USING MACHINE LEARNING TO PREDICT THE RISK OF HUMAN-ELEPHANT CONFLICT IN THE NEPAL-INDIA TRANSBOUNDARY REGION

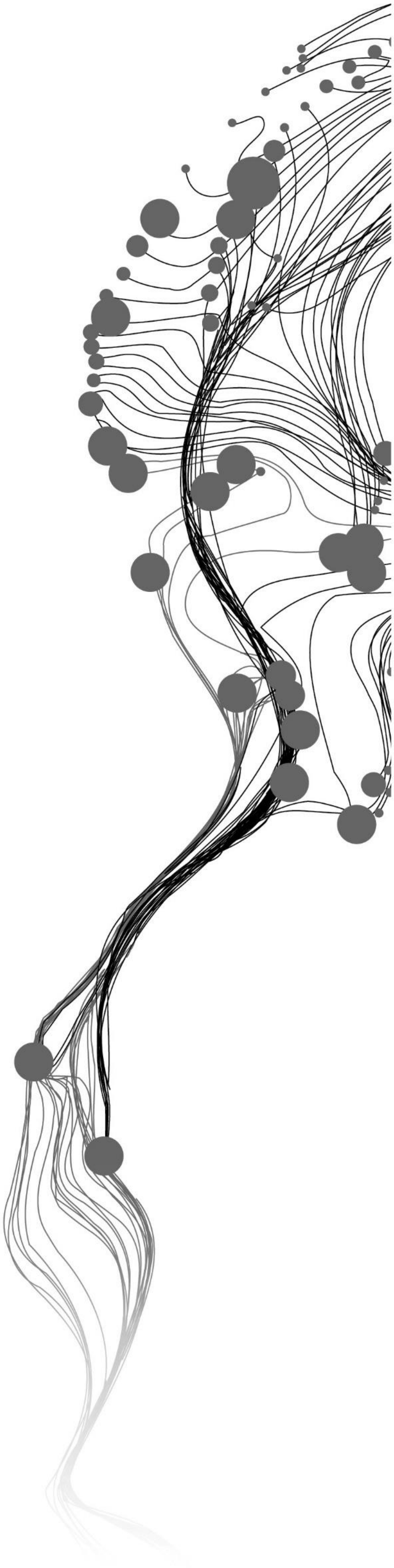
BINITA KHANAL

June 2022

SUPERVISORS:

Assoc. Prof. Dr. Tiejun Wang

Drs. Raymond G. Nijmeijer



USING MACHINE LEARNING TO PREDICT THE RISK OF HUMAN-ELEPHANT CONFLICT IN THE NEPAL-INDIA TRANSBOUNDARY REGION

BINITA KHANAL

Enschede, The Netherlands, June 2022

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.

Specialisation: Natural Resources Management

SUPERVISORS:
Assoc. Prof. Dr. Tiejun Wang
Drs. Raymond G. Nijmeijer

ADVISOR:
Dr. Ashok Kumar Ram (Department of National Parks and Wildlife Conservation, Nepal)

THESIS ASSESSMENT BOARD:
Assoc. Prof. Dr. Thomas A. Groen (Chair, University of Twente)
Dr. Shadrack Ngene (External Examiner, Wildlife Research and Training Institute, Kenya)
Assoc. Prof. Dr. Tiejun Wang (First supervisor, University of Twente)
Drs. Raymond G. Nijmeijer (Second supervisor, University of Twente)

ABSTRACT

Human-elephant conflict (HEC) is a common form of human-wildlife conflict in African and Asian countries where wild elephants are present. The cross-border regions between Nepal and India are the natural habitats for Asian elephants. However, this region has experienced dramatic land cover and land use changes due to human pressure and infrastructure development over the last several decades. Habitat loss and fragmentation drive elephants closer to human settlements and cause more frequent human-elephant conflicts. The study of this phenomenon does not just concern the conservation of wildlife but also human security. This study aims to predict the risk of human-elephant conflict along the transboundary landscape of Nepal-India using machine learning algorithms. To do so, I first modelled the habitat suitability of elephants using an ensemble species distribution modelling approach and identified key factors determining habitat suitability. Then I predicted the risk of human-elephant conflict using a random forest algorithm and identified major factors contributing to the risk. The results of my study show that 26,679 km^2, approximately one-third of the total transboundary landscape area, is predicted to be suitable for Asian elephants. Only twenty per cent of the predicted suitable habitat is located within the protected areas. Elevation, precipitation of the driest month and wettest month, and temperature of the warmest month are the key variables determining the habitat suitability for elephants in this region. The result of the predicted human-elephant conflict indicated high human interference in the remaining suitable habitats of Asian elephants. Human settlements and agricultural fields near protected areas experienced a high risk of conflict. The human disturbances and the expansion of settlements in the migratory route of elephants are expected to intensify human-elephant conflict. This is the first study that attempts to use state-of-the-art machine learning algorithms to predict the risk of human-elephant conflict along the cross-border landscape of Nepal-India. The suitable elephant habitat and the human-elephant risk areas identified by this study are important, which could serve as a basis for developing transboundary conservation of elephants as well as strategies for mitigating man-elephant conflicts. The study recommends that the transboundary conservation efforts need to be strengthened, and special attention should be paid to human colonisation around the protected area while implementing measures to mitigate the risks of conflict between humans and elephants.

