hence a variable is removed one by one (ITC_handouts).

$$VIF = \frac{1}{1 - R_i^2}$$

The correlation coefficient and VIF were calculated using the *usdm* package in R (Naimi, 2013). Variables with VIF greater than 10 were identified and deleted.

2.5.2. Selecting pseudo-absences

Records on both presences and absences are required for the most efficient species distribution models. But absence data for a species is rarely available. Therefore, Barbet-Massin et al (2012) have explored various techniques to generate pseudo absences. In their study on how, where and how many pseudo absences to select, the results suggest that random selection of pseudo absences produces the most reliable SDMs. They also explored the effect of number of pseudo absences on the predictive performance of the model. As per their recommendations, in the present study large number (800 to 1000) of pseudo absences were selected for **GLM** giving equal weights to presences and absences. For **MARS** and **FDA** less (100-200) number of pseudo-absences were selected giving equal weightage to both presences and absences and averaging several runs. Finally in, case of **CTA, GBM** and **RF** exactly the same number of pseudo absences were selected as the presences available and averaging several runs (Barbet-Massin et al., 2012). An array was generated giving the specific requirement for pseudo absence selection for each model as discussed above.

Modelling with biomod2

Biomod2 package in R is an ensemble platform for modelling the potential distribution of a species by running 10 different models as well as run all the models within an ensemble forecast framework in order to understand the ecological niche of a species. Currently, 7 techniques for ensemble are available in this package. Apart from this, biomod2 also performs variable importance per model by a randomize procedure, evaluates the performance of models using these metrics – Cohen's kappa, True skill statistics (TSS) and Area Under the Curve (AUC). Out of the 10 modelling techniques available, 8 of them were implemented in this study and will be discussed briefly in the next section.

2.5.2.1. Individual models

- <u>Generalized linear models(GLM)</u>
  This technique is based on fitting a linear relationship between the predictor variables and response variable. If the relationship is not linear biomod2 allows to perform transformation using either polynomial or quadratic function. Model selection is done using Bayseian information Criterion (BIC) and Akaike Information Criterion (AIC). The stepwise procedure helps in removing redundancy in predictor variables if any (McCullagh and Nelder, 1989).

- <u>Generalized boosted models (GBM) or Boosted Regression trees (BRT)</u>
  GBM's try and fit several, simple individual models whose predictions are later on combined in order to provide a more robust response. The BRT algorithm developed by Friedman (2001) is implemented in biomod2. It uses a number of simple regression tress in each model and recursively builds the model by adding more trees, reclassifying the data as per the new tree.

- <u>Classification tree Analysis (CTA)</u>

This technique doesn't rely on the hypotheses about relationship between predictor and response variables. The tree is built by recursively splitting the data governed by a simple rule. Each split is a binary one where the data is partitioned into two groups based on the similarities. It is an iterative algorithm and it seeks to reduce variance as much as possible within the subset (Breiman et al., 1984)

- Artificial Neural Networks (ANN)
  It is a machine learning algorithm which basically establishes a linear relationship between outputs and input with the help of a hidden function. In case of SDMs, the output is the response variable, the input is the explanatory variable and the hidden layer is the hidden composite variable. A linear relation is established between the hidden composite variable and the response variable (Ripley, 1996)

- Flexible Discriminant Analysis (FDA)
  FDA is an extended version of linear discriminant analysis. It is a supervised classification method that combines different models like MARS based on mixture models (Hastie and Tibshirani, 1996).

- Multivariate Adaptive Regression Splines (MARS)
  MARS is useful when the optimum value across various levels of explanatory layers is different. It was introduced by Friedman (1991) and it will identify and estimate the model whose coefficients differ at various levels of explanatory layers.

- Random Forest (RF)
  RF builds an ensemble of un-correlated classification trees and averages them. The number of classification trees built is equal to the number of data present. The validation of each tree is performed by a subset of data that has not been used. Accuracy and variable importance of each tree is done by calculating the rate of misclassification (Breiman, 2001)

- Surface Range Envelope (SRE)
  The concept of SRE is similar to that of BioClim. It works by identifying the maximum and minimum value of each explanatory variable and any area with all its variables falling in this range is considered included (Beaumont and Hughes, 2002)

2.5.2.2. Ensemble models

There are 7 ensemble-model algorithms implemented in biomd2:

- Mean of probabilities: This corresponds to calculating the mean of probabilities over a range of specified models.
- Coefficient of variation of probabilities (CV): This model corresponds to calculating standard deviation (sd) of probabilities over the range of specified models. This ensemble-model is more of a measure of uncertainty. Therefore, higher the value, higher is the uncertainty where the species was observed. CV complements mean probability quite well.
- Confidence interval: This ensemble-model is confidence interval around the probability of means. It complements mean of probability as well. Two probabilities will be calculated in this case – the upper one and the lower one with respect to a user defined threshold *alpha*.
- Median of probabilities: This model corresponds to calculating median of probabilities over the range of specified models. This ensemble-model is slightly sensitive to outliers than the mean of probabilities.

- Models committee averaging: In this ensemble-model the binary predictions are averaged and called committee averaging score. Then in each pixel the sum of 1 is divided by the number of models. It gives the prediction as well as measure of uncertainty.
- Weighted mean of probabilities: In this ensemble-algorithm the wighted mean of probabilities is calculated. The weights are estimated as per the evaluation metric decided by the user (kappa, TSS or ROC). Therefore, better the performance of a model higher the weightage it has the ensemble.

2.5.3. Variable Importance

The variable importance function in biomod2 R package performs a backwards elimination, procedure by removing one variable at a time and analysing the effect on SDM's (Thuiller et al., 2013). This procedure is used to identify a smaller subset of variables that is most influential and increases the models predictive performance. This technique also tries to identify the relation between variables and the prediction. In this research 5 different Presence-absence (PA) set were selected and each PA set was iterated 4 times. Therefore in total, each model iterated 20 times for all the species that were studied in this research.

2.5.4. Model Evaluation

In this research, two statistical measures were used to evaluate the predictive performance of the models: (1) Kappa statistics and (2) Area under curve (AUC) – by calculating the area under receiver operating curve (ROC). Cohen's kappa evaluates the error by quantifying the (a) Commission error and (b) omission error. Commission error measures the absences classified as presences and omission error measures the presences that have been defined as absences. The equation to calculate cohen's kappa is given as:

Model evaluation and comparison is based on an independent dataset. But generally separate independent datasets are not available (Historical data for example). Therefore, techniques like cross-validation or data partitioning are then used. The most commonly used technique for model evaluation is data partitioning where the independent data is divided into two parts – (1) training data to calibrate the model and (2) testing data to validate the model. The most widely accepted range for partition is 70/30 to train and test the model. Data partitioning of 70/30 was used in this study (Franklin, 2010).

In the present study cross-validation was performed wherein the 20 iterations of 70/30 data subset were made. Each time a different subset was used to calibrate and test the model. The total number of runs are equal to the number of subsets made.

AUC measures the goodness of fit of the model on the dataset. AUC is a plot between (1-sensitivity), the commission error in the x-axis versus (sensibility), omission error in the y-axis. The values of AUC range from 0.5 to 1, more you move towards 1, better is the classification (Franklin, 2010).

2.5.5. Model Comparison

Models were compared based on their AUC. The model with highest maximum AUC was selected as the best model. If the maximum AUC was same in case of two or more models, the mean AUC was then considered as a measure of predictive performance of models.

2.5.6. Model Transferability

Just as species occurrence can be forecasted in time, forecasts can be made in space too. With this concept, models calibrated and tested in one geographical location could be used to forecast the occurrence of species in another area too. This is called model transferability. In this study a common dataset (11 variables) was used to perform model transferability in CTR and vice-versa (Wenger and Olden, 2012)

2.6. Assumptions and sources of errors

A few assumptions have been made in this research work and they are as follows:

- It is assumed that the prey species selected to model the distribution of tiger are the only source of food.
- Tiger occurrence data in Manas had been collected for the years 2010-2011 and 2012-13. The datasets have been pooled together with the assumption that there has been no change in the distribution pattern of tigers in the year 2011-12. Similarly, tiger data in Corbett for the years 2003 and 2005-2007 have been pooled together with the same assumption.
- There is sampling bias in the tiger occurrence records of CTR which may induce some uncertainty in the results. The presence points in CTR are over sampled towards the eastern part of CTR and under sampled in the western part of CTR.
- Most prey species in both CTR and MNP have very less presence points which means there is insufficiency of data for prey.
- Almost all camera traps in Manas (althought well spread) were placed on (non-metallic) roads used by forest officials and near the drainages. Therefore the model could pick up strong correlation between the tiger presence point and roads & drainages which is not the actual case. This may lead to erroneous results.

3. RESULTS

In this chapter the main findings of this research have been discussed briefly. Also, the results have been examined and compared with the findings of other similar research works. The chapter has been divided into 4 main sections:

- Analysis of collinearity
- Prey base generation
- Distribution modelling of Bengal tiger
- Sympatric carnivore relationships

3.1. Analysis of collinearity

Multicollinearity tests reveal that there were no collinear variables among the 11 original set of predictor variables in both MNP and CTR. The results of VIF test are shown in Appendix 3. Since no collinearity was found among predictor variables, all 11 variables were considered as a common dataset to perform Model transferability.

3.2. Generated prey base

As discussed in chapter 1, in order to ensure long term survival of tigers, conservation of its prey base is of utmost importance. Tigers can adapt to varied environments but cannot survive in prey scarce regions. Therefore, in this section the outcomes of predictive statistical modelling of individual prey species for both CTR and MNP have been described briefly. Also the combined prey density map which will serve as an input for the distribution modelling of tigers has been explained in this section.

3.2.1. Prey density in CTR

Five prey species were selected in CTR on the basis of expert knowledge about prey preference of tigers. It is assumed that these species contribute about 80% to the diet of tigers (Karanth and Sunquist, 1995). The results of predictive modelling for each species is as follows:

- <u>Barking deer</u>
  Figure 3-1 (a) shows the Habitat suitability map (HSM) for Barking deer in CTR by GLM. As described in chapter 1, barking deer can inhabit varied environments and the HSM of barking deer concurs with this statement. On visual inspection it was noted that the distribution map showed high probability of distribution of barking deer in almost all forest types present in CTR – riverine grasslands, pure Sal forests and deciduous forests. It has also been discussed in chapter 1 that barking deer prefers logged environments. The HSM also shows high probability distribution in areas which are slightly disturbed by anthropogenic activities. Table 3-1 shows the model accuracy for barking deer. It can be seen that all models except SRE have performed very well with maximum AUC above 0.82. GLM had the highest AUC (=0.953) in this case. But the Cohen's kappa was highest in case of GBM (=0.762) and the mean AUC was highest in case of FDA (=0.855). Appendix 4 shows the variable importance graph for barking deer.
- <u>Chital</u>
  Figure 3-1 (b) shows the HSM for chital by RF. As discussed in chapter 1, chital is known to thrive in forest grassland edges. The highest probability of distribution as per HSM is in the vicinity of rivers. These areas are riverine grasslands and are generally associated with forest patches nearby. Also, areas near villages have been classified as high probability areas. As per table 3-2, RF performed the best in case of chital with maximum AUC = 0.924. All models except

SRE performed reasonably well. Although the Kappa accuracy is for chital is not very high. Appendix 4 shows the variable importance graph for chital.

– Nilgai

Figure 3-1 (c) shows the HSM for nilgai by FDA. Nilgai prefers lower elevated areas and is more comfortable near agricultural patches. The HSM for nilgai shows highest probability only in areas near villages and agricultural patches. But this could also be because of insufficient sampling and lack of presence points. The variable importance graph in Appendix 4 shows the most important variable that explains the distribution of nilgai as elevation. Also, as per the response curve of nilgai, higher the elevation, lesser is the probability of finding nilgai there. So there is a negative relationship. As per table 3-3, the maximum AUC in this case is 1 as per three models (MARS, FDA and GBM) which is quite unlikely.

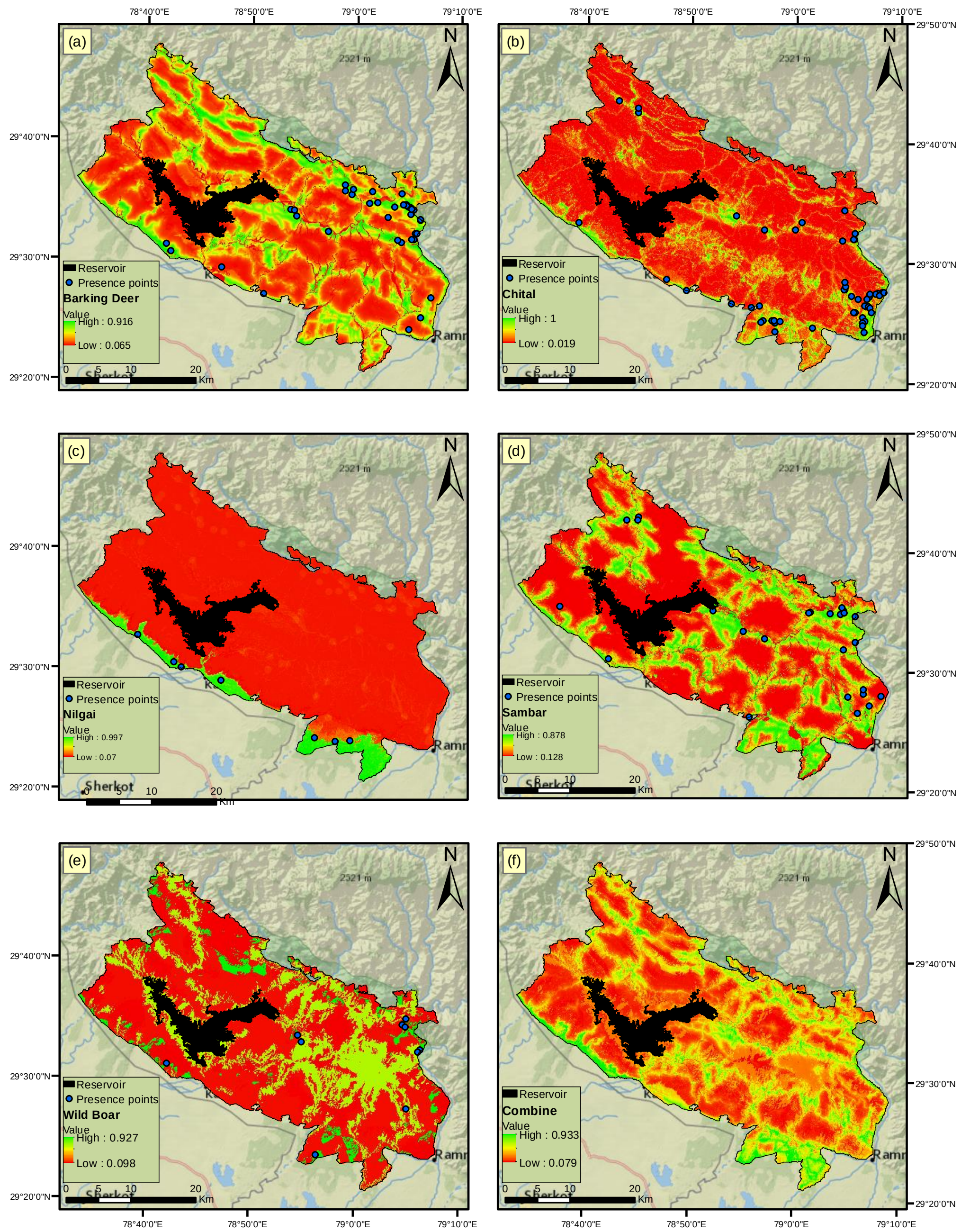


Figure 3-1: Individual prey distribution maps for CTR

- <u>Sambar</u>
  Figure 3-1 (d) shows the HSM for Sambar by GLM. Sambar prefers higher elevated areas and can thrive in a variety of habitats. The HSM shows the probability distribution to be very high in agricultural areas or forest plantation which is quite unlikely because it is generally difficult to spot sambar. But the HSM also shows high distribution in many elevated areas which is the preference of Sambar. As per table 3-4, the models for sambar did not give much accurate results but GLM (AUC=0.782) gave the highest accuracy amongst others. Appendix 4 shows the variable importance for Sambar.

- <u>Wild boar</u>
  Figure 3-1 (e) shows the HSM for wild boar by ANN. It prefers open areas like grasslands and agricultural areas. Most of the distribution of wild boar as per ANN is in the riverine grasslands. The Kappa accuracy is not very high for the wild boar model. ANN performed the best with maximum AUC=0.856. Appendix 4 shows the variable importance for wild boar.

  Figure 3-1 (f) shows the combined prey density map that has been prepared by averaging all individual prey density maps. The resultant map is also a probability distribution map with values ranging from 0 to 1. Higher the value, higher is the probability of finding all 5 prey species and hence higher prey density.