Keywords: Asian elephant, species distribution modelling, ensemble model, habitat suitability, human-wildlife conflict, cross border

ACKNOWLEDGEMENTS

The journey of two years from *I am missing my home (Nepal)* to *I will miss my second home (Enschede, ITC)* has been surreal. I joined ITC with the hope of learning about GIS, remote sensing, and programming to solve environmental problems. I spent my two years at ITC learning things I wanted with immense joy and rich experiences than I could have imagined.

The successful completion of my MSc degree would not have been possible without Dr. Tiejun Wang – my esteemed first supervisor. From the first meeting when I talked about my interest in elephant conservation to designing research and completing a research thesis on human-elephant dimensions, you have inspired me to further excel in my academic research career. I will forever be grateful for the invaluable advice, guidance, and patience you showed towards me. You understood that this is a learning phase and allowed me to learn from my mistakes while showing immense support throughout the journey. Thank you for everything, Dr. Wang. I would also like to extend my gratitude to my second supervisor, Drs. Raymond G. Nijmeijer for providing comments and grooming my capabilities of writing a good research thesis. I am thankful to Prof. Dr. Andrew K. Skidmore for his critical comments during my research proposal and mid-term evaluation phase.

My sincere appreciation also goes to Dr. Ashok Kumar Ram for making it possible to start a journey at ITC. Assisting you in your PhD project fuelled my passion for elephants. I would like to thank you for your advice, guidance, and valuable comments and for providing elephant presence and human-elephant conflict location data that helped finalise my research thesis.

I thank Faculty ITC, the University of Twente, for an opportunity to be part of one of the best MSc programmes in geo-information science and earth observation. I thank the ITC Excellence Scholarship programme for funding my study and stay in the Netherlands. ITC has provided a platform, resources, connection with research communities and most importantly, with friends from around the world who will remain associated with my life.

I have the most profound respect for all the Department of Natural Resources professors who shaped my thinking and analytical skills. I am thankful for Dr. Thomas A. Groen for helping me understand the concepts of species distribution models. I am grateful to Dr. Mariana Belgiu for helping me understand machine learning/random forest concepts and addressing my questions, even in the middle of the aisle of ITC restaurant.

I would like to thank all my Nepalese community and NRM friends. I would like to thank Enzo, Sahara, Sachita, Shruti, Meenal for supporting me throughout the study. A special thanks to Biman for helping with GIS software, fixing code bugs, and lifting my spirit on bad days. To all my friends in studio 5-036, we made it, guys. Cheers!

I would not have made it this far without the blessings and support from my friends and family back home. To my best friend Nikita, thank you for listening and encouraging me. I would like to thank my brothers Basanta, Prashanta and Praman for supporting me in every way possible. Big thanks to my sisters-in-law, Jyoti, and Nikki Valzi, for cheering me on my success. Special love to my nephew Samyak, who keeps me grounded. Last but not least, I owe my mother and father everything. I can never find a word to describe my gratitude for the love, cherish, support, blessings, and prayers you shower upon me.

This is dedicated to my dear mama, Dhana Kumari. Your prayers have done more miracles than I can think.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ALOS	Advanced Land Observing Satellite
BRT	Boosted Regression Trees
EC	European Commission
FAO	Food and Agriculture Organization of the United Nations
GIS	Geographic Information System
HEC	Human-Elephant Conflict
HWC	Human-Wildlife Conflict
IUCN	International Union for Conservation of Nature
MaxEnt	Maximum Entropy Model
NASA	National Aeronautics and Space Administration
NTNC	National Trust for Nature Conservation
PA	Protected Area
RF	Random Forest
SDM	Species Distribution Model
SRTM	Shuttle Radar Topography Mission
UNEP	United Nations Environment Programme

1. INTRODUCTION

1.1. Background

Humans have a long evolutionary history of interactions with wildlife. Such interactions can be positive or negative. The positive interaction is observed when wildlife or humans benefit each other without posing substantial problems (Buijs and Jacobs, 2021). The negative interaction is when the encounter between wildlife and humans causes adverse effects on each other, threatening the survival of respective species (Nyhus, 2016). Such negative interaction is termed human-wildlife conflict (HWC). Currently, HWC is a serious problem escalating globally attributed to increased competition for the overlapping niche, habitat, and resource uses, thus endangering human well-being and wildlife conservation (Gross et al., 2021). There are visible and hidden impacts associated with HWC. The visible impacts of HWC on humans are manifested as human death/injury, crop and property losses or economic losses, and livestock losses to wildlife (Leslie et al., 2019; Treves et al., 2006). The hidden negative impacts include diminished psycho-social well-being caused due to human death/injury, food insecurity due to crop damage and transactions cost associated with while pursuing compensation for the wildlife damage (Barua et al., 2013). The additional impact of HWC on wildlife is the retaliatory killing, which threatens wildlife conservation. Such retaliatory killing phenomena are common in areas where crop damage and livestock killed by wildlife are high (Distefano, 2005).

There are multiple factors associated with HWC, including ecological, anthropogenic and social factors (Dickman, 2010; Gross et al., 2021), as well as political-economic structure (Fletcher and Toncheva, 2021). The negative consequences of these factors get intertwined, making the issue of HWC complex. Nonetheless, anthropogenic factors are directly linked with forest encroachment and the conversion of forests into the urban and agricultural areas, which results in habitat loss and fragmentation (Lamarque et al., 2009). As the human population grows, the demand for space and resources increases; there is a continuum of conflict between the need for wildlife and human. An extreme consequence of HWC is the significant decline in the wildlife population (Gross et al., 2021) or even the extinction of already endangered species (Nyhus, 2016). Overall, HWC compromises the sustainable development goals of protecting, restoring and promoting sustainable use of natural ecosystems (Secretariat of the Convention in Biological Diversity, 2016). Although not particularly indicated in the Strategic Plan for Biodiversity 2011-2020 and the Aichi Targets, the ongoing negotiations towards the post-2020 Global Biodiversity Framework have highlighted the issue of HWC (Secretariat of the Convention in Biological Diversity, 2016). This allows an opportunity to draw attention from the international and national level to analyse the multiple contributing factors of HWC and focus on mitigation of HWC for bio-diversity conservation and human-wildlife co-existence.

The impacts of a fight between humans and wildlife for space and resource use are prominent in large mammals like elephants with a relatively large home range. The protected areas (PAs) provide the intact forest habitat for elephants, whereas the remaining forest outside the PAs sustains extensive human pressure. The elephants utilise forest cover primarily for their movement triggered by habitat and food requirements (Williams et al., 2008). However, the shrinking of natural habitats, the expansion of human settlements, and the development of infrastructures have brought elephants into frequent contact with humans (Shaffer et al., 2019). Habitat loss, habitat fragmentation and the expansion of farmland have been attributed to HEC (Distefano, 2005; Lamarque et al., 2009; Leslie et al., 2019; Naughton et al., 1999; Nyhus, 2016; Shaffer et al., 2019). The extent of the interface between forest and agriculture, as well as the percentage of farmers involved in crop cultivation each year, contribute to the increased frequency of HEC (Baskaran et al., 2013). When the native habitat is disturbed and unable to provide sufficient resources, elephants get lured toward nutritious crops such as rice, wheat and sugarcane (Sukumar, 2003). Given elephants' colossal body size and weight, crop damage is severe by trampling or consuming crops from agricultural fields and destroying stored grain. The incidents of human casualties occur when people try to chase off attacking elephants. The deaths of 1,713 humans were reported between 2015 and 2018 due to HEC in India (FAO, 2021), whereas USD 245,431 was spent in Sri Lanka in 2018 to compensate for the human casualty and physical property damage caused by elephants (Prakash et al., 2020).

Transboundary landscapes are often characterised by ecologically connected areas involving at least two national borders. Most national borders are imposed based on geographical features like mountainous terrain or other geologically complex landscapes that support high species richness and endemism (Liu et al., 2020). The conservation effort from the individual country is essential to protect threatened and endemic species. However, the transboundary prevalence of species demands combined conservation initiatives beyond the national borders. The management plan for a specific protected area along the transboundary landscape is limited to a local level. It, therefore, overlooks the connectivity and linking corridors to many other protected areas essential for animal movement (Plumptre et al., 2007). Many international communities like the International Union for Conservation of Nature (IUCN), the United Nations Convention on Biological Diversity (CBD) and the Convention on International Trade of Endangered Species (CITES) have been advocating for coordinated conservation actions to avert biodiversity losses, reduce human-wildlife conflicts, stop illegal wildlife trade and promote international governance (Mason et al., 2020; Sandwith et al., 2001; Vasilijević et al., 2015).

The habitat range of more than 50% of terrestrial species shares the international border globally, of which Asia harbours approximately 82% of the border hotspots (Mason et al., 2020). In the context of elephant conservation in Southern Asia, contiguous and connected habitats between countries like Bhutan-India, Bangladesh-India, and India-Nepal is pivotal, ensuring the conservation and free movement of elephants and other wildlife. The open border between Nepal and India allows the free movement of people and animals. The practice of transboundary conservation between India and Nepal is exercised, and there are multiple protected areas and corridors across the transboundary region managed for biodiversity conservation. The Terai arc landscape (TAL) and the Kanchenjunga landscape (KL) are products of such transboundary conservation efforts initiated by Nepal and India (Ministry of Forests and Soil Conservation, 2016), which are home to endangered species like Tiger in TAL (Chanchani et al., 2014), One-horned rhinoceros and Asian elephants in both TAL and KL (Chanchani et al., 2014; Sharma et al., 2020a). However, transboundary conservation is particularly challenging on the Nepal-India border due to the high human pressure (Mason et al., 2020). The effects of habitat loss and fragmentation, road and railway construction, deteriorating habitat quality, forest encroachment, and erection of solar-powered fences compromise the movement of Asian elephants in these landscapes, which contributes to increased human-elephant conflict (Tiwari, 2017). Understanding suitable habitat and resource use by elephants in the landscape and how human pressure modifies those habitats are essential for reducing human-elephant conflict.