Table 3-1: CTR models accuracy for Barking Deer

CTR models accuracy for Barking Deer					
Models	**Max AUC**	**Min AUC**	**Mean AUC**	**Std AUC**	**Kappa**
SRE	0.772	0.524	0.637	0.080	0.295
CTA	0.821	0.566	0.726	0.059	0.234
RF	0.920	0.744	0.835	0.061	0.717
MARS	0.920	0.527	0.769	0.114	0.717
FDA	0.944	0.757	0.855	0.055	0.717
GLM	0.953	0.678	0.811	0.083	0.702
GBM	0.952	0.662	0.828	0.071	0.762
ANN	0.917	0.611	0.777	0.092	0.691

Table 3-2: CTR models accuracy for Chital

CTR models accuracy for Chital					
Models	**Max AUC**	**Min AUC**	**Mean AUC**	**Std AUC**	**Kappa**
SRE	0.737	0.557	0.637	0.056	0.311
CTA	0.794	0.475	0.679	0.081	0.340
RF	0.924	0.751	0.816	0.046	0.586
MARS	0.865	0.665	0.772	0.059	0.529
FDA	0.882	0.723	0.783	0.049	0.547
GLM	0.867	0.693	0.800	0.048	0.510
GBM	0.900	0.770	0.823	0.036	0.559
ANN	0.805	0.512	0.670	0.074	0.378

Table 3-3: CTR models accuracy for Nilgai

CTR models accuracy for Nilgai					
Models	**Max AUC**	**Min AUC**	**Mean AUC**	**Std AUC**	**Kappa**
SRE	0.500	0.483	0.493	0.006	0.000
CTA	0.983	0.492	0.873	0.147	0.652
RF	0.992	0.875	0.956	0.032	0.792
MARS	1.000	0.008	0.646	0.276	1.000
FDA	1.000	0.917	0.971	0.022	1.000
GLM	0.983	0.425	0.711	0.191	0.659
GBM	1.000	0.825	0.950	0.045	1.000
ANN	0.971	0.358	0.785	0.175	0.483

Table 3-4: CTR models accuracy for Sambar

CTR models accuracy for Sambar					
Models	**Max AUC**	**Min AUC**	**Mean AUC**	**Std AUC**	**Kappa**
SRE	0.748	0.457	0.556	0.090	0.206
CTA	0.771	0.568	0.667	0.054	0.263
RF	0.720	0.477	0.609	0.068	0.377
MARS	0.756	0.367	0.578	0.113	0.377
FDA	0.769	0.478	0.660	0.088	0.473
GLM	0.782	0.479	0.636	0.081	0.390
GBM	0.699	0.537	0.633	0.045	0.297
ANN	0.733	0.510	0.621	0.072	0.404

Table 3-5: CTR models accuracy for Wild Boar

CTR models accuracy for Wild Boar					
Models	**Max AUC**	**Min AUC**	**Mean AUC**	**Std AUC**	**Kappa**
SRE	0.667	0.433	0.547	0.092	0.488
CTA	0.833	0.417	0.680	0.127	0.160
RF	0.797	0.414	0.586	0.106	0.488
MARS	0.783	0.300	0.541	0.155	0.376
FDA	0.764	0.328	0.547	0.126	0.376
GLM	0.808	0.400	0.623	0.120	0.468
GBM	0.794	0.453	0.647	0.092	0.488
ANN	0.856	0.458	0.679	0.114	0.376

3.2.2. Prey density in MNP

- Barking deer
  FDA and RF gave the highest accuracy In terms of highest AUC (=0.875) as shown in table 3-7. Barking deer prefers logged areas but as per the HSM, highest probabilities were in areas with dense forest. Appendix 4 shows the variable importance for wild boar.

- Water Buffalo
  As per table 3-7, ANN, GBM, FDA, MARS and RF have very high accuracies in terms of maximum AUC. This phenomenon is observed when occurrence points are very low (10-15 points). The model with highest mean AUC is GBM and therefore it performs the best in this case. Water buffaloes are often found with herds of Cattle. As per the variable importance graph in Appendix 4. Distance to water and distance to settlements are the most significant variables to determine the distribution of water buffalo.

- Gaur
  As per table 3-8, ANN, GBM, FDA, MARS and RF performed well each having the highest maximum AUC. The highest mean AUC was observed in GBM. It was noticed that for low presence points (10-15 points), GBM performed the best and SRE did not run at all. As per the variable importance of gaur elevation, distance to water bodies and distance to settlements were the most significant variables. But as per the response curve of elevation, higher probabilities were found in case of higher elevated areas which is not true in case of gaur because it prefers lower elevated regions. But the HSM map showed highest probability areas for gaur to be grasslands. Appendix 4 shows the variable importance for gaur

- Hog deer
  As per the HSM given in figure 3-2 (d), none of the grasslands are shown as high probability distribution areas whereas hog deer prefer only floodplains or grasslands. Only a small patch of forest is considered most suitable area for Hog deer. As per table 3-9, FDA performed best with maximum AUC (=0.926). Appendix 4 shows the variable importance for hog deer.

- Sambar
  Figure 3-2 (e), shows the probability distribution map for Sambar in MNP. The areas of high probability distribution are generally in dense forests and at higher elevation, towards the Bhutan hills. Sambar prefers hilly areas and therefore this concurs with our results. Table 3-10 shows that GB performs the best because it has the highest mean AUC (=0.586). Appendix 4 shows the variable importance for sambar.

- Wild boar
  Table 3-11 shows that ANN have performed the best to describe the distribution of wild boar in ANN. Although GBM might have a better performance because it has a higher kappa value as well as higher mean AUC. Stretches of grasslands have been considered best for wild boar as per figure 3-2 (f). Appendix 4 shows the variable importance for wild boar.

  Figure 3-2 (g) shows the combined prey density map that has been prepared by averaging all individual prey density maps. The resultant map is also a probability distribution map with values ranging from 0 to 1. Higher the value, higher is the probability of finding all 6 prey species and hence higher prey density.

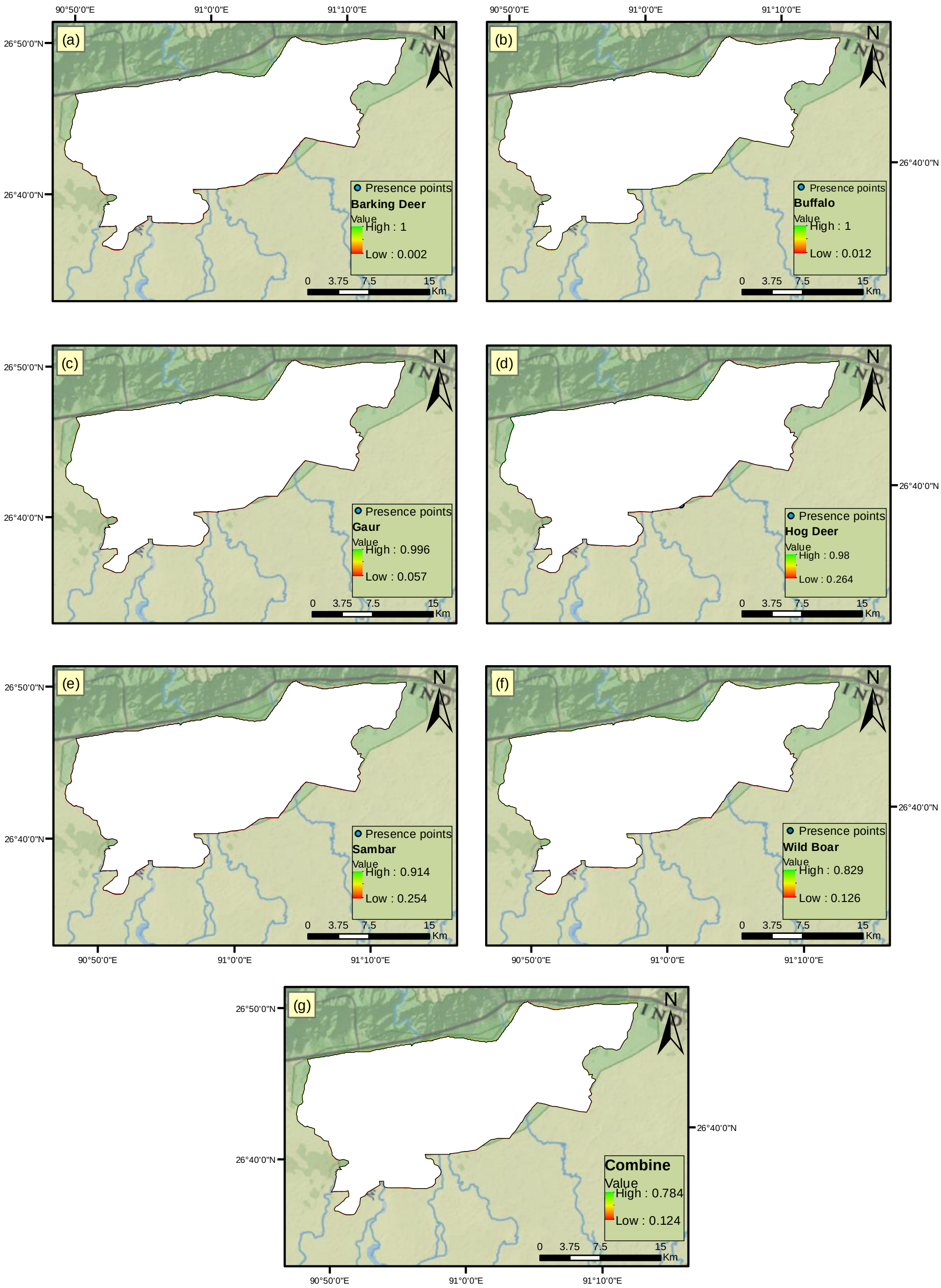


Figure 3-2: Individual prey distribution maps for MNP

Table 3-6: MNP models accuracy for Barking Deer

MNP models accuracy for Barking Deer					
Models	**Max AUC**	**Min AUC**	**Mean AUC**	**SD AUC**	**Kappa**
SRE	0.500	0.383	0.449	0.036	0.000
CTA	0.758	0.283	0.544	0.147	0.313
RF	0.875	0.446	0.665	0.118	0.638
MARS	0.812	0.317	0.585	0.146	0.717
FDA	0.875	0.433	0.655	0.111	0.436
GLM	0.783	0.367	0.551	0.139	0.465
GBM	0.842	0.321	0.673	0.133	0.465
ANN	0.638	0.354	0.503	0.101	0.370

Table 3-7: MNP models accuracy for Buffalo

MNP models accuracy for Buffalo					
Models	**Max AUC**	**Min AUC**	**Mean AUC**	**SD AUC**	**Kappa**
CTA	0.978	0.196	0.686	0.192	0.779
RF	1.000	0.696	0.887	0.115	1.000
MARS	1.000	0.228	0.628	0.250	1.000
FDA	1.000	0.457	0.801	0.171	1.000
GLM	0.804	0.413	0.571	0.126	0.648
GBM	1.000	0.739	0.897	0.096	1.000
ANN	1.000	0.359	0.792	0.177	1.000

Table 3-8: MNP models accuracy for Gaur

MNP models accuracy for Gaur					
Models	**Max AUC**	**Min AUC**	**Mean AUC**	**SD AUC**	**Kappa**
CTA	0.978	0.196	0.686	0.192	0.779
RF	1.000	0.696	0.887	0.115	1.000
MARS	1.000	0.228	0.628	0.250	1.000
FDA	1.000	0.457	0.801	0.171	1.000
GLM	0.804	0.413	0.571	0.126	0.648
GBM	1.000	0.739	0.897	0.096	1.000
ANN	1.000	0.359	0.792	0.177	1.000

Table 3-9: MNP models accuracy for Hog Deer

MNP models accuracy for Hog Deer					
Models	**Max AUC**	**Min AUC**	**Mean AUC**	**SD AUC**	**Kappa**
SRE	0.648	0.383	0.536	0.073	0.295
CTA	0.800	0.469	0.644	0.089	0.471
RF	0.895	0.493	0.758	0.103	0.720
MARS	0.900	0.479	0.711	0.134	0.824
FDA	0.926	0.460	0.731	0.131	0.684

GLM	0.867	0.431	0.693	0.139	0.604
GBM	0.886	0.490	0.740	0.115	0.684
ANN	0.874	0.443	0.720	0.117	0.604

Table 3-10: MNP models accuracy for Sambar

MNP models accuracy for Sambar					
Models	**Max AUC**	**Min AUC**	**Mean AUC**	**SD AUC**	**Kappa**
CTA	0.933	0.200	0.623	0.186	0.605
RF	0.900	0.433	0.726	0.141	0.638
MARS	1.000	0.333	0.526	0.182	1.000
FDA	0.783	0.333	0.622	0.164	0.433
GLM	1.000	0.400	0.586	0.162	1.000
GBM	0.933	0.433	0.765	0.155	0.767
ANN	0.867	0.350	0.534	0.142	0.433

Table 3-11: MNP models accuracy for Wild Boar

MNP models accuracy for Wild Boar					
Models	**Max AUC**	**Min AUC**	**Mean AUC**	**SD AUC**	**Kappa**
SRE	0.686	0.388	0.541	0.077	0.319
CTA	0.774	0.443	0.616	0.102	0.604
RF	0.829	0.574	0.695	0.076	0.684
MARS	0.845	0.393	0.625	0.133	0.684
FDA	0.786	0.440	0.611	0.091	0.604
GLM	0.869	0.295	0.678	0.136	0.604
GBM	0.848	0.524	0.702	0.084	0.720
ANN	0.881	0.395	0.630	0.135	0.549

3.3. Distribution Modelling of bengal tiger

The combined prey density map obtained in the above section is used as an input for distribution modelling of tiger in both MNP and CTR. Also models trained and calibrated in CTR are transferred to MNP and vice versa. The results for distribution modelling of tiger are as follows:

3.3.1. Tiger models in CTR

Figure 3-3 (a) shows the probability distribution map of tigers in CTR. It can be noticed that there is sampling bias in the presence points collected. Most of the points are biased towards the east and most of them are present in the vicinity of human settlements. On visual inspection, it can be noted that areas in proximity of drainages and rivers show higher probability of occurrence. Most of these areas are forest grassland edges or simply grasslands. This suggests that tigers prefer relatively open areas and not very dense forests if abundant prey is available and this concurs to the results obtained by Kushwaha et al (2008) which confirms the importance of canopy density in explaining the distribution of tigers. Although, on visual interpretation of the Habitat Suitability map prepared by Kushwaha et al (2008), it was noted that, the total area that was shown as suitable habitat for tigers was relatively higher than the map prepared in this research. The reason could be insufficient sampling in this case. There was no sampling done in the entire western part of Corbett which has dense forest and is at a higher altitude. Another interesting finding was that in the present study many agricultural patches and forest plantations near settlements were considered suitable habitat for tigers. On the other hand, in the study by Kushwaha et al. (2008) these areas are unsuitable for tigers. There maybe two reasons for this – (1) As discussed in chapter 2, the buffer zone of Corbett has been oversampled, so almost all the points are near the villages or agricultural lands. (2) As per a study by Musavi et al. (2006) there are serious cases of human tiger conflict in the buffer zone (mainly cattle depredation by tigers) in Corbett. Therefore, our results that villages and agricultural patches are highly suitable for tigers may suggest high levels of human tiger conflict in the buffer zone.

Table 3-12 shows the accuracy of all the models in CTR. Random forests have the best predictive power with maximum AUC=0.913 as well as a high kappa value of 0.714. All models except SRE have performed well. This concurs with the study by Lazo (2013), who also found that SRE performed poorly in most cases.

Figure 3-2 shows the variable importance of tiger model in CTR. It can be seen that four variables seem to contribute most in explaining the distribution of tigers in CTR – Elevation, slope, distance to water bodies and distance to settlements. As per the response curves, with an increase in elevation after 600 m the probability drops immediately suggesting that tigers prefer lower elevations. Also, as the slope increases the distribution of tiger also decreases which states that tigers prefer gentle slopes. As the distance from water bodies increases the distribution of tiger decreases which suggest that tigers prefer areas with plenty water sources. All these results concur with the results obtained by Kushwaha et al. (2008). But, as the distance from anthropogenic stress (settlements) increases the distribution of tiger decreases sharply (negative relationship). This however conflicts with the results obtained by Kushwaha et al. (2008) which suggests that "tigers avoid human settlements to the maximum extent possible". However, Corbett experiences high instances of human tiger-conflicts as discussed above, so it could be inferred that tigers frequently visit the buffer zone of CTR to prey upon livestock and cattle. This may be because livestock are relatively easy prey to catch than the wild preys. This reveals an opposite trend in the distribution of tigers in terms of proximity to anthropogenic stress than normally observed. The combined prey density does not significantly contribute in explaining the distribution of tigers in CTR as anticipated. Although the response curve shows an increase in the probability distribution of tigers with an increase in prey density. This could be because the prey base itself is not sufficiently accurate in terms of the distribution of various prey species in the study area.

Models trained and validated in CTR were transferred to MNP. Only the models with the highest AUC were transferred. Figure 3-3 (b) shows the probability distribution of tigers in MNP by CTR RF transferred model. On visual interpretation, it may be noted that the transferred model performs fairly well in MNP. The grassland and forest patches show the highest probability of tiger distribution which are the natural habitat of tigers. An interesting thing to note is that the CTR transferred model in MNP shows low probability distribution in areas near settlements and agriculture. But the same model shows very high probability distribution near villages and agriculture in CTR.