The advancement in technology and improvement of data science coupled with a wide range of remotely sensed data acquisition has allowed the potential application of machine learning in ecology and wildlife conservation (Tuia et al., 2022). Machine learning techniques have been applied to perform a multitude of tasks such as land cover change mapping, (Gislason et al., 2006), comparing the performance of different machine learning algorithms in forest fire burn area mapping (Bar et al., 2020), identifying animal species (Willi et al., 2019), species distribution modelling (Gobeyn et al., 2019) and conflict prediction (Kitratporn and Takeuchi, 2019). The growing use of machine learning is the direct result of their ability to deal with the multi-dimensionality of data and model complex nonlinear relationships of variables (Belgiu and Drăgu, 2016; Olden et al., 2008). Machine learning tools and techniques have proved to excel in prediction and enhancing the capabilities to model complex ecological systems (Olden et al., 2008).

Species distribution models (SDMs) are widely used in modelling habitat suitability and extrapolating species distribution over space and time (Guisan and Thuiller, 2005). The earliest use of SDMs can be dated back to the 1920s, assessing the role of climatic determinants in species distribution (Guisan and Thuiller, 2005). Since then, SDMs have evolved to be used in diverse fields (Elith and Graham, 2009) using both statistical and machine learning algorithms (Gobeyn et al., 2019). SDMs are used to develop prediction maps by creating relationships between species occurrence and sets of corresponding environmental variables (Franklin, 2012). These predictive output maps are used to draw conclusions on the species-environment relationship. However, each machine learning based SDM has its unique way of learning from a given training dataset. Each algorithm functions on different assumptions and has its strengths and flaws.

The ensemble learning method is a type of machine learning technique that combines the result number of base learners to produce more generalised final results (Zhou, 2009). The application of ensemble learning is evolving and has been applied in diverse fields, including predicting disease hotspots (Simons et al., 2019), mineral mapping (Wang et al., 2020), landslide susceptibility mapping (Hu et al., 2020), predicting range shift of elephants due to climate change (Kanagaraj et al., 2019) and mapping human-carnivore conflict (Mpakairi et al., 2018). The concept of the ensemble method revolves around the idea that combining the result of individual forecasts lowers the mean error of any of the particular constituent forecasts (Araujo and New, 2007). The ensemble methods in SDMs have an advantage over the single model for the comprehensive assessment, given the plethora of data developed for species modelling (Araujo and New, 2007). The common ensemble methods include bagging, boosting and stacked generalisation (also called stacking) (Sewell, 2008). Ensemble stacking methods refers to combining the prediction of base learners developed from multiple machine learning algorithms by the second-level learner, also known as the meta-learner (Zhou, 2009).

The focus and study on human-wildlife conflict have grown in the last 20 years (König et al., 2020). However, the literature review suggests that studies on habitat suitability and HEC are either focused on small pockets that primarily look at the elephant habitat inside the protected region (Lamichhane et al., 2017) or are limited to individual countries. The cross-border regions between Nepal and India, which provides natural habitats for Asian elephants, have experienced dramatic land cover land-use changes due to human pressure and infrastructure development over the last few decades (Chanchani et al., 2014). The consequence is human-elephant conflict, and it is a problematic issue for both countries. The future of Asian elephants and their habitat conservation, as well as promoting livelihood and well-being of people living in the transboundary region of Nepal and India, are intertwined. Therefore, it is crucial to have knowledge of the distribution of suitable habitats and human-elephant conflict risk in the transboundary landscape for effective management of the landscape in a holistic and ecologically sound manner.

1.2. Asian elephant and its habitat use

The historical records indicate that the distribution of the Asian elephant (*Elephas maximus*) ranged from Tigris-Euphrates in western Asia, throughout the southern region of the Himalayas in South Asia and north into China as far as the Yangtze river (Olivier, 1978). At present days, the habitat of Asian elephants is lost in western Asia, and the status of the species remains extinct (Williams et al., 2020). The current distribution is now limited to 13 ranging countries, including Bangladesh, Bhutan, Burma, China, Cambodia, India, Indonesia, Laos, Malaysia, Nepal, Sri Lanka, Thailand, and Vietnam (Sukumar, 2003;

Williams et al., 2020) with a population of 48,323-51,680 in the wild (Menon and Tiwari, 2019). India has the highest number of Asian elephants, with a population of 29,964, while the population ranges from 109 to 145 in Nepal (Menon and Tiwari, 2019). However, these elephant populations are widespread in isolated landscapes and persist as fragmented populations. The Asian elephant is listed as endangered on the IUCN Red List (Williams et al., 2020). It is also listed in Appendix I of the Convention on International Trade in Endangered Species of Wild Flora and Fauna (UNEP-WCMC, 2021) and protected under the Wildlife Protection Act (1972) of India and the National park and wildlife conservation act 2029 (1973) of Nepal (Singh et al., 2019).

Asian elephants are mega-herbivores and generalist feeders. Given the large body size, elephants spend most of their time (up to 14-19 hrs a day) feeding, during which they can consume up to 150 kg (Williams et al., 2020) and between 38-57 litres (10-15 gallons) of water daily (Eisenberg, 1980). Elephants are both grazers and browsers and feed on a variety of plants. Elephants primarily graze on tall and newly grown grasses during the wet season and browse on tall trees during dry seasons (Sukumar, 1990). The study in Southern India reports that elephants feed on fresh foliage of species like *Acacia pennata*, fruits of species like *Tamarindus indica*, culms and lateral shoots of bamboo species, leaves and twigs of species like *Acacia* and *Ficus* species and bark of certain plant species like *Tectona grandis* and *Eucalyptus* spp. (Sukumar, 1990). Depending on the availability and seasonality, elephants also raid different crops, including maize, paddy, wheat and garden fruits like banana, sugarcane, jackfruit, and mango (Pant et al., 2016; Sukumar, 1990).

Asian elephants exhibit distinct seasonal movements to optimise the use of resources in different habitat types (Sukumar, 2003). During their movement, elephants make use of diverse vegetation types whose spatial distribution is not only influenced by topographic and climatic factors but also by human intervention. Even though the primary native habitat of Asian elephants consists of a naturally vegetated forest habitat (Sukumar, 1989; ten Velde, 1997), they have learned to take advantage of the human-modified agrarian landscape in the present scenario (Krishnan et al., 2019; Lamichhane et al., 2017). In fact, the study finding by Calabrese et al. (2017) suggests that forest and agricultural mix habitat promotes elephant abundances in the given landscapes. The fragmented forest remnants in such habitat mixtures serve the purpose of shelter and foraging needs in the natural habitat while providing an opportunistic switch to the agricultural field for nutritious and palatable crops (de la Torre et al., 2021; Sitompul et al., 2013; Sukumar, 2003). The indulgence of elephants in such crop-raiding for nutritional benefits, however, is counteracted by the HEC as it increases the contact with humans and cultivation (Calabrese et al., 2017; Sukumar, 2003).

The home range of Asian elephants differs according to the available habitat, resource requirement, sex, and reproductive necessity. Previous studies showed that the home range of elephants in Southern India was estimated to be 105 to 320 km^2 (Sukumar, 1989), 188 to 400 km^2 in northwest India (Williams et al., 2008) and 34 to 232 km^2 in Sri Lanka (Fernando et al., 2008). Elephants can travel up to 10-20 km in a day (Sukumar, 2003).

1.3. Problem statement

An important step in understanding the distribution of human-elephant conflict risk is to develop knowledge about the ecology of elephants, their habitat preferences, and the distribution of suitable habitats. Since the Asian elephant is a large migratory animal that knows no political boundary, it becomes crucial to understand habitat preference in larger landscapes and, in this case, the transboundary landscape of Nepal and India. While most of the research is focused in either a single country, there are a very few numbers of researches on Asian elephant being done that focuses on the transboundary level. Kanagaraj et al. (2019) studied the current suitable habitat and predicted range shift under climate change along with the elephant's range in Nepal and India. Padalia et al. (2019) and Ram et al. (2021b) assessed forest loss and fragmentation in the historical range of Asian elephants in India and Nepal, respectively. The study by Ram et al. (2022) predicted HEC in the Chure Terai Madhesh landscape located in Nepal, whereas Sharma et al. (2020a) mapped human-wildlife conflict hotspots in the Kanchenjunga transboundary landscape between Nepal, India and Bhutan in eastern Himalayan. However, to my best knowledge, a study on human-elephant conflict integrating the information of habitat suitability covering the entire Nepal-India transboundary landscape has not been done before.

Humans influence land-use dynamics and modify the landscape attributes, causing socio-economic, climatic, and biophysical changes (Shaffer et al., 2019). The conceptual diagram (Figure 1) describes the intricacies between the elephant and people's attributes and how they contribute to human-elephant conflict. HEC incidents are expected to be observed in the area where the common interest of humans and elephants in resource use overlaps. The delicate harmony between human and elephant existence is broken down because of anthropogenic factors like population growth and forest encroachment, thus, exacerbating negative interactions as the overlapping area of shared resource use expands. Every year, there is a record of elephants raiding crops and killing elephants either by electrocution (Onlinekhabar, 2021) or explosion (The Guardian, 2020).

Effective conservation planning requires a deeper understanding of the use of landscape attributes by elephants and the impacts of humans on those attributes. While it is crucial to analyse the relationship between elephant habitat preference and HEC to embrace the HEC mitigation approach promoting human-elephant co-existence, the issue is sensitive given the national and international conservation status of the Asian elephant. The use of machine learning algorithms in modelling habitat suitability and predicting human-elephant conflict risk is expected to produce accurate and reliable information to assist in informed decision-making for the management of HEC mitigation strategy.