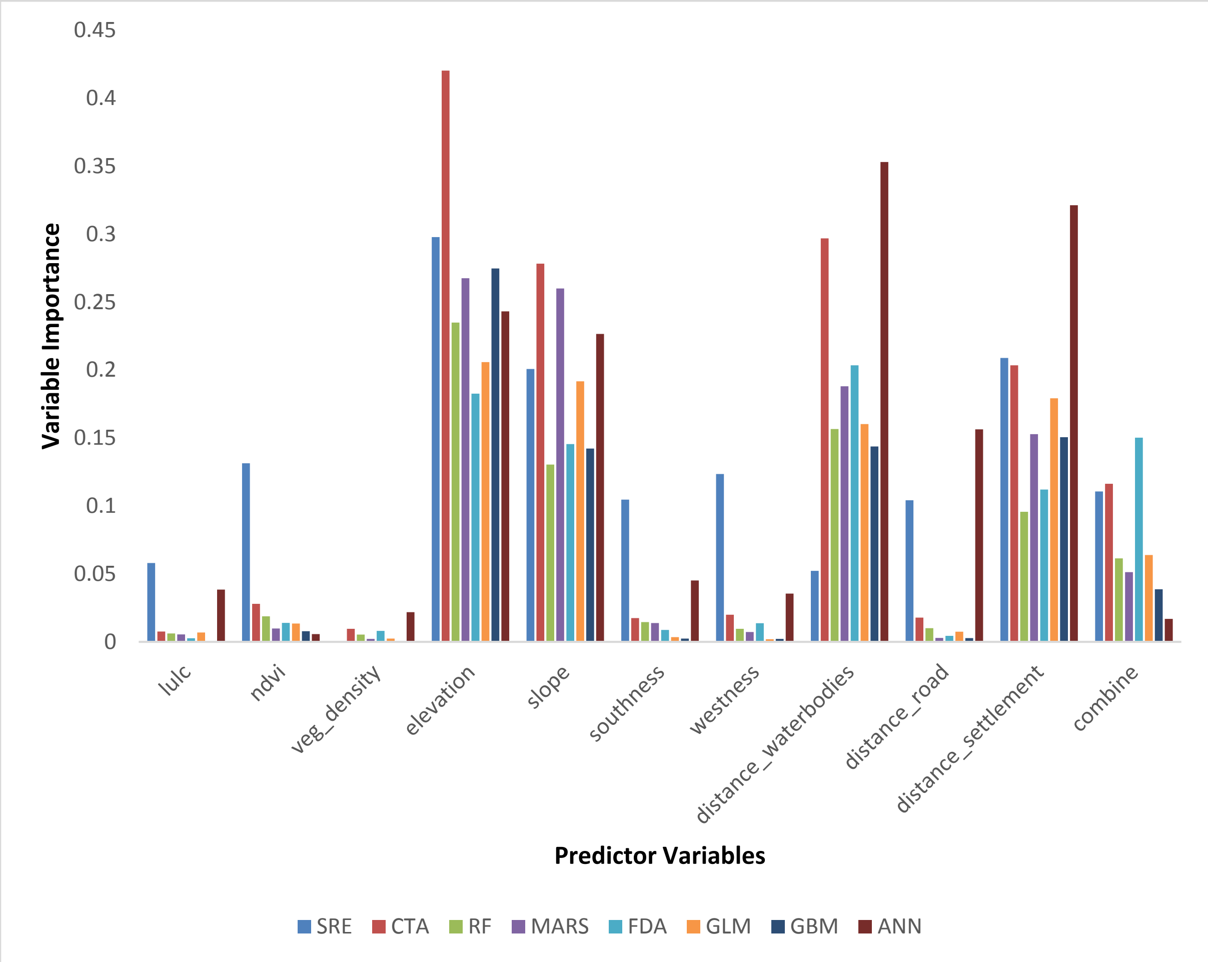


Figure 3-3: CTR variable importance for tiger

Table 3-12: CTR models accuracy for Tiger

CTR models accuracy for Tiger					
Models	**Max AUC**	**Min AUC**	**Mean AUC**	**SD AUC**	**Kappa**
SRE	0.690	0.612	0.656	0.022	0.348
CTA	0.837	0.746	0.808	0.022	0.603
RF	0.913	0.881	0.897	0.010	0.714
MARS	0.895	0.825	0.854	0.019	0.609
FDA	0.875	0.821	0.848	0.016	0.597
GLM	0.872	0.822	0.843	0.014	0.612
GBM	0.900	0.865	0.881	0.011	0.671
ANN	0.862	0.718	0.817	0.045	0.605

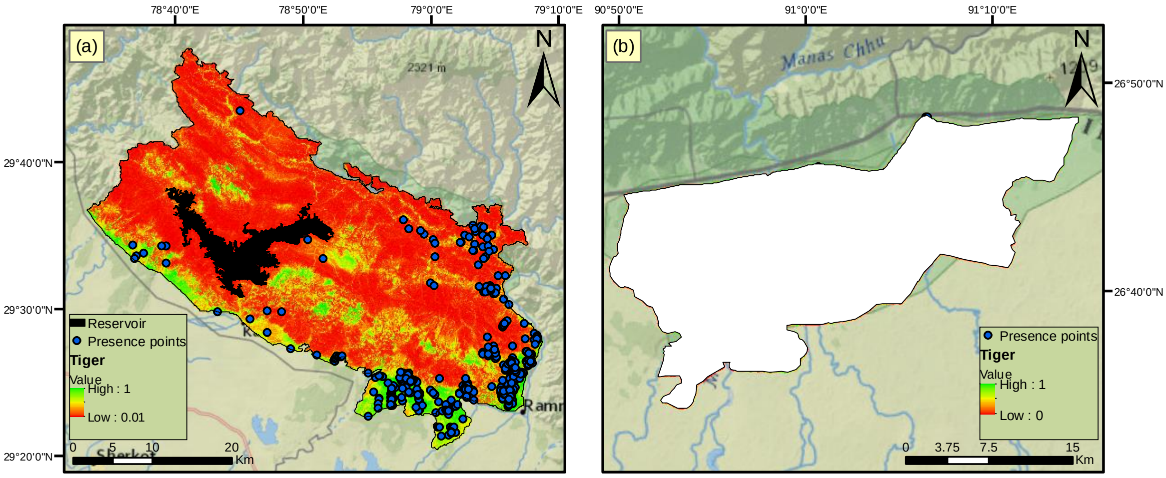


Figure 3-4: Distribution maps for tiger in CTR and corresponding transferred model in MNP

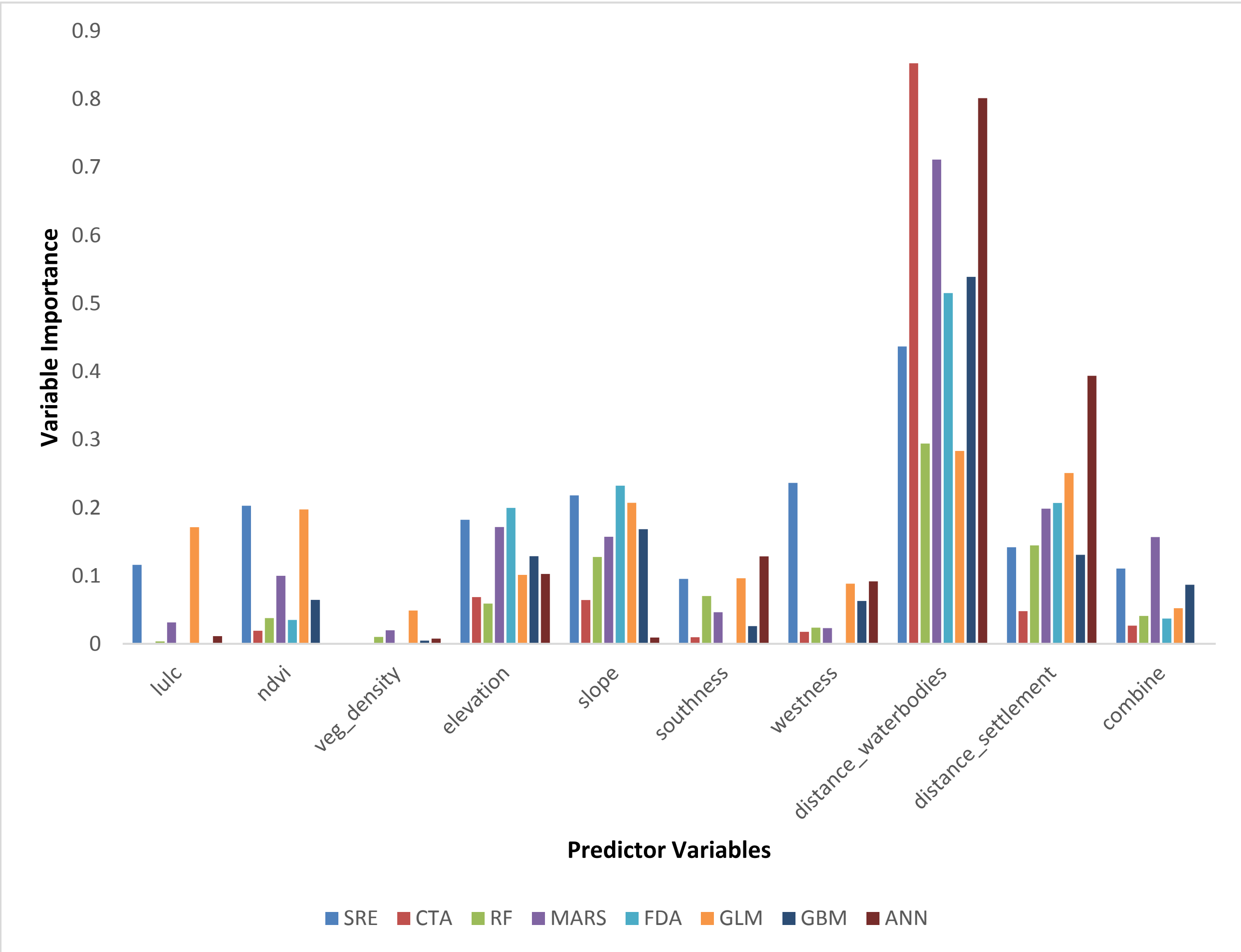


Figure 3-5: MNP variable importance for tiger

3.3.2. Tiger models in MNP

Figure 3-5 (a) shows the probability distribution map of tigers in MNP. According to the map highest probability distribution areas are mostly grasslands and very few patches of forests. The probability distribution was very low near settlements and agricultural areas. Habitat suitability of tigers in Indo-Bhutan Manas Tiger Conservation Landscape (IBMTCL) was studied by Ghosh (2013). The habitat suitability map produced in the present study is very similar to the map produced by that of Ghosh (2013) except that many patches of forests (mostly deciduous) were not classified as highly suitable habitats for tigers in the present study. Not much study has been conducted on the habitat suitability of tigers and other species in Manas.

Table 3-13 shows the accuracy of all the models in MNP. MARS had the best predictive power with maximum AUC=0.913 as well as a maximum kappa value of 0.549. RF and GBM also performed fairly well. GBM had the highest mean AUC among all other models.

Figure 3-4 shows the variable importance of tiger model in MNP. As per the graph two predictor variables have the highest contribution in explaining the distribution of tigers in MNP – Distance to water bodies and Distance to settlement. Elevation and slope fairly contribute to the predictive model. According to the response curve of distance to water bodies, there is an increase in the probability distribution of tigers with increase in distance to water bodies up to 500 m, with further increase in distance the probability reduces. This suggests that tigers prefer vicinity of water bodies. With increase in distance to settlements the probability of tiger also increases which concurs with the studies Kushwaha et al, (2008) and Ghosh (2012). The combined prey density did not contribute much in explaining the distribution of tigers in MNP as hypothesized. This might be because the distribution maps for each prey was not accurate because of insufficiency of presence points. Also as per the response curve the distribution of tiger decreases in high prey density areas.

Models trained and validated in MNP were transferred to CTR. Only the models with the highest AUC were transferred. Figure 3-5 (b) shows the probability distribution of tigers in CTR by MNP MARS transferred model. On visual interpretation, it may be noted that the transferred model does not perform well in CTR. Most of the area is classified as unsuitable for tigers whereas the core zone of Corbett is known to inhabit numerous tigers.

Table 3-13: MNP models accuracy for Tiger

MNP models accuracy for Tiger					
Model	**Max AUC**	**Min AUC**	**Mean AUC**	**SD AUC**	**Kappa**
SRE	0.752	0.510	0.638	0.069	0.408
CTA	0.733	0.557	0.666	0.055	0.240
RF	0.859	0.507	0.738	0.086	0.469
MARS	0.902	0.493	0.688	0.103	0.549
FDA	0.810	0.543	0.679	0.081	0.457
GLM	0.821	0.510	0.664	0.080	0.339
GBM	0.879	0.570	0.767	0.079	0.547
ANN	0.789	0.345	0.627	0.122	0.321

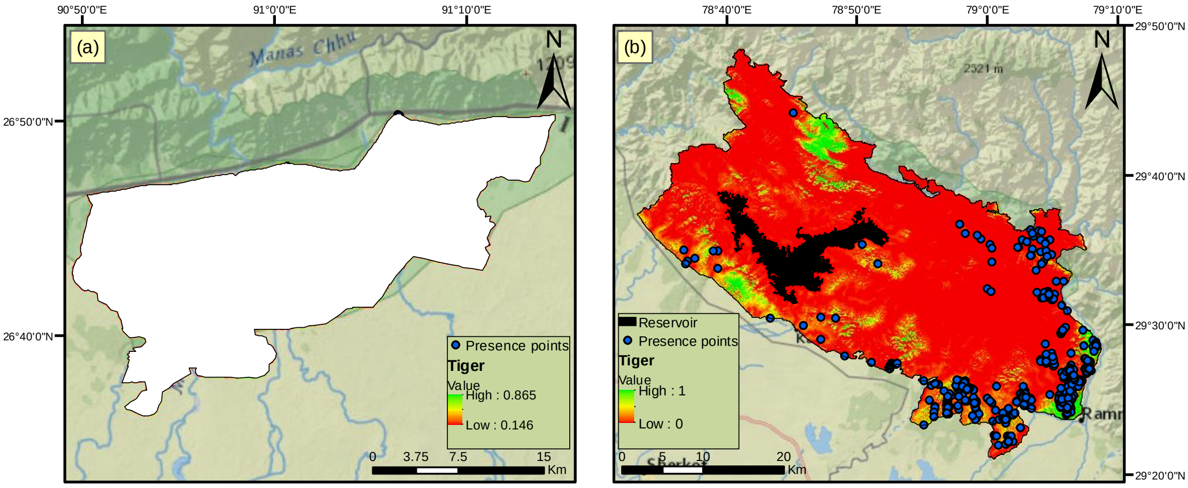


Figure 3-6: Distribution map for tiger in MNP and corresponding transferred model in CTR

3.3.3. Is there any difference in the distribution pattern of tigers between CTR and MNP?

The level of anthropogenic stress is higher in CTR as there are approximately 140 villages within the park boundary. On the other hand the levels of anthropogenic stress is slightly less in MNP. Undoubtedly there are villages in Manas but not as many as in CTR. The distribution pattern of tigers in CTR is affected by presence of human settlements. There have been numerous reports on human-tiger conflict and cattle depredation by tigers in the buffer zone of CTR. One outcome of the present study is that a positive relation exists between tiger presence and human settlements in CTR. So, the anthropogenic stress does affect the distribution pattern of tigers in CTR. Although, this could also be because the sampling of occurrence points was concentrated in near settlements. But other studies have indeed reported the problem of livestock depredation of tigers in the buffer zone of Corbett (Musavi et al., 2006) which supports the results obtained in this research.

MNP on the other hand has had its own share of anthropogenic stressors affecting the landscape. The major one being decades of insurgency problems prevailing in Manas. But, the situation has greatly improved since then and the population of tigers are not affected (Ghosh, 2013). Therefore, the levels of anthropogenic stress is relatively less in manas. The presence of villages and agricultural practices does not affect the distribution of tigers as it does in Corbett. Results show that distance to settlements has a positive effect on tigers which means more the distance between villages higher is the probability occurrence of tigers. Therefore, the distribution of tigers is more natural in MNP.

3.3.4. Ensemble modelling of Bengal tiger distribution in CTR and MNP

Ensemble models are known to be more robust than individual models (Naimi, 2015). Combined predictions of multiple models are more accurate than at least half of the original models (Araujo and New, 2007). Ensemble models have lower standard deviation and high performance and therefore have higher accuracy than individual models (Naimi, 2015).

Figures 3-6 and 3-8 shows the variable importance of all ensemble models for tiger distribution in CTR and MNP respectively. The most important predictor variables are same as that of individual models.

Tables 3-14 and 3-15 show the accuracy of ensemble models for tigers in CTR and MNP respectively. It can be seen that ensemble models have clearly outperformed individual models. Out of the seven ensemble techniques, Ensemble by weighted mean of probabilities have performed the best with AUC=0.922 and kappa=0.546 in CTR and AUC=0.959 and kappa=0.452 in MNP.

Figures 3-7 (a) and 3-9 (a) show the probability distribution maps by ensemble by weighted mean for CTR and MNP respectively. On the other hand, Figures 3-7 (b) and 3-9 (b) show ensemble by standard deviation, which is a measure of uncertainty. It can be seen that in both CTR and MNP the areas where the probability distribution is high, the uncertainty is low.

3.3.4.1. CTR

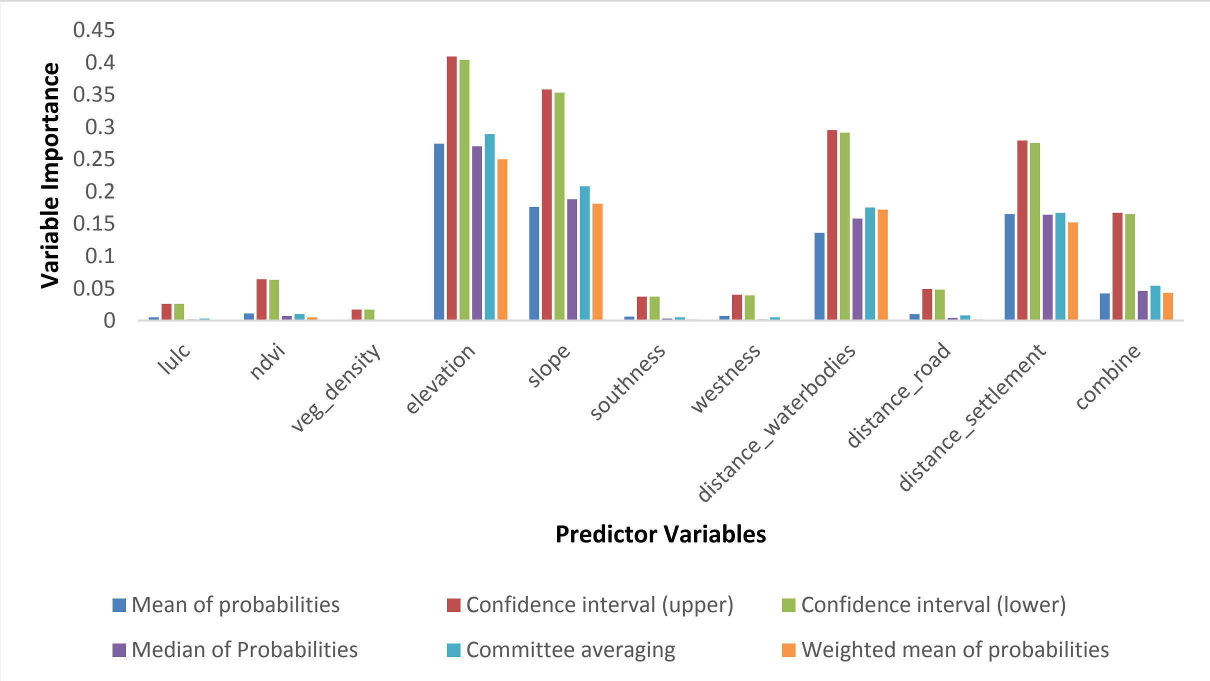


Figure 3-7: CTR variable importance for tiger

Table 3-14: Ensemble models accuracy for Tiger (CTR)

Ensemble models accuracy for Tiger (CTR)		
Ensemble models	**AUC**	**Kappa**
Mean of probabilities	0.912	0.533
Confidence interval (upper)	0.911	0.53
Confidence interval (lower)	0.913	0.534
Median of Probabilities	0.902	0.522
Committee averaging	0.915	0.542
Weighted mean of probabilities	0.922	0.546

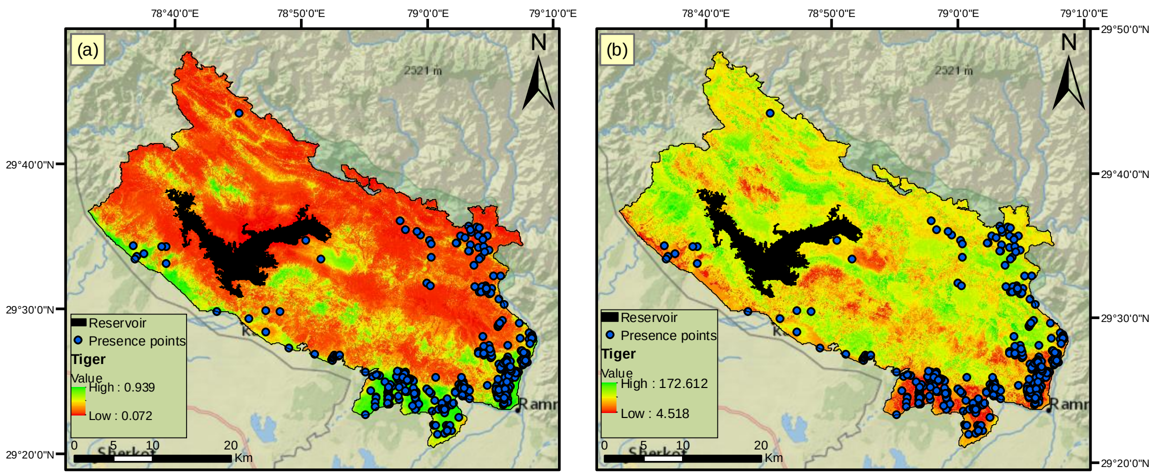


Figure 3-8: Ensemble models

3.3.4.2. MNP

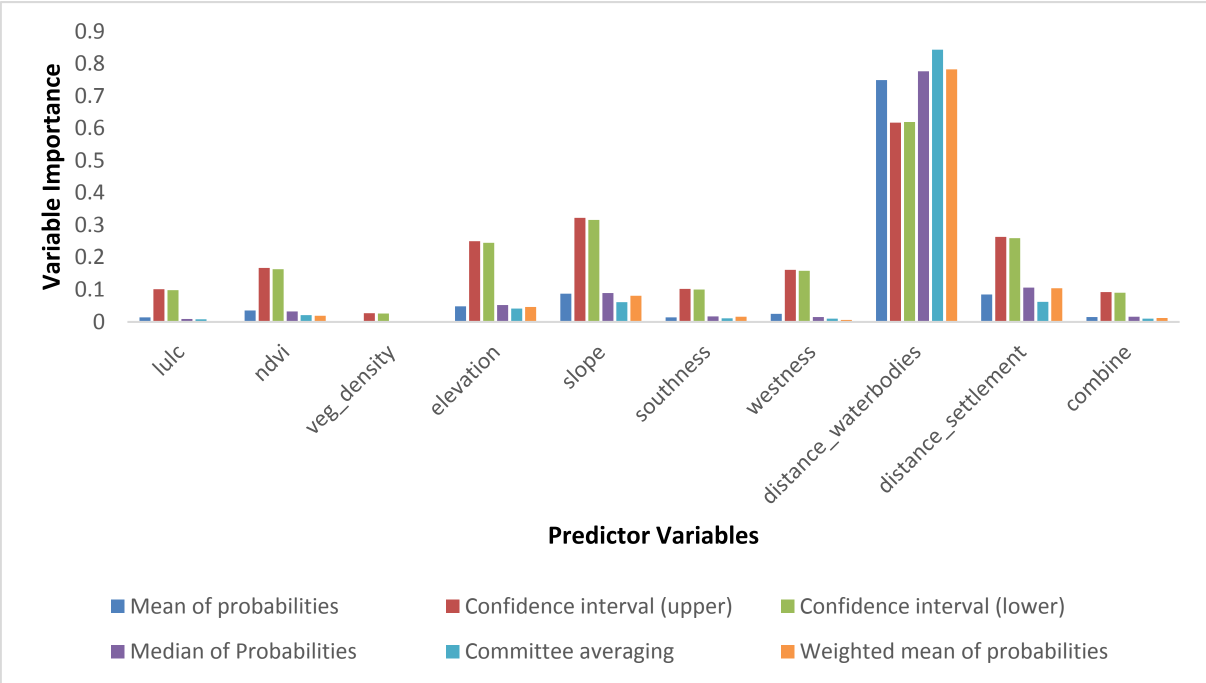


Figure 3-9: MNP variable importance for tiger

Table 3-15: Ensemble models accuracy for Tiger (MNP)

Ensemble models accuracy for Tiger (MNP)		
Ensemble model	**AUC**	**Kappa**
Mean of probabilities	0.95	0.461
Confidence interval (upper)	0.946	0.431
Confidence interval (lower)	0.954	0.473
Median of Probabilities	0.934	0.362
Committee averaging	0.932	0.36
Weighted mean of probabilities	0.959	0.452

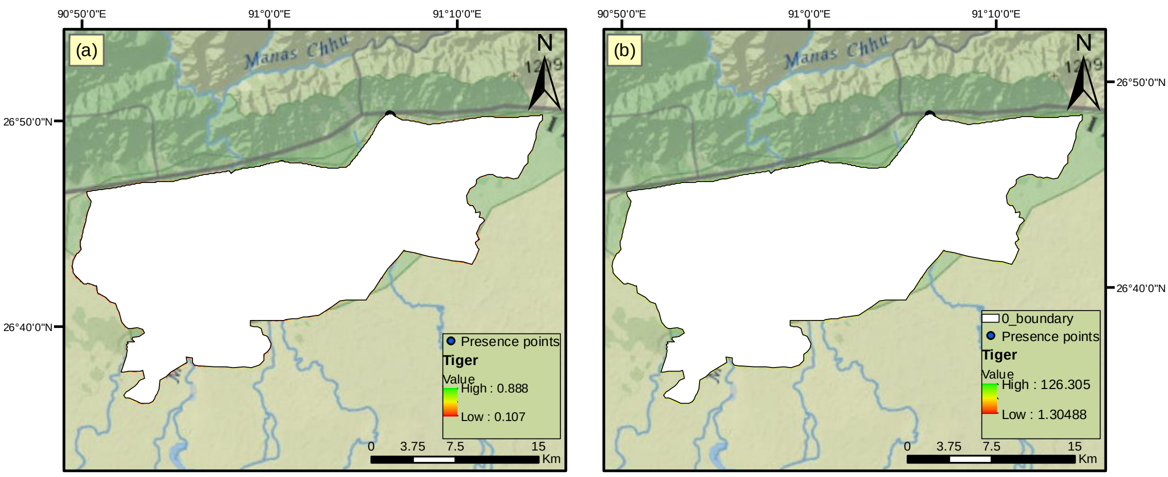


Figure 3-10: Ensemble models

3.3.5. Comparing the distribution of tigers in both study areas

3.4. Sympatric carnivore relationship

Tigers are known to co-exist with other carnivores such as leopards and clouded leopards. In this section the distribution of leopards and clouded leopards are modelled in order to understand their ranges in CTR and MNP. Also it is an attempt to look at the overlap in the home ranges of sympatric carnivores.

3.4.1. Distribution modelling of leopard in CTR

Table 3-16 shows accuracy of the model for leopards in CTR. FDA have performed the best with Maximum AUC=0.983 and kappa value of 0.659. Other models which have high predictive power are MARS, GBM and ANN.

Figure 3-10 shows the probability distribution map for leopards in CTR. Very little area has been classified as high probability distribution for leopards which is not the case. This is because of poor and insufficient sampling, only 8 leopard presence points were available and all of them were concentrated in southern part near settlements. The variable importance of leopards is shown in Appendix

Table 3-16: CTR models accuracy for Leopard

CTR models accuracy for Leopard					
Model	**Max AUC**	**Min AUC**	**Mean AUC**	**SD AUC**	**Kappa**
SRE	0.500	0.475	0.491	0.008	0.000
CTA	0.867	0.367	0.664	0.150	0.483
RF	1.000	0.329	0.735	0.159	0.792
MARS	0.971	0.242	0.657	0.242	0.659
FDA	0.983	0.225	0.738	0.211	0.659
GLM	0.742	0.417	0.580	0.124	0.483
GBM	0.950	0.208	0.785	0.187	0.659
ANN	0.958	0.112	0.611	0.220	0.483

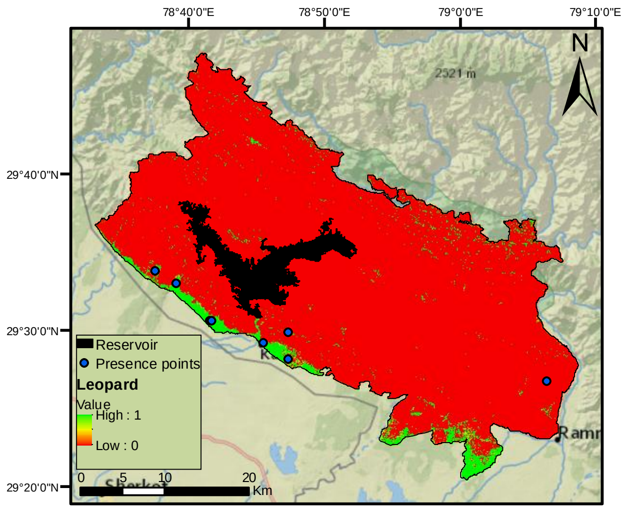


Figure 3-11: Distribution of leopards

3.4.2. Distribution modelling of leopard and clouded leopard in MNP

Tables 3-17 and 3-18 shows accuracy of the model for leopards and clouded leopard in MNP. GLM have performed the best for leopards with Maximum AUC=0.859 and kappa value of 0.558. RF have performed the best for clouded leopards with Maximum AUC=0.844 and kappa value of 0.476.

Figures 3-11 (a) and (b) shows the probability distribution map for leopards and clouded leopards in MNP respectively. Mainly grasslands and small forest patches have the highest probability distribution for both species. The variable importance of leopards and clouded leopards are shown in Appendices

There is certainly an overlap in the home ranges of tigers and its sympatric carnivores but not much inference can be drawn by analysing these maps because home range is dependent on individual carnivore.

Table 3-17: MNP models accuracy for Leopard

MNP models accuracy for Leopard					
Model	**Max AUC**	**Min AUC**	**Mean AUC**	**SD AUC**	**Kappa**
SRE	0.694	0.433	0.569	0.075	0.340
CTA	0.722	0.420	0.533	0.094	0.341
RF	0.565	0.289	0.436	0.077	0.261
MARS	0.744	0.298	0.518	0.114	0.414
FDA	0.706	0.344	0.488	0.115	0.340
GLM	0.859	0.296	0.612	0.141	0.558
GBM	0.641	0.335	0.523	0.100	0.340
ANN	0.806	0.270	0.501	0.137	0.340

Table 3-18: MNP models accuracy for Clouded Leopard

MNP models accuracy for Clouded Leopard					
Model	**Max AUC**	**Min AUC**	**Mean AUC**	**SD AUC**	**Kappa**
SRE	0.833	0.483	0.555	0.103	0.784
CTA	0.733	0.350	0.523	0.108	0.267
RF	0.844	0.244	0.500	0.148	0.476
MARS	0.767	0.339	0.561	0.143	0.784
FDA	0.772	0.322	0.581	0.127	0.522
GLM	0.750	0.400	0.592	0.096	0.476
GBM	0.706	0.300	0.522	0.103	0.522
ANN	0.728	0.350	0.545	0.087	0.476

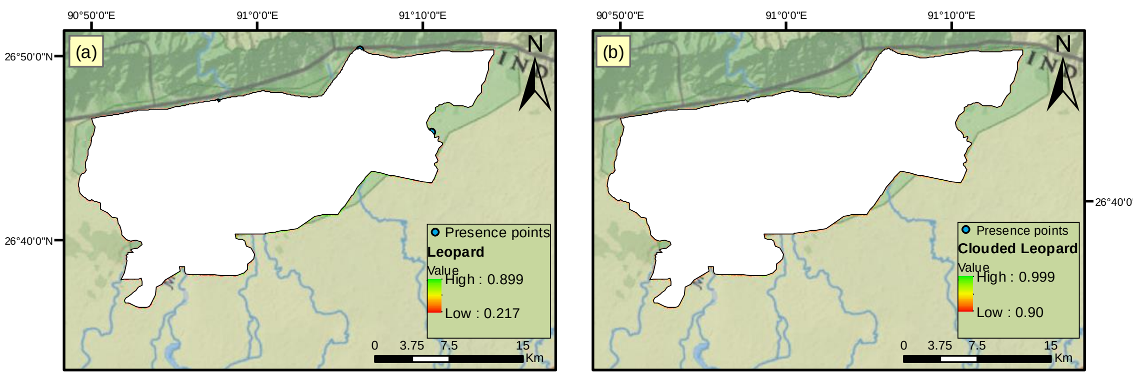


Figure 3-12: Distribution of leopard and clouded leopard

4. CONCLUSIONS AND RECOMMENDATION

4.1. CONCLUSION

SDM's have been useful in accurately defining the ecological niche of a species and help answer complex questions in ecology. In this study also SDMs have been successful in predicting the potential distribution of tigers in two areas (CTR and MNP). They have also helped to understand the difference in distribution patterns of tigers in both study areas and answer a few critical questions. In this chapter, first the research questions will be answered followed by providing specific conclusions and recommendations.

4.1 Answering the research questions

Q1: Do the anticipated explanatory variables (Distance from anthropogenic stress, Distance from water, Prey density, NDVI, altitude, slope) significantly contribute in the predictive model for distribution of tigers?

The explanatory variables (1) distance from anthropogenic stress (roads and settlements), (2) distance from water, (3) altitude, and (4) slope significantly contributed to the predictive model for distribution of tigers. The contribution of each variable is explained in detail in the results and discussion chapter. The variable prey density did not contribute significantly in explaining the distribution of tigers as hypothesized. NDVI contributed significantly in explaining the distribution of various prey species studied in this research.

Q2: Is the presence of anthropogenic stress a more important factor in explaining the distribution of tigers in CTR than MNP?

Yes, the presence of anthropogenic stress (settlements and agricultural areas) is a more important factor in explaining the distribution of tigers in CTR than MNP. The high levels of anthropogenic stress in CTR leads to more instances of human-tiger conflicts.

4.2 Specific conclusions

- Ensemble models performed better than any of the individual models.
- Apart from SRE all the models performed fairly well.
- Selection of pseudo-absences (how, where and how many) significantly determines the performance of the model.
- Models developed in CTR were accurately transferred to MNP and vice versa.

4.2. RECOMMENDATIONS

- Use of better sampling design for collection of species occurrence data.
- Incorporating expert knowledge on tiger ecology in order to improve the prey base.
- Studying tigers at territorial level not at the landscape level. Studying the territory of individual tiger and its relation to the prey base and other sympatric carnivore by employing telemetry devices will give better insight on its ecology.

LIST OF REFERENCES

Barbet‐Massin, M., Jiguet, F., Albert, C. H., & Thuiller, W. (2012). *Selecting pseudo‐ absences for species distribution models: how, where and how many?* Methods in Ecology and Evolution, 3(2), 327-338.

Bentley, A. (1978). *An introduction to the deer of Australia.* Koetong Trust Service, Victoria, Australia.

Borah J., Wangchuk D., Swargowari A., Wangchuk T., Sharma T., Das D., Rabha N., Basumatari A., Kakati N., Ahmed M. F., Sharma A., Sarmah A., Dutta D. K., Lahkar B., Dorji T., Brahma P. K. Ramchiary L., Tempa T., Wangdi Y., Nedup T., Wangdi T., Tharchen L., Dhendup P., Bhobora C. R., Pandav B. & Vattakaven J. (2012). *Tigers across Borders: Tigers in Indo-Bhutan Transboundary Manas Conservation Complex (TMCC).* Technical report.