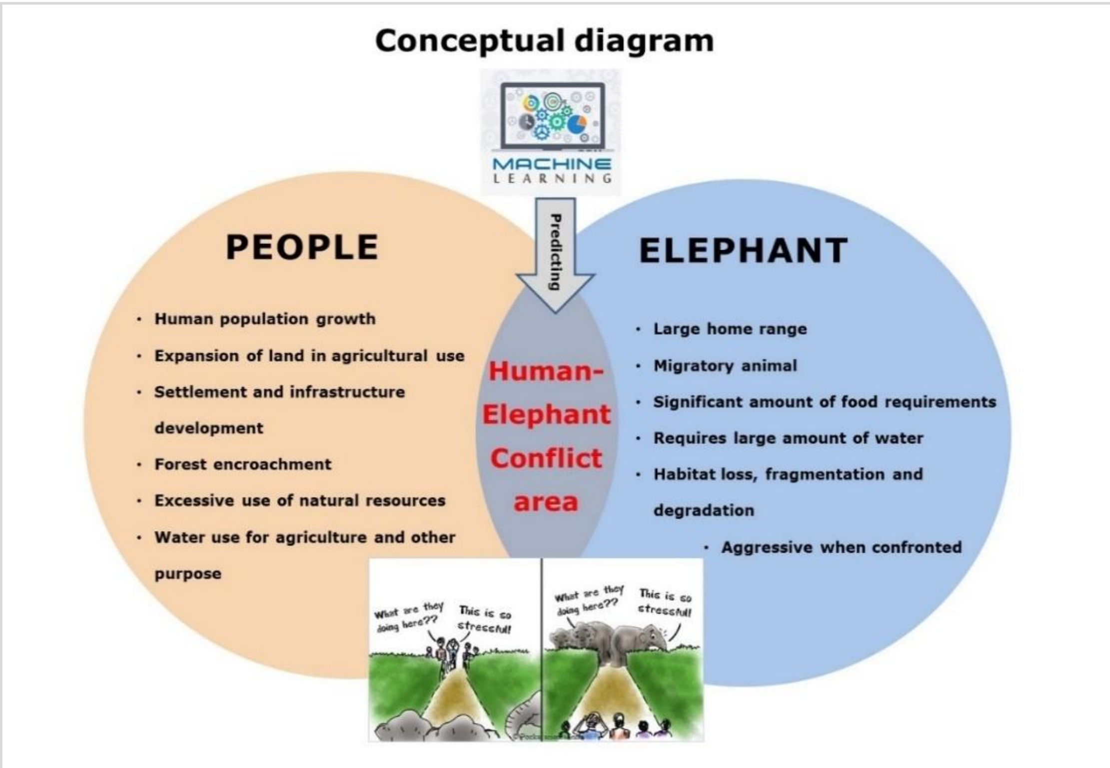


Figure 1: Conceptual diagram illustrating human-elephant conflict.

1.4. Research objective

The overall objective of this study is to predict the risk of human-elephant conflict along the transboundary landscape of Nepal-India using machine learning algorithms. The specific objectives are:

1. To map the suitable habitat for Asian elephants
2. To determine the key factors influencing the habitat suitability of Asian elephants
3. To predict the risk of human-elephant conflict
4. To identify key factors shaping the spatial distribution of human-elephant conflict

1.5. Research questions

1. Where and how many suitable habitats are available for the Asian elephant in the transboundary landscape of Nepal-India?
2. What are the key factors determining the habitat suitability for the Asian elephant in the transboundary landscape of Nepal-India?
3. Where are the risk areas for human-elephant conflict in the transboundary landscape of Nepal-India?
4. What are the key factors determining the human-elephant conflict in the transboundary landscape of Nepal-India?

1.6. Research hypothesis

H_0: Habitat suitability of elephants is not a key factor (i.e., top four factors) contributing to the risk of human-elephant conflict.

H_1: Habitat suitability of elephants is one of the key factors (i.e., top four factors) contributing to the risk of human-elephant conflict.

2. MATERIALS AND METHODS

2.1. Study area

The study area includes the transboundary region of Nepal and India (Figure 2). The total size of the study area is approximately 74,808 km². Elephants in India are distributed in four general regions: northeastern India, central India, northwestern India and southern India, whereas in Nepal, they are found along southern lowland Terai, which borders India (Williams et al., 2020). The elephant populations in Nepal can be found in four isolated groups: eastern, central, western and far-western (Pradhan et al., 2011). The transboundary movement of elephants is observed mainly between northeastern India and the eastern border of Nepal and between northwestern India and the southwest border of Nepal, extending along the foothills of the Shivalik/Churia and lowland Terai floodplains (Pradhan et al., 2011; Williams et al., 2020). Within this area, there are fifteen protected areas (PAs) (Figure 2). Nine PAs lie in India: Nandhaur Wildlife Sanctuary(1), Pilibhit Tiger Reserve(2), Kishanpur Wildlife Sanctuary(4), Dudhwa National Park(5), Katarniaghat Wildlife Sanctuary(6), Sohelwa National Park(9), Sohagibarwa Wildlife Sanctuary(10), Valmiki Tiger Reserve(11) and Mahananda Wildlife Sanctuary(15), whereas six PAs are located in Nepal: Suklaphanta National Park(3), Bardiya National Park(7), Banke National Park(8), Chitwan National Park(12), Parsa National Park(13) and Koshi Tappu Wildlife Reserve(14). A total of 7,835 km² of the study area is protected under the PAs. The habitat corridors in the study area connect the PAs and other forest habitats allowing the safe passage for wildlife movements.