Carbone, C., Mace, G.M., Roberts, S.C., & Macdonald, D.W. (1999). *Energetic constraints on the diet of terrestrial carnivores.* Nature. 402, 286-288.

Champion, F. W. (1929). *Tiger tracks.* Journal of the Bombay Natural History Society, *33*, 284-287.

Chanchani P., Lamichhane B. R., Malla S., Maurya K., Bista A., Warrier R., Nair S., Almeida M., Ravi R., Sharma R., Dhakal M., Yadav S. P., Thapa M., Jnawali S. R., Pradhan N. M. B., Subedi N., Thapa G. J., Yadav H., Jhala Y. V., Qureshi Q., Vattakaven J. and Borah J. (2014). *Tigers of the Transboundary Terai Arc Landscape: Status, distribution and movement in the Terai of India and Nepal.* National Tiger Conservation Authority, Government of India, and Department of National Park and Wildlife Conservation, Government of Nepal.

Chapron, G., Miquelle, D.G., Lambert, A., Goodrich, J.M., Legendre, S., & Clobert, J. (2008). *The impact on tigers of poaching versus prey depletion*. Journal of Applied Ecology 45, 1667-1674.

Choudhury, A. (2002). *Distribution and conservation of the Gaur Bos gaurus in the Indian Subcontinent.* Mammal Review, *32*(3), 199-226.

Chundawat, R.S., Habib, B., Karanth, U.K., Kawanishi, K., Ahmad Khan, J., Lynam, T., Miquelle, D., Nyhus, P., Sunarto, S., Tilson, R., & Wang, S. (2011). *Panthera tigris.* The IUCN Red List of Threatened Species. Version 2014.3. <www.iucnredlist.org>. Downloaded on **10 December 2014**.

Corbett, J. (1944). *Man-eaters of Kumaon*. Oxford University Press, London, UK.

Dinerstein, E., Loucks, C., Wikramanayake, E., Ginsberg, J., Sanderson, E., Seidensticker, J., Forrest, J., Bryja, G., Heydlauff, A., Klenzendorf, S., Leimgruber, P., Mills, J., O'Brien, T.G., Shrestha, M., Simons, R., & Songer, M. (2007). *The fate of wild tigers*. Bioscience, 57, 508-514.

Duckworth, J.W., Kumar, N.S., Anwarul Islam, Md., Hem Sagar Baral & Timmins, R.J. (2008). *Axis axis*. The IUCN Red List of Threatened Species. Version 2014.3. <www.iucnredlist.org>. Downloaded on **12 January 2015**.

Duckworth, J.W., Steinmetz, R., Timmins, R.J., Pattanavibool, A., Than Zaw, Do Tuoc & Hedges, S. (2008). *Bos gaurus*. The IUCN Red List of Threatened Species. Version 2014.3. <www.iucnredlist.org>. Downloaded on **12 January 2015**.

Dunbar-Brander, A. A. (1923). *Wild animals in central India*. Edward Arnold, London, UK

Elith, J., Graham, C. H., Anderson, R. P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R. J., Huettmann, F., Leathwick, J. R., Lehmann, A., Li, J., Lohmann, L. G., Loiselle, B. A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J. McC., Peterson, A. T., Phillips, S. J., Richardson, K. S., Scachetti-Pereira, R., Schapire, R. E., Soberon, J., Williams, S., Wisz, M. S., & Zimmermann, N. E. (2006). *Novel methods improve prediction of species' distributions from occurrence data*. Ecography, 29(2), 129-151.

Elith, J., & Leathwick, J. R. (2009). *Species distribution models: ecological explanation and prediction across space and time*. Annual Review of Ecology, Evolution, and Systematics, 40(1), 677.

Ghosh, S. (2013). *Habitat Suitability analysis for Tigers (Panthera tigris) in the Indo-Bhutan Manas Tiger Conservation Landscape*. (Ph.D. Thesis). Department of geography and earth Sciences, Aberystwyth University, U.K.

Grinnell J. (1904). *The origin and distribution of the chestnut-backed chickadee*. Auk, 21:364–65

Guisan, A., & Thuiller, W. (2005). *Predicting species distribution: offering more than simple habitat models*. Ecology letters, 8(9), 993-1009.

Guisan, A., & Zimmermann, N. E. (2000). *Predictive habitat distribution models in ecology. Ecological modelling*, 135(2), 147-186.

Harihar, A., Pandav, B., & Goyal, S.P. (2009). *Subsampling photographic capture-recapture data of tigers (Panthera tigris) to minimize closure violation and improve estimate precision: a case study*. Population Ecology, 51, 471-479.

Hedges, S., Sagar Baral, H., Timmins, R.J. & Duckworth, J.W. (2008). *Bubalus arnee*. The IUCN Red List of Threatened Species. Version 2014.3. <www.iucnredlist.org>. Downloaded on **12 January 2015**.

Henschel, P., Hunter, L., Breitenmoser, U., Purchase, N., Packer, C., Khorozyan, I., Bauer, H., Marker, L., Sogbohossou, E. & Breitenmoser-Wursten, C. (2008). *Panthera pardus*. The IUCN Red List of Threatened Species. Version 2014.3. <www.iucnredlist.org>. Downloaded on **12 January 2015**.

Imam, E., Kushwaha, S.P.S., & Singh, A. (2009*). Evaluation of suitable tiger habitat in Chandoli National Park, India, using multiple logistic regression*. Ecological Modelling, 220, 3621-3629.

Jhala, Y.V., Qureshi, Q., Gopal, R., & Sinha, P.R. (2011). *Status of the Tigers, Co-predators and Prey in India*. Dehradun, National Tiger Conservation Authority, Government of India, New Delhi and Wildlife Institute of India.

Karanth, K.U. (2003). *Tiger ecology and conservation in the Indian subcontinent*. Journal of the Bombay Natural History Society, 100(2 & 3), 169-189.

Karanth, K.U., & Chellam, R. (2009). *Carnivore conservation at the crossroads*. Oryx, 43, 1-2.

Karanth, K.U., Chundawat, R.S., Nichol, J.D., & Kumar, N.S. (2004). *Estimation of tiger densities in the tropical dry forests of Panna, Central India, using photographic capture-recapture sampling*. Animal Conservation, 7, 285-290.

Karanth, K.U., Goodrich, J.M., Vaidyanathan, S., & Reddy, G.V. (2009). *Landscape-scale, ecology-based management of wild tiger populations*. Paper presented to GTI Workshop, Kathmandu.

Karanth, K.U., Nichols, J.D., Kumar, N.S., & Hines, J.E. (2006). *Assessing tiger population dynamics using photographic capture-recapture sampling*. Ecology, 87, 2925-2937.

Karanth, K.U., & Sunquist, M.E. (1995). *Prey selection by tiger, leopard and dhole in tropical forests*. Journal of Animal Ecology, 64, 439–450.

Kushwaha, S. P. S., Khan, A., Habib, B., Quadri, A., & Singh, A. (2004). *Evaluation of sambar and muntjak habitats using geostatistical modelling*. Current Science Bangalore, 86(10), 1390-1400.

Kushwaha, S., Quadri, A., Singh, A., & Khan, A. (2008). *Habitat Preference Assessment of the Tiger and its Prey Ungulates in the Corbett Tiger Reserve*. Fauna of Corbett Tiger Reserve, Uttarakhand, 35, 33.

Mallon, D.P. (2008). *Boselaphus tragocamelus*. The IUCN Red List of Threatened Species. Version 2014.3. <www.iucnredlist.org>. Downloaded on **12 January 2015**.

Niamir, A. (2009). *Distribution modelling of the Short-toed Eagle in relation to potential food availability*. (Master's thesis). ITC, University of Twente, Enschede, Netherlands. Retrieved from: http://www.itc.nl/library/papers_2009/msc/gem/niamir.pdf.

Oliver, W. & Leus, K. (2008). *Sus scrofa*. The IUCN Red List of Threatened Species. Version 2014.3. <www.iucnredlist.org>. Downloaded on **12 January 2015**.

Pearson, R. G., & Dawson, T. P. (2003). *Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful?* Global ecology and biogeography, 12(5), 361-371.

Sanderson, E., Forrest, J., Loucks, C., Ginsberg, J., Dinerstein, E., Seidensticker, J., Leimgruber, P., Songer, M., Heydlauff, A., O'Brien, T., Bryja, G., Klenzendorf, S., & Wikramanayake, E. (2006). *Setting Priorities for the Conservation and Recovery of Wild Tigers: 2005-2015*. The Technical Assessment. WCS, WWF, Smithsonian, and NFWF-STF, New York and Washington, DC, USA.

Sanderson, J., Khan, J.A., Grassman, L. & Mallon, D.P. (2008). *Neofelis nebulosa*. The IUCN Red List of Threatened Species. Version 2014.3. <www.iucnredlist.org>. Downloaded on **12 January 2015**

Sankar, K., & Goyal, S. P. (2004). *Ungulates of India. ENVIS Bulletin: Wildlife.*

Schaller, G. B. (1967). *The deer and the tiger*. University of Chicago Press.

Singh, A., & Kushwaha, S. P. S. (2011). *Refining logistic regression models for wildlife habitat suitability modeling—A case study with muntjak and goral in the Central Himalayas, India*. Ecological Modelling, 222(8), 1354-1366.

Skidmore, A. K., Toxopeus, A. G., Bie, K. D., Corsi, F., Venus, V., Omolo, D. P., Marquez, J., & Giménez, R. R. (2006). *Herpetological species mapping for the Meditteranean*. In ISPRS Commission VII Mid Term Symposium on Remote Sensing from Pixel to Processes, Enschede.

Sunquist, M. E., Karanth, K. U., & Sunquist, F. (1999). *Ecology, behaviour and resilience of the tiger and its conservation needs. Riding the tiger: tiger conservation in human-dominated landscapes*. Cambridge University Press, Cambridge, UK.

Thapar, V. (1989). *Tigers: the secret life*. Elm Tree.

Thuiller, W., Lafourcade, B., Engler, R. & Araujo, M. B. 2009. *BIOMOD - a platform for ensemble forecasting of species distributions*. Ecography 32: 369-373 (Version 3.1-64).

Timmins, R.J., Duckworth, J.W., Pattanavibool, A., Steinmetz, R., Samba Kumar, N., Anwarul Islam, Md. & Sagar Baral, H. (2008a). *Muntiacus vaginalis*. The IUCN Red List of Threatened Species. Version 2014.3. <www.iucnredlist.org>. Downloaded on **12 January 2015**.

Timmins, R., Duckworth , J.W., Samba Kumar, N., Anwarul Islam, M., Sagar Baral, H., Long, B. & Maxwell, A. (2012). *Axis porcinus*. The IUCN Red List of Threatened Species. Version 2014.3. <www.iucnredlist.org>. Downloaded on **12 January 2015**.

Timmins, R.J., Steinmetz, R., Sagar Baral, H., Samba Kumar, N., Duckworth, J.W., Anwarul Islam, Md., Giman, B., Hedges, S., Lynam, A.J., Fellowes, J., Chan, B.P.L. & Evans, T. (2008b). *Rusa unicolor*. The IUCN Red List of Threatened Species. Version 2014.3. <www.iucnredlist.org>. Downloaded on **12 January 2015**.

5. APPENDICES

APPENDIX 1

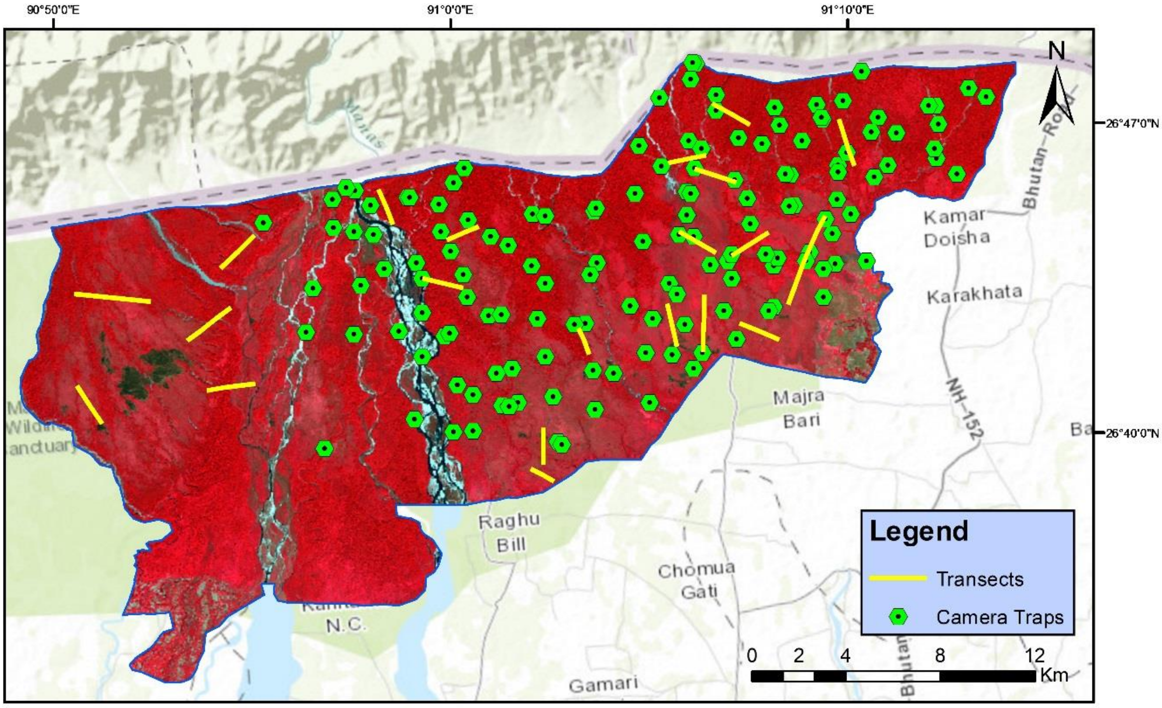


Figure 5-1: Transects and camera traps laid across MNP

APPENDIX 2

Table 5-1: LULC MNP

LULC type	Barking Deer	Gaur	Buffalo	Sambar	Wild Boar	Hog Deer
Cropland	0	0	0	0	1	0
Deciduous	1	1	1	1	1	1
Fallow	0	0	0	0	0	0
Grassland	0	1	1	0	1	1
Inland wetlands	0	0	1	0	0	0
Agricultural Plantation	0	1	1	0	1	0
Rivers	0	0	0	0	0	0
Rural	0	0	0	0	0	0
Sandy Areas	0	1	1	0	1	0
Scrub land	0	1	1	0	1	1
Semi evergreen	1	1	1	1	1	1

Table 5-2: LULC CTR

LULC type	Barking Deer	Chital	Nilgai	Sambar	Wild Boar
Cropland	0	0	1	0	1
Deciduous	1	1	0	1	1
Fallow	0	0	1	0	0
Forest Plantation	1	1	1	0	1
Grasslands	0	1	1	0	1
Mining	0	0	0	0	0
Agricultural Plantation	0	0	1	0	1
Reservoir / Lakes	0	0	0	0	0
Rivers	0	0	0	0	0
Rural	0	0	0	0	0
Scrub land	0	0	1	0	1
Urban	0	0	0	0	0

Table 5-3: LULC COPREDATORS

LULC type	Tiger	Leopard	Clouded Leopard
Cropland	0	0	0
Deciduous	1	1	1
Fallow	0	0	0
Forest Plantation	0	1	0
Grassland	1	1	0
Inland wetlands	1	0	0
Agricultural Plantation	0	1	0
Rivers	0	0	0
Reservoir / Lakes	0	0	0
Rural	0	0	0
Sandy Areas	0	0	0
Scrub land	0	0	0
Semi evergreen	1	1	1
Urban	0	0	0

APPENDIX 3

Table 5-4: MNP

Variables	VIF
lulc	1.797297
ndvi	1.640376
veg_density	1.583691
elevation	3.690009
slope	1.154406
southness	1.124934
westness	1.009262
distance_waterbodies	1.222549
distance_road	1.162126
distance_settlement	3.490421
combine	1.77309

Table 5-5: CTR

Variables	VIF
lulc	2.503435
ndvi	2.478668
veg_density	1.879534
elevation	1.856513
slope	1.68566
southness	1.23198
westness	1.072314
distance_waterbodies	1.213012
distance_road	1.596651
distance_settlement	1.503258
combine	2.460771

APPENDIX-4

A

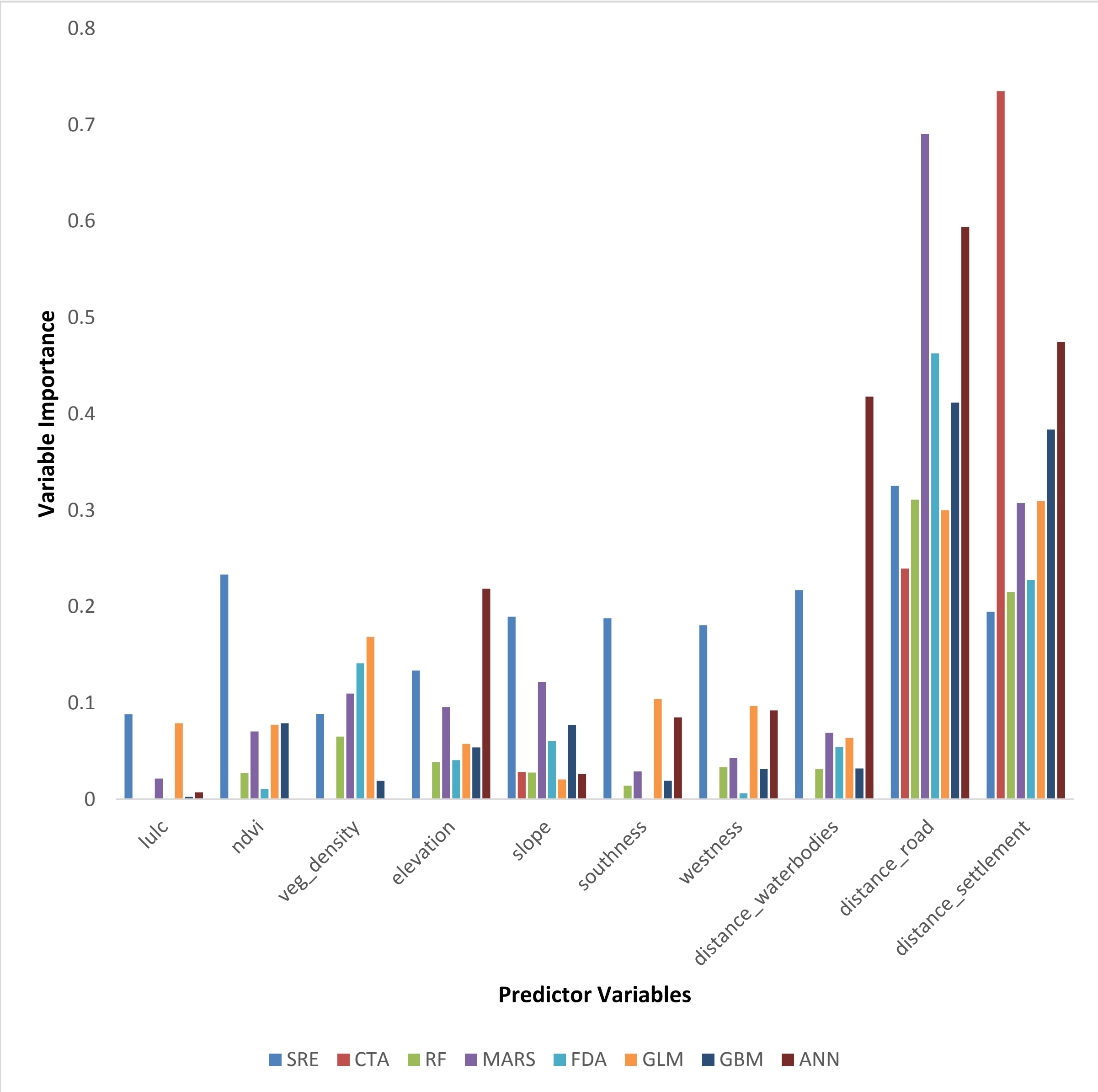


Figure 5-2: CTR variables importance for barking Deer

B

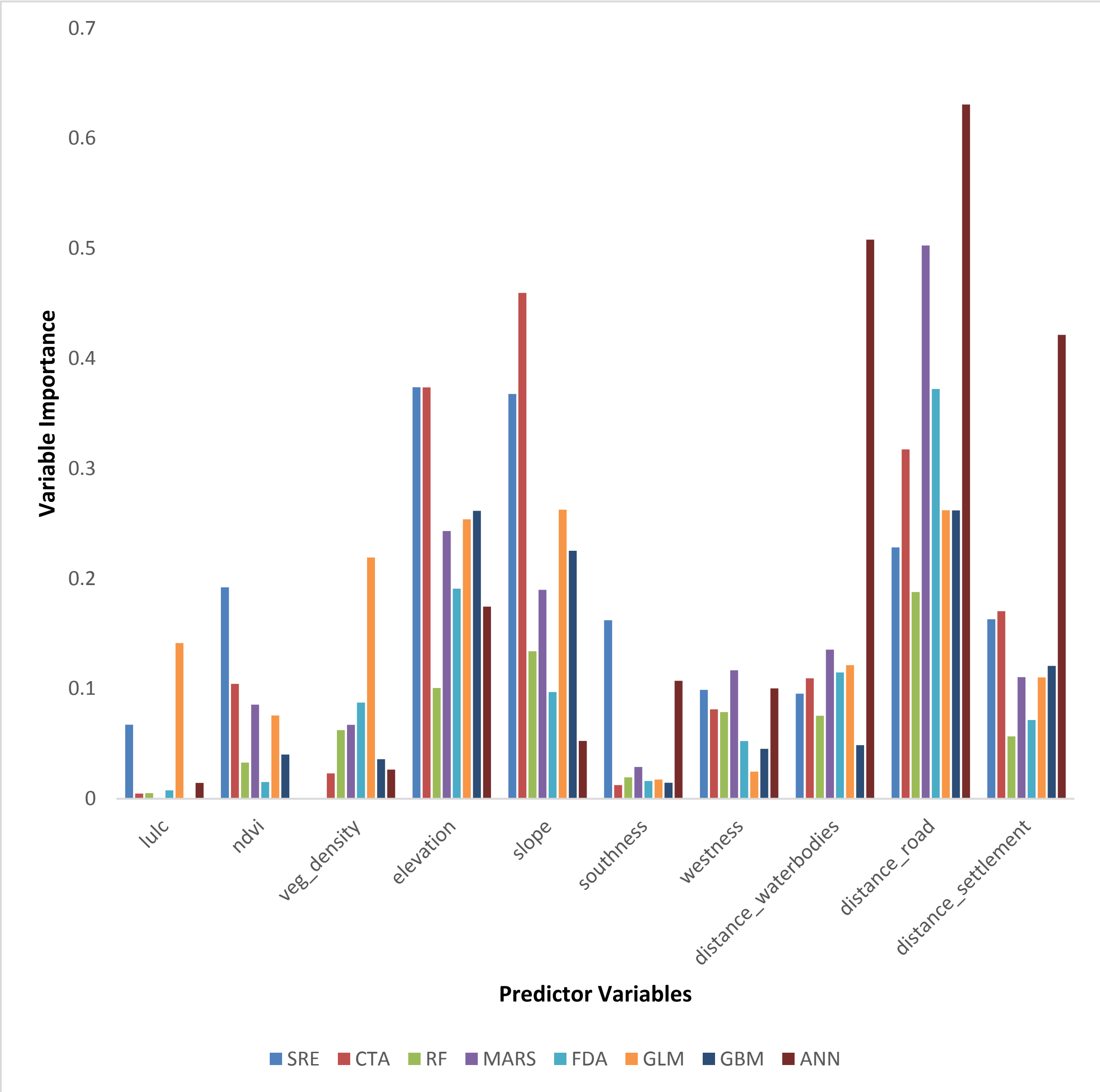


Figure 5-3: CTR variables importance for Chital

C

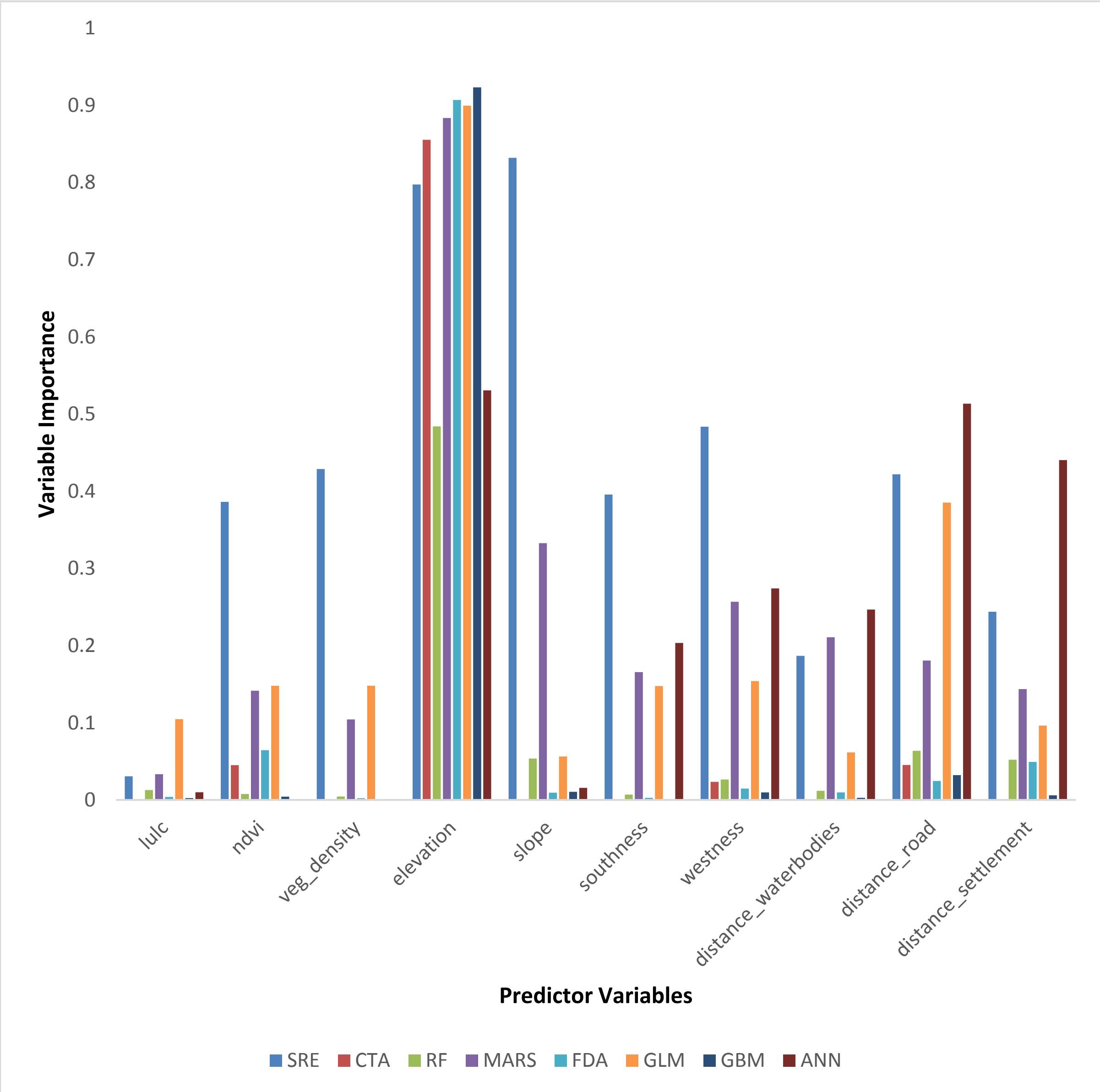


Figure 5-4: CTR variables importance for Nilgai

D

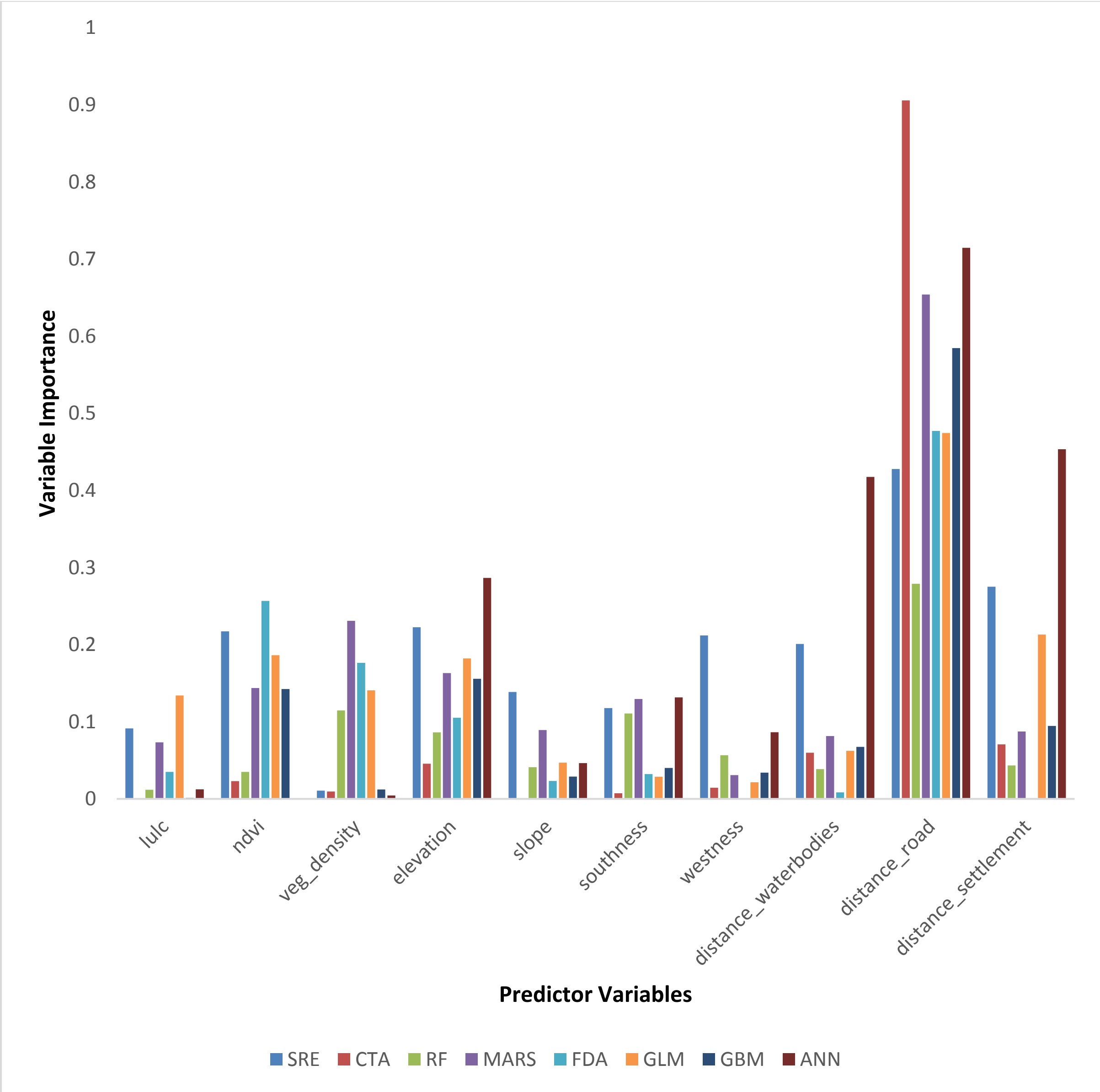


Figure 5-5: CTR variables importance for Sambar

E

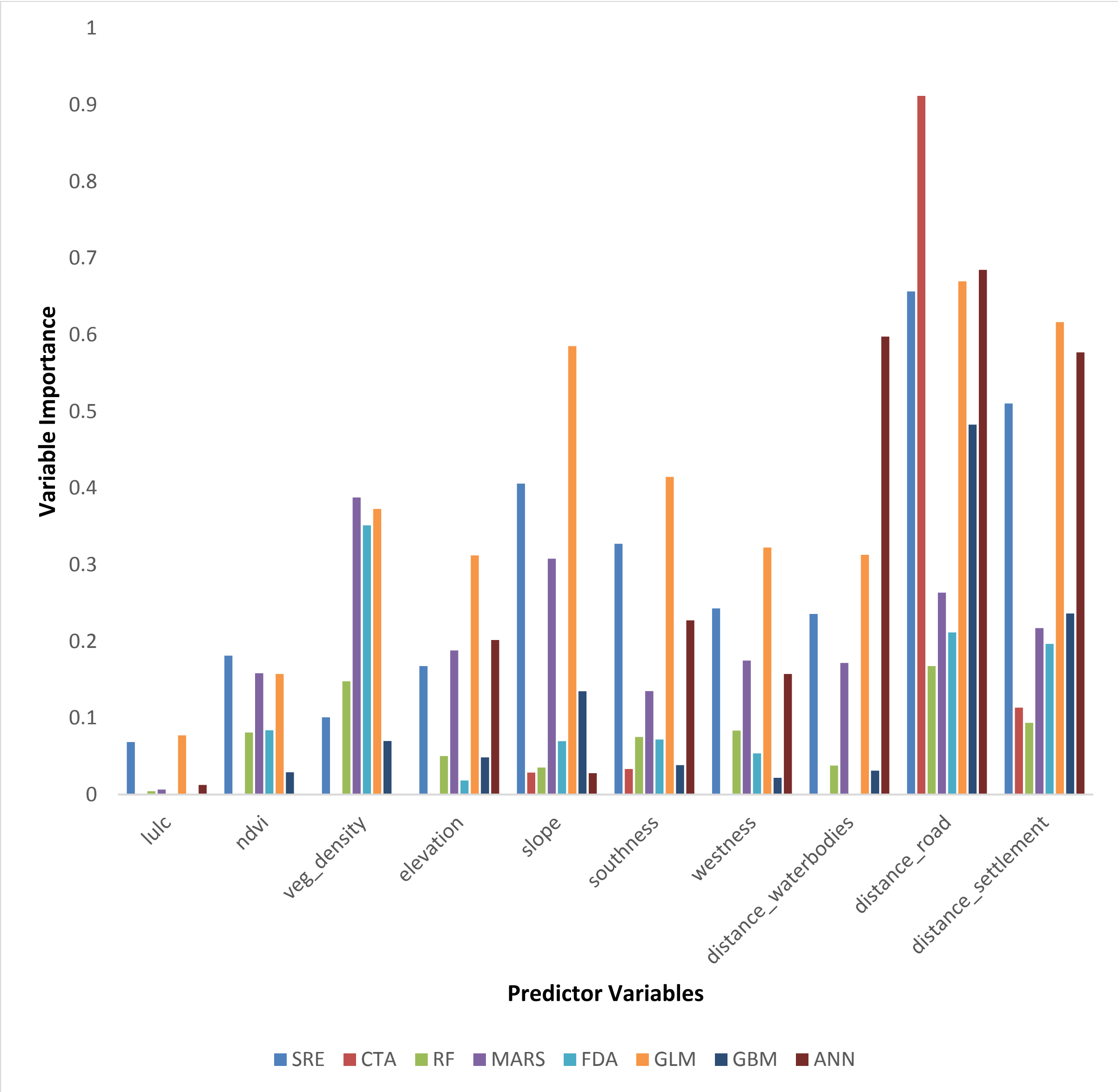