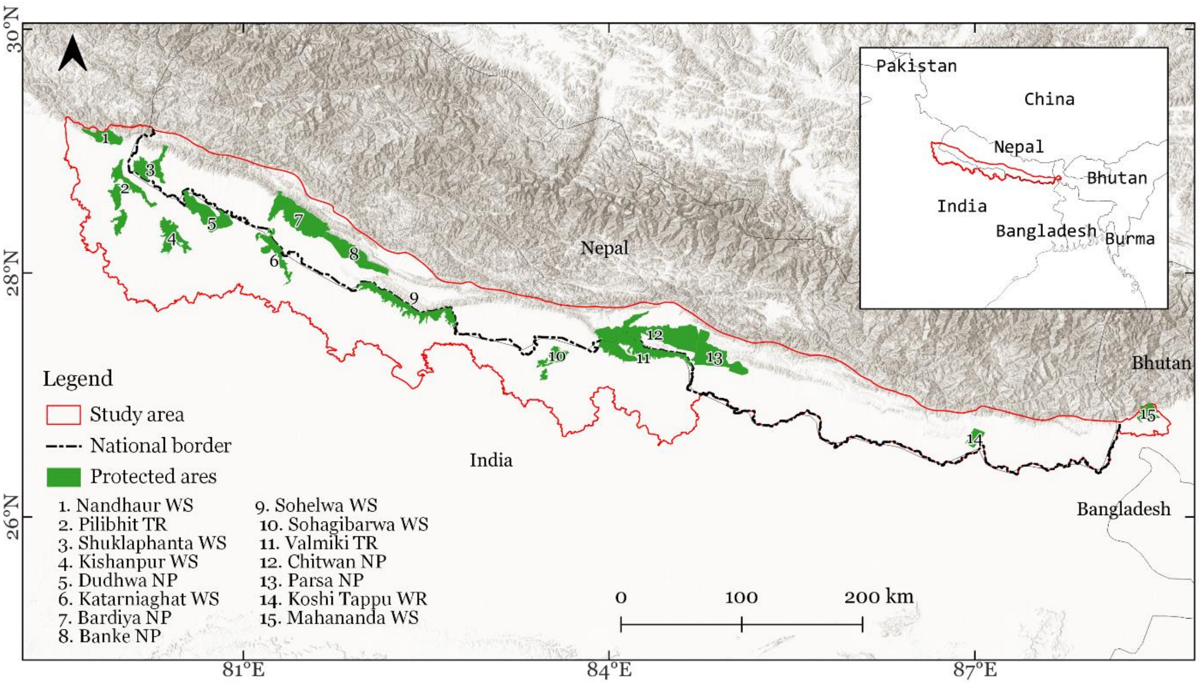


Figure 2: Study area map in the transboundary region of Nepal and India. The map shows protected areas which serve as the prime natural habitat of Asian elephants.

The climate in the study area is divided into four distinct seasons: winter (Nepal: mid-December/mid-March; India: January-February), pre-monsoon or hot weather season (Nepal: mid-March/mid-June;

India: March-May), monsoon (Nepal: mid-June/mid-September; India: June-September) and autumn or post-monsoon (Nepal: mid-September/mid-December, India: October-December) (De et al., 2005; Lamichhane et al., 2017). The mean temperature ranges from 35-40°C during the hot weather season and 14-16°C during the winter season (Lamichhane et al., 2017).

The vegetation is characterised by moist tropical and subtropical forests (Naha et al., 2019). It is composed of mostly *Shorea robusta* and *Terminalia* species forests, riverine forests of *Acacia catechu* and *Dalbergia sissoo*, plantations of *Tectona grandis* and *Eucalyptus* spp., and grasslands, including tall floodplain and open grasslands (Chanchani et al., 2014). This region supports a diverse ecosystem. Apart from the elephant, the region is home to mammalian species like a tiger (*Panther tigris*), greater one-horned rhinoceros (*Rhinoceros unicornis*), gaur (*Bos gaurus*), wild buffalo (*Babulus bubalis*), blackbuck (*Antilope cervicapra*), sloth bear (*Ursus ursinus*), dhole (*Cuon alpinus*) and many other species (Chanchani et al., 2014).

The Nepal-India transboundary landscape is a human-dominated landscape where most of the area is used for agriculture cultivation. The forest cover is mainly observed in the northern part along the foothills of Shivalik/Churia, as shown in the landcover/land-use map of the study area (Figure 3) developed by the European Space Agency (ESA) for the year 2020.