Figure 5-6: CTR variables importance for Wild Boar

F

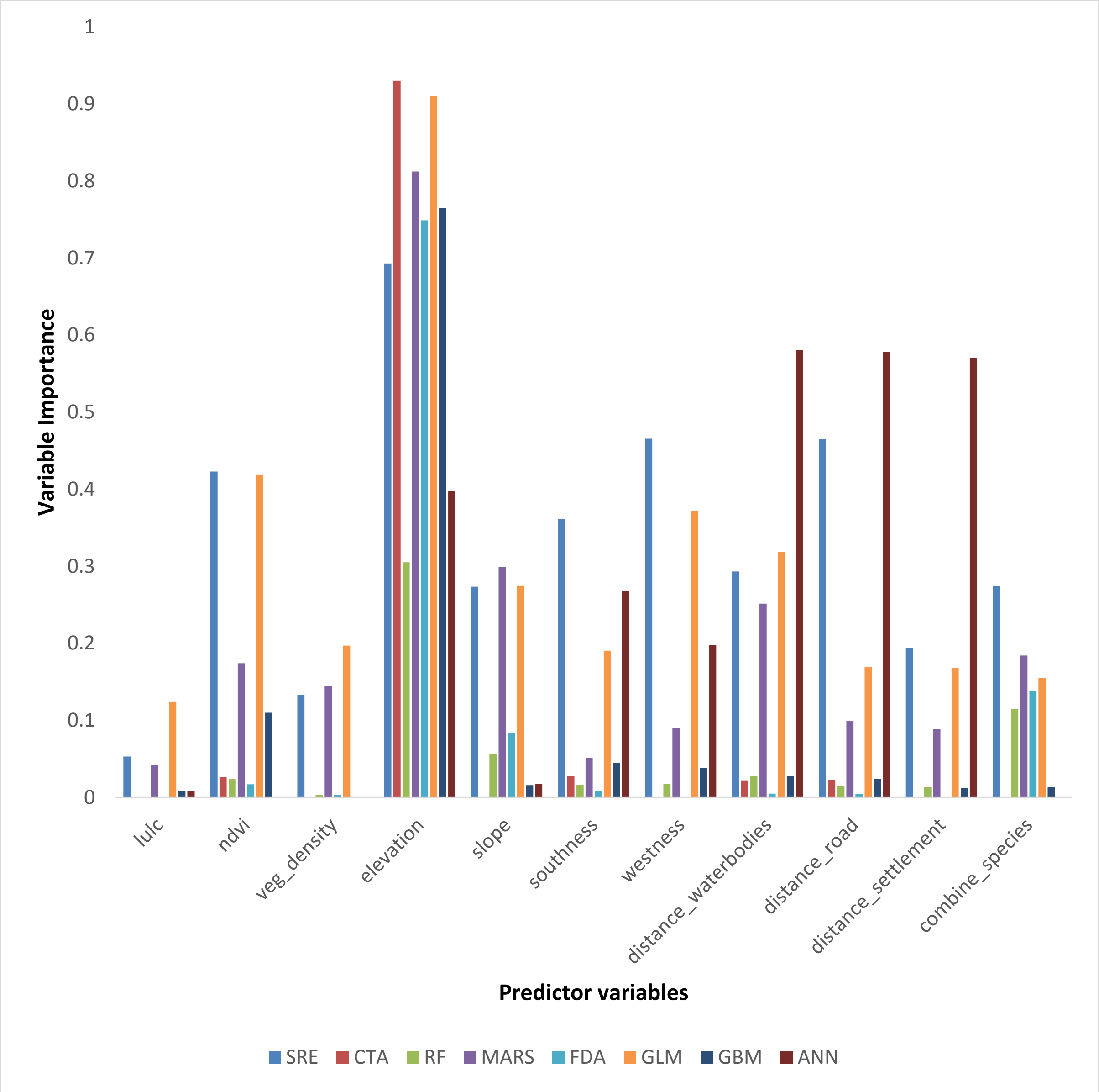


Figure 5-7: CTR variables importance for Leopard

G

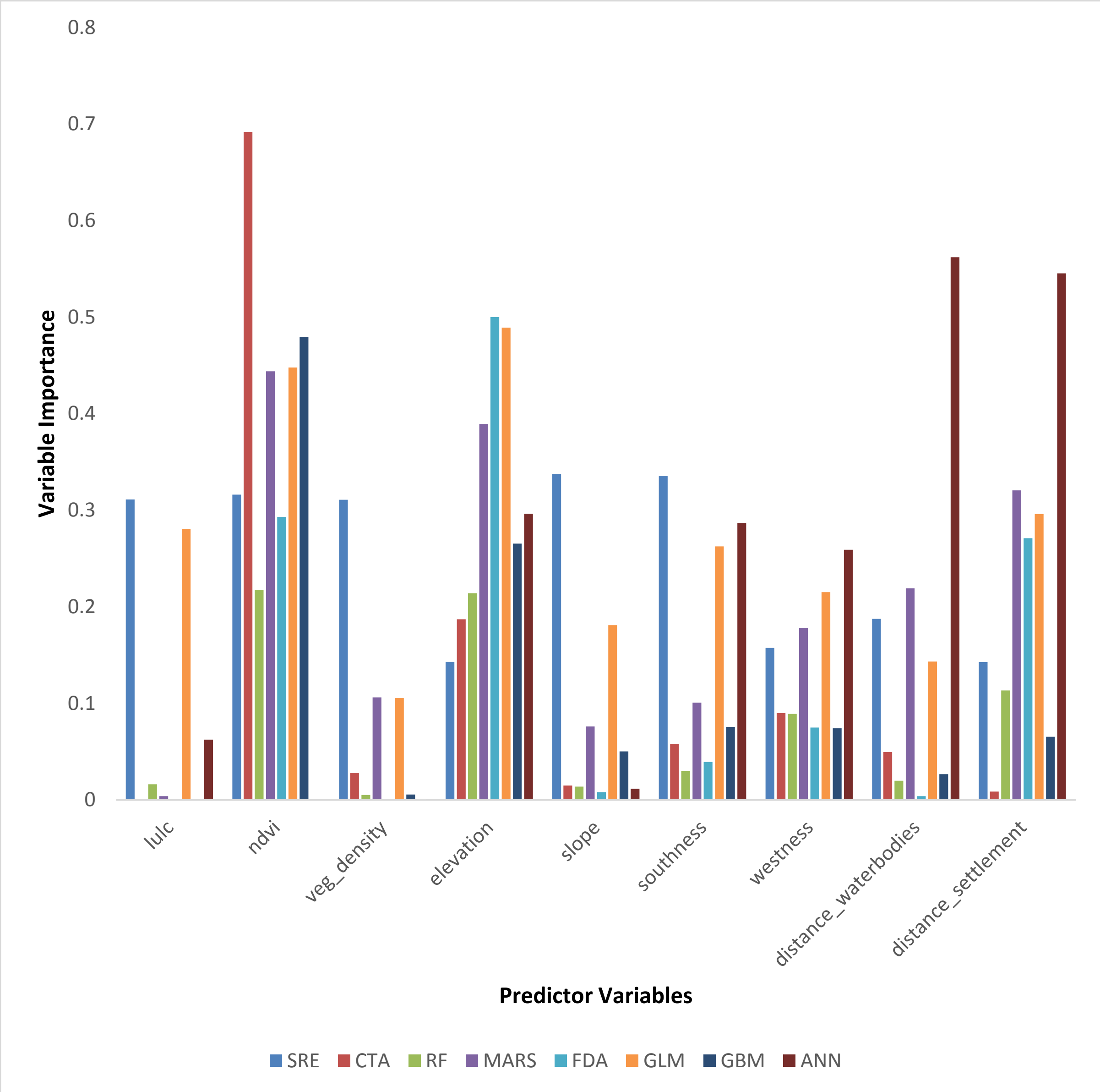


Figure 5-8: MNP variables importance for Barking Deer

H

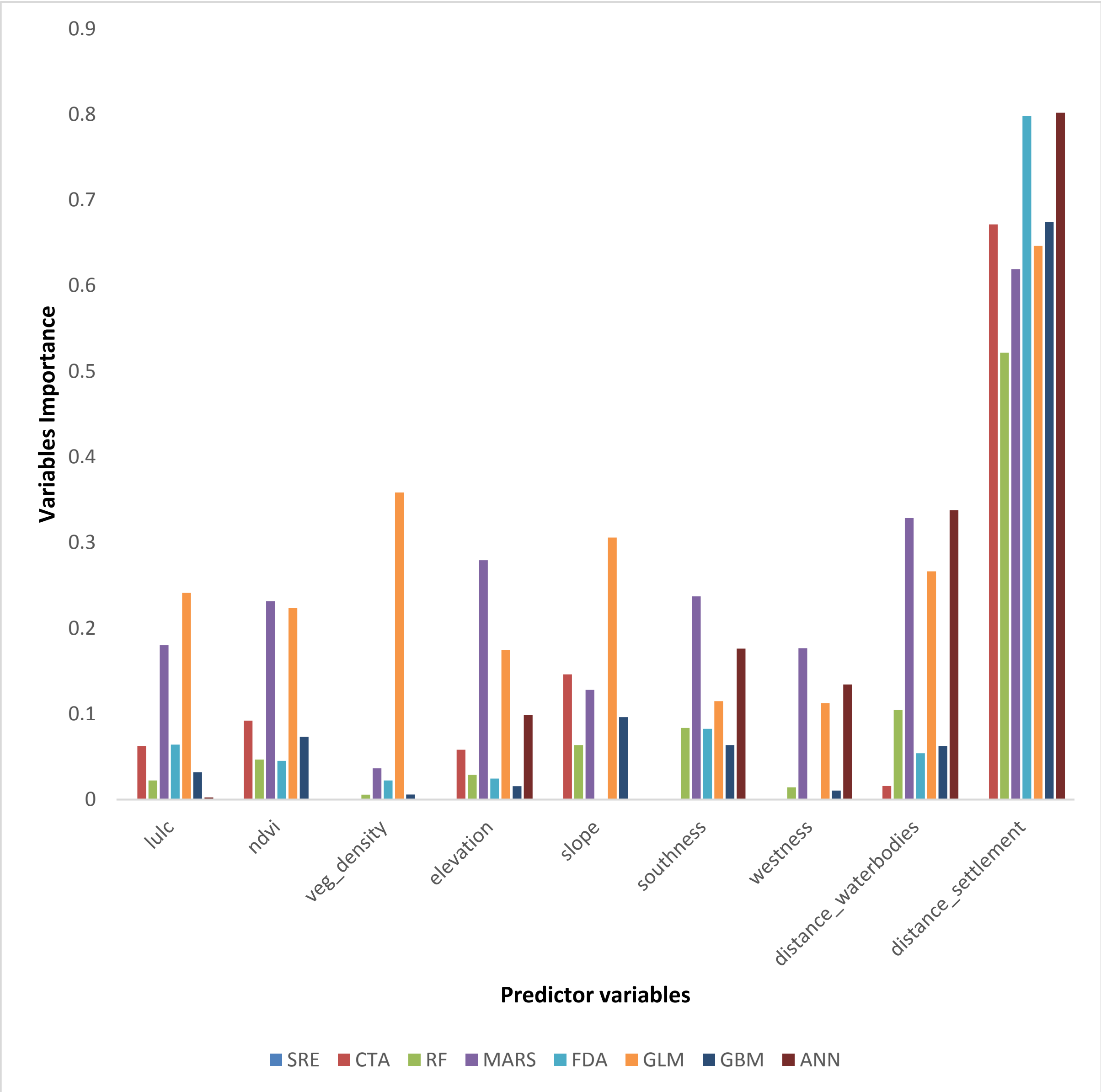


Figure 5-9: MNP variables importance for Buffalo

I

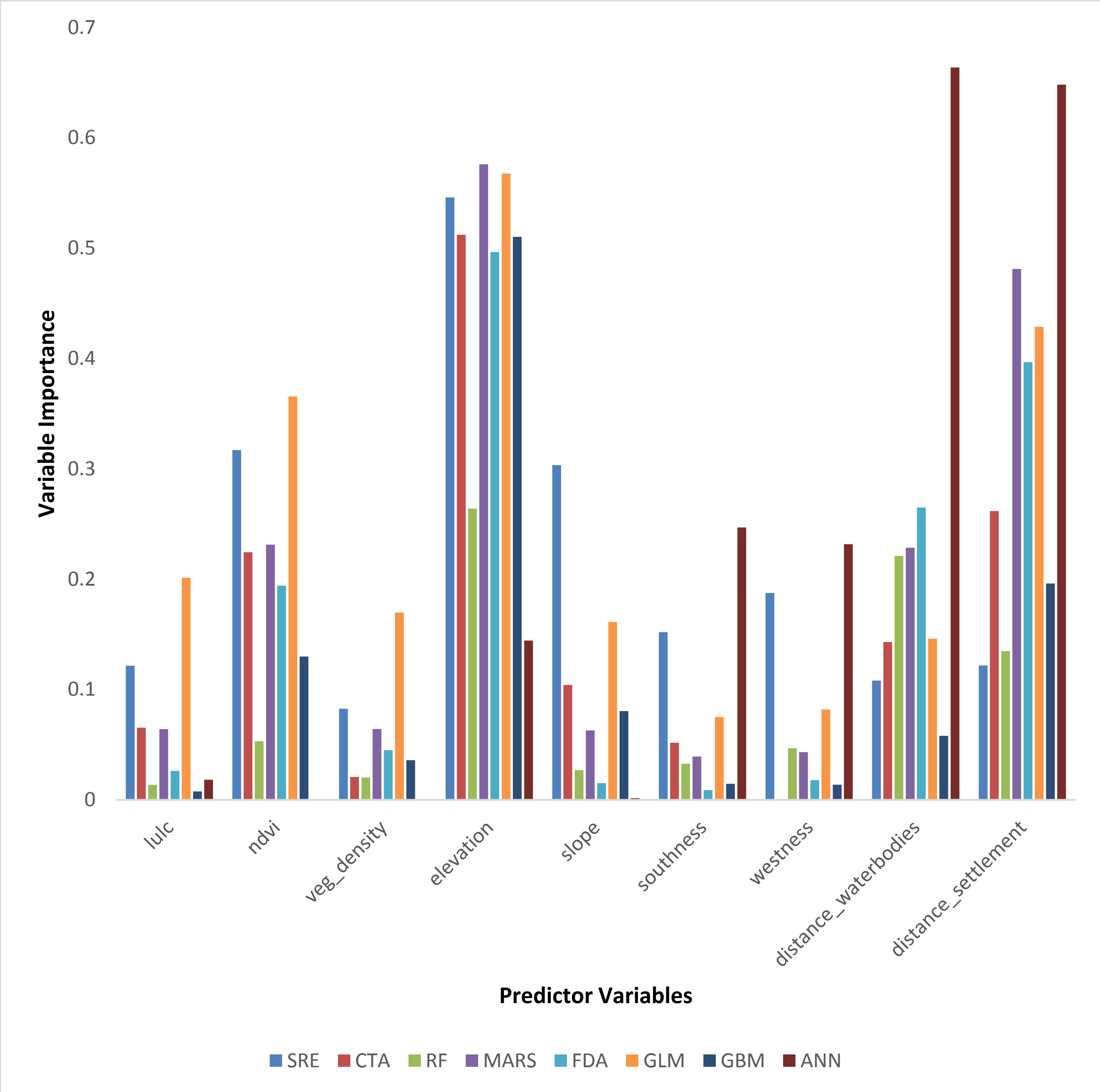


Figure 5-10: MNP variables importance for Gaur

J

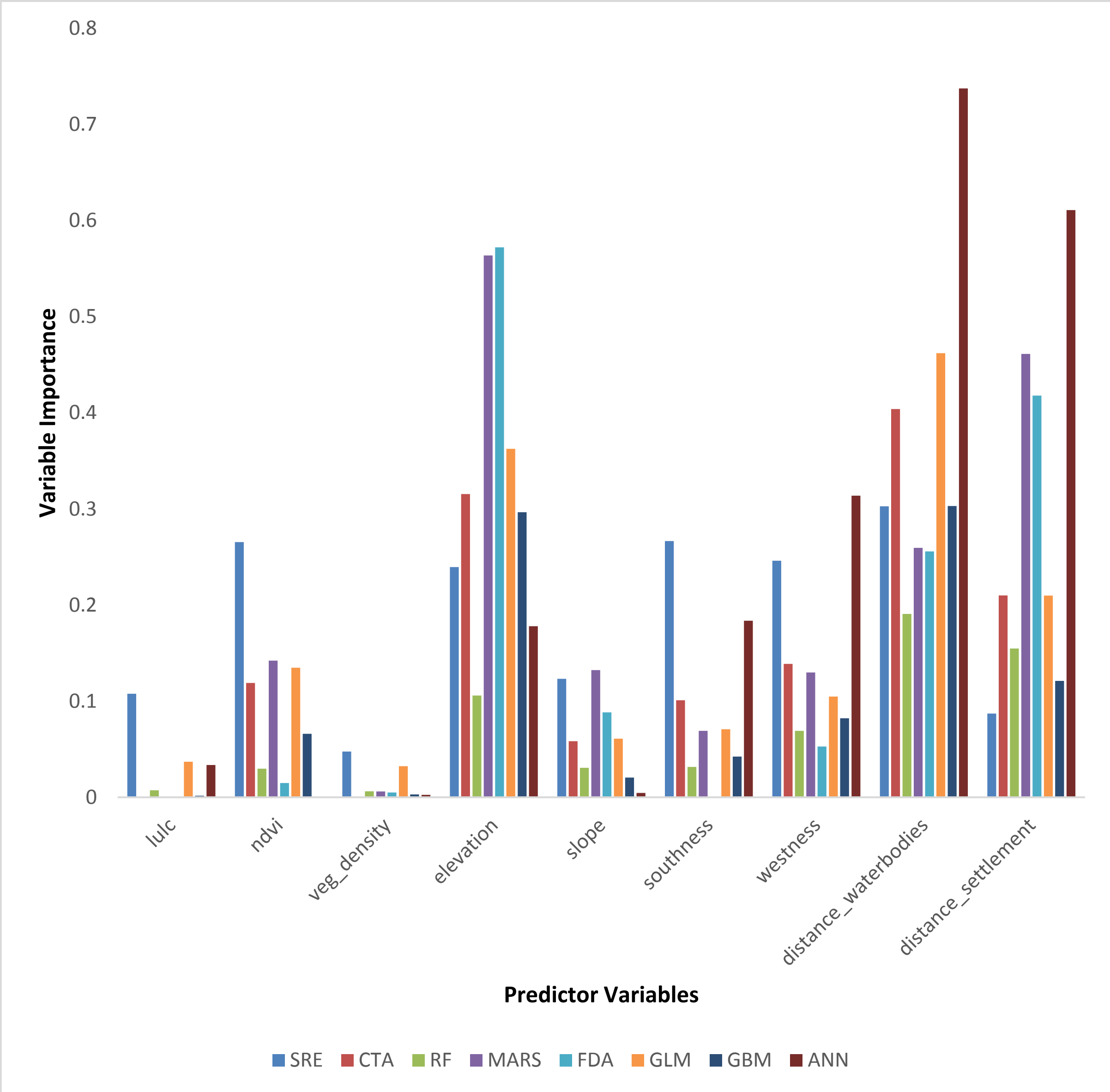


Figure 5-11: MNP variables importance for Hog Deer

K

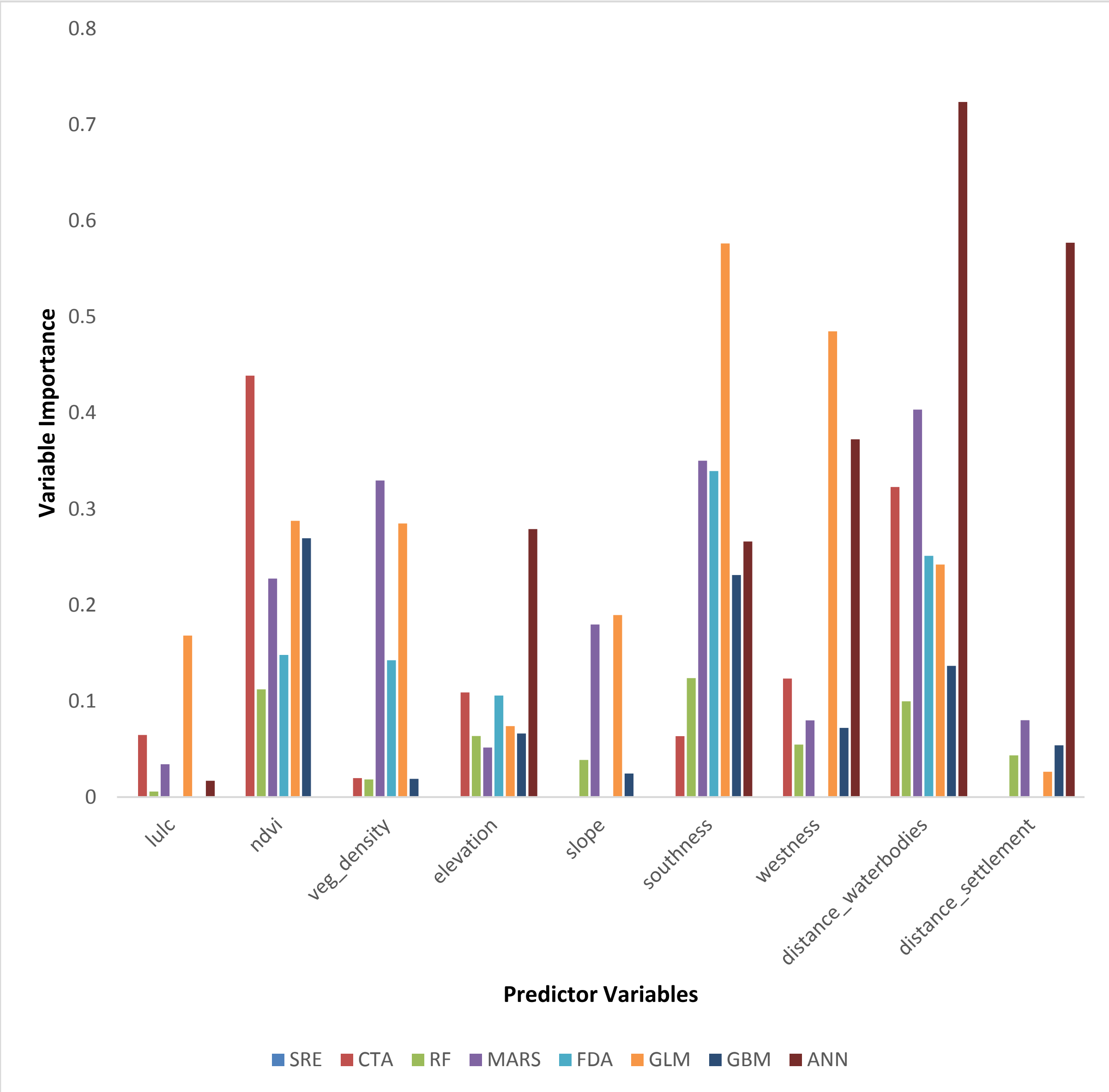


Figure 5-12: MNP variables importance for Sambar

L

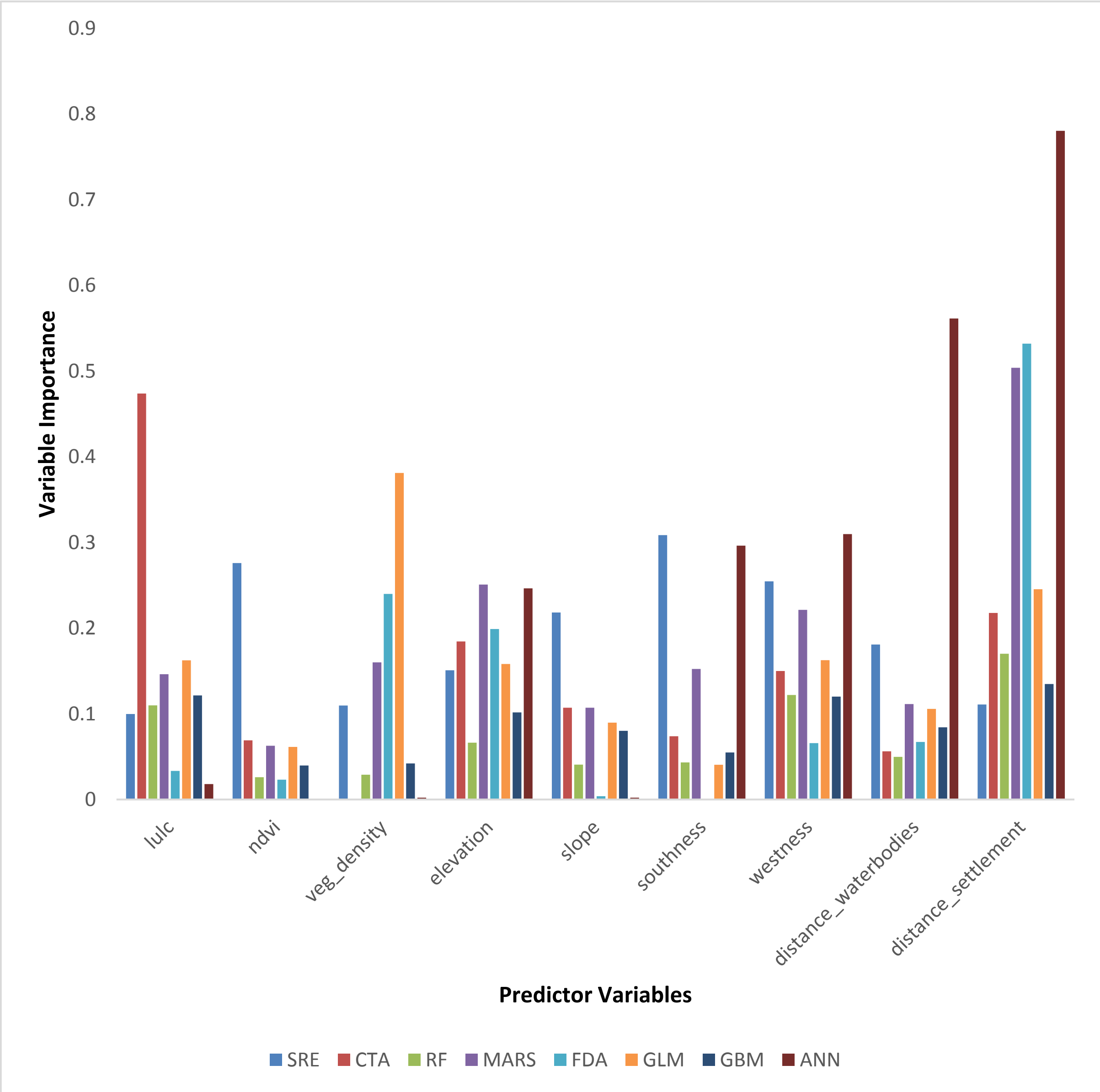


Figure 5-13: MNP variables importance for Wild Boar

M

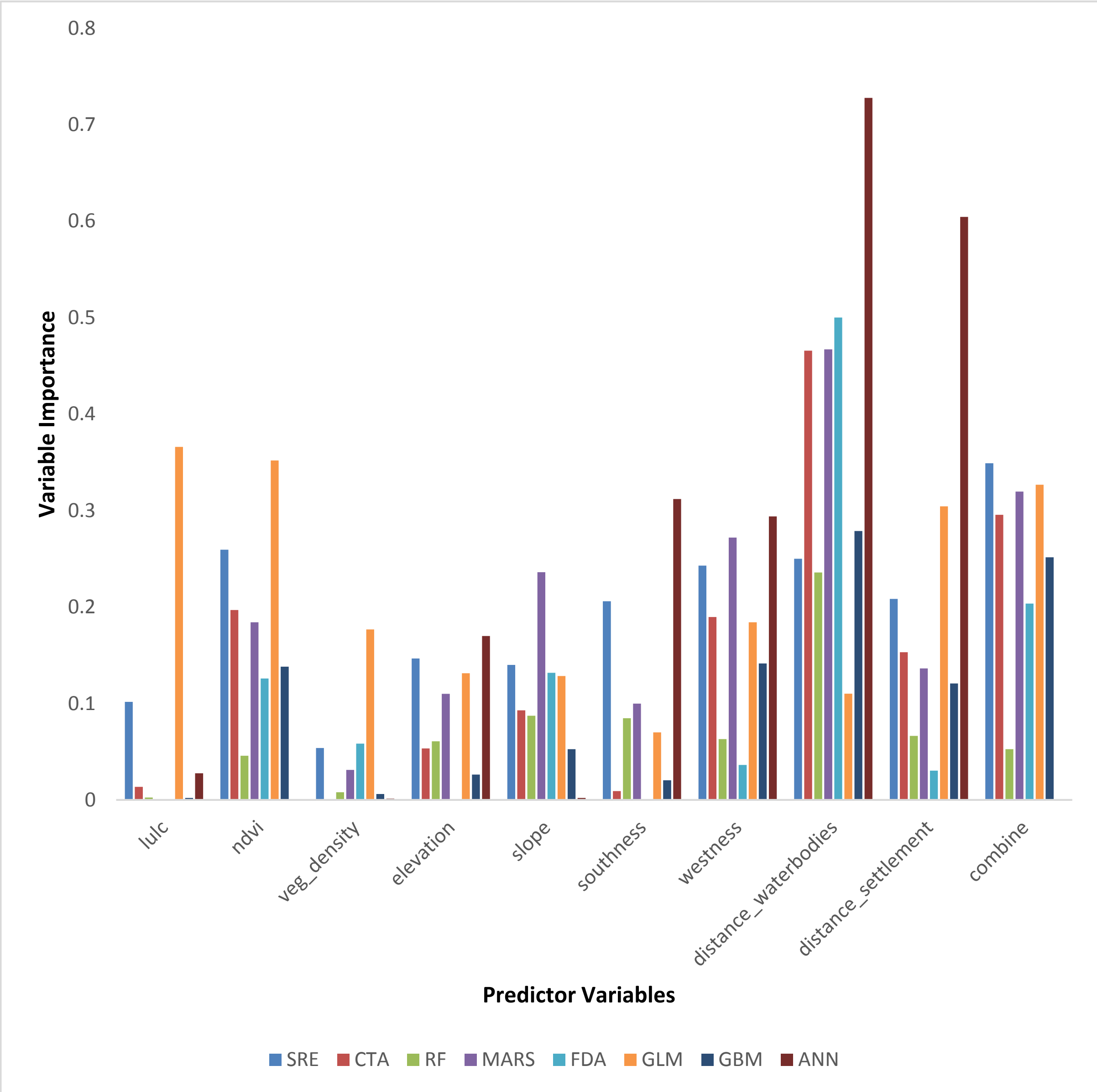


Figure 5-14: MNP variables importance for Leopard

N

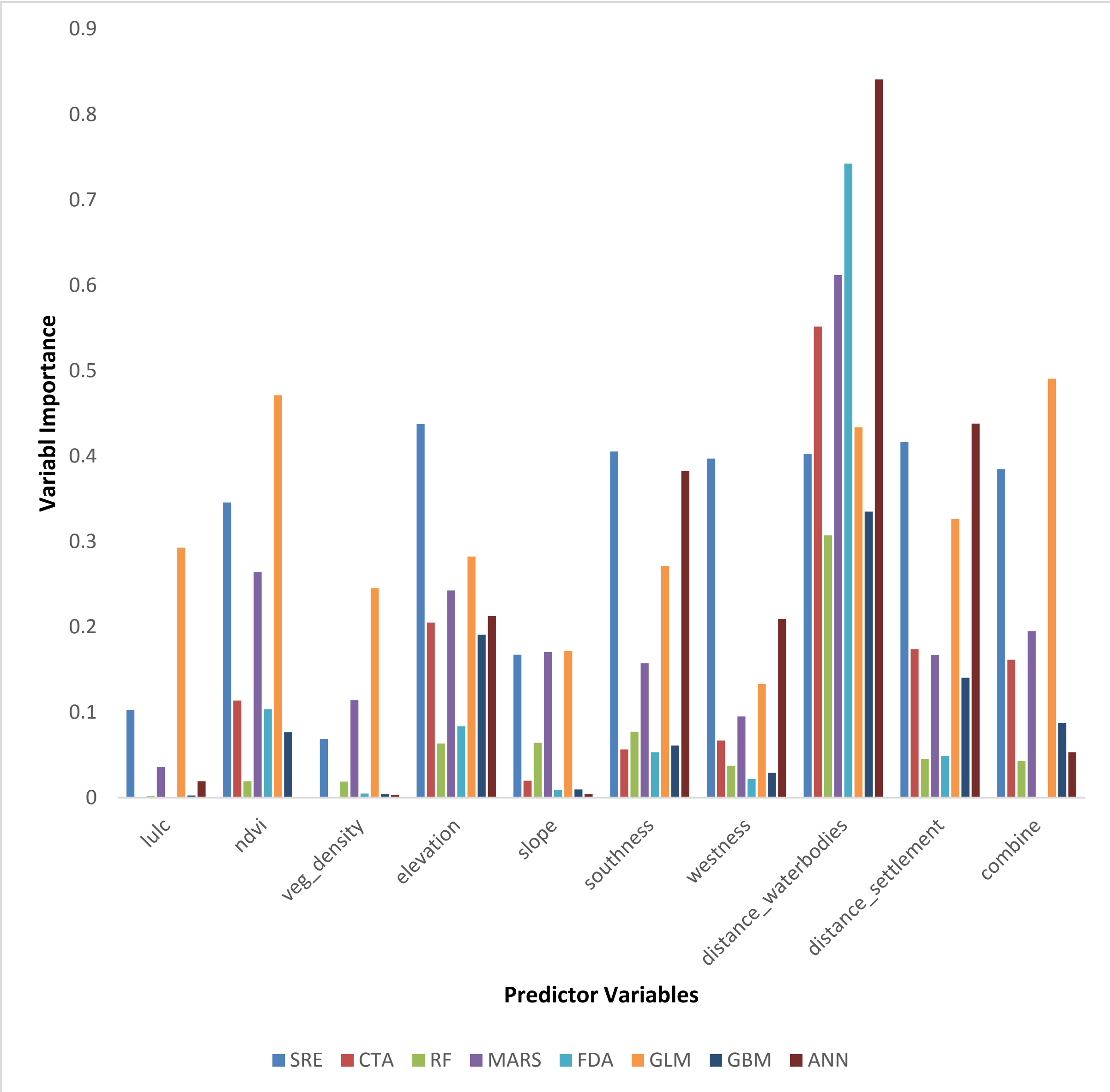


Figure 5-15: MNP variables importance for Clouded Leopard