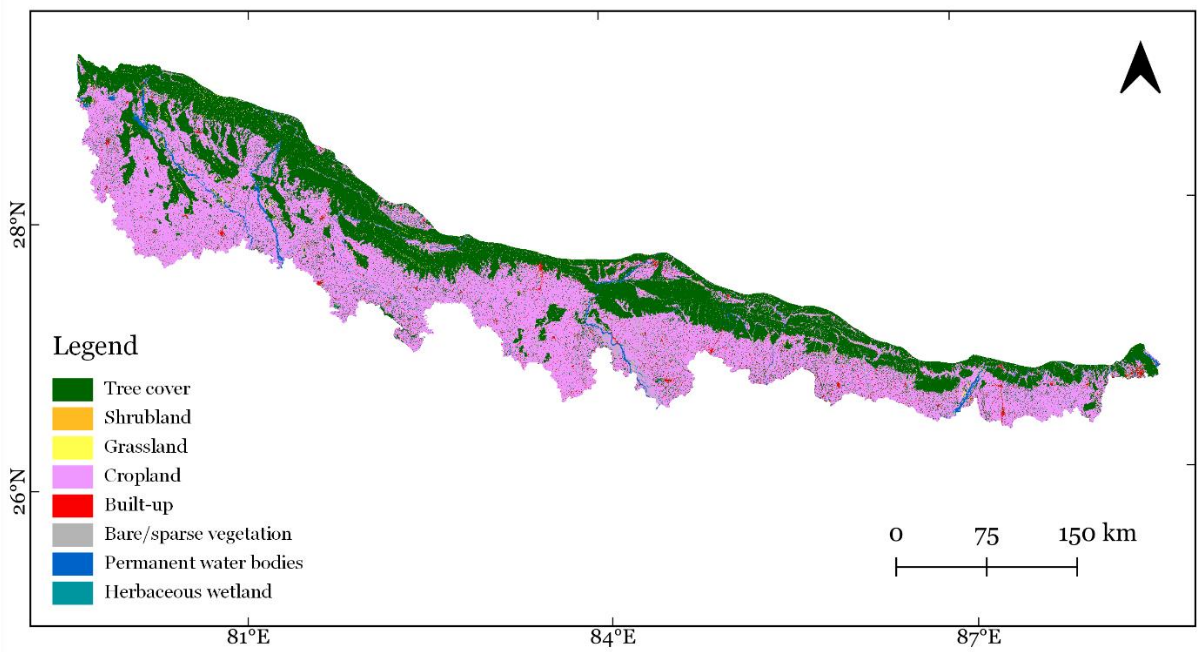


Figure 3: Land cover/land-use map of the study area.

2.2. Datasets

2.2.1. Elephant occurrence data

The elephant occurrence data were collected from various sources from January 2017 to June 2020. The occurrence points of Asian elephants for Nepal were compiled from occupancy field survey data of the PhD research project that I assisted in data compilation and analysis (Ram, 2021). I took permission from principal researcher Ashok Kumar Ram to partially use the data owned by him for my research work. For an occupancy field survey, the study area in Nepal was divided into 15×15 km grids and 7.5×7.5 km sub-grids. Elephant presence signs were surveyed along the transects within the grids. A total of 136 presence points were collected from the occupancy survey in Nepal. Additional two presence points were collected from the published article (Sharma et al., 2020b). In the case of India, a total of 20 presence points were collected from local newspapers which reported the sighting of elephants. The smallest administrative bodies, like the village name mentioned in the newspaper, were identified from local knowledge. An additional five elephant occurrence points were collected from the global biodiversity information facility[1]. A total of 163 elephant presence points were used in habitat suitability modelling.

In order to reduce spatial autocorrelation in the dataset, which may produce biased results, elephant occurrence points were selected at least 5.6 km apart from each other, which is the radius of the minimum circular home range size of elephants, taken as 100 km^2 in the study area (Kanagaraj et al., 2019).

2.2.2. Human-elephant conflict incidence data

The HEC incidence data were also collected from various data sources from January 2017 to June 2020. The HEC data for Nepal was collected from a PhD research project (Ram, 2021). I was involved in the research project for data collection, compilation, and analysis. I partially used the primary field data with permission from the principal investigator. The details of the data and data collection method can be sourced back to the published paper by the same author (Ram et al., 2022). The additional data on HEC incidence was collected from other published articles (Naha et al., 2019) and local news sources. The total number of incidence points collected from published articles was 62, whereas 38 points were collected from news sources.

In order to reduce spatial autocorrelation in the dataset, which may produce biased results, HEC incidence points were selected at least 5.6 km apart from each other.

2.2.3. Predictor variables for elephant habitat suitability modelling

The choice of predictor/explanatory variables is the centre step in species distribution modelling (SDM) because the choice drives the modelled spatial distribution output (Araújo and Guisan, 2006). In this study, potential relevant predictor variables were selected based on the literature on the ecology, range and habitat distribution of elephants (Baskaran et al., 2013; Chen et al., 2016; de la Torre et al., 2021; Kanagaraj et al., 2019; Kitratporn and Takeuchi, 2019; Lamichhane et al., 2017; Shaffer et al., 2019; Sharma et al., 2020a; Sukumar, 1989). The identified twenty-seven predictor variables were grouped into five categories: bio-climatic, vegetation, topographic, and water (Table 1). All predictor variables were downloaded from different sources with different spatial resolutions and projections. It is essential to maintain the same extent and spatial resolution of all layers for analysis using SDMs. The spatial resolution of ~1km was maintained for all the variables justifiable based on the large area of study area and behaviour of the Asian elephant. Thus, all the other variables with different spatial resolutions were resampled to ~1km using the bilinear method for continuous data and the majority votes method for categorical data. They were then reprojected to EPSG: 32644 – WGS 84/UTM 44N zone and clipped to the study area.

The nineteen bioclimatic variables represent the climatic factors derived from monthly temperature and precipitation data and aggregated across the temporal range of 1970 to 2000 (Fick and Hijmans, 2017).

The topography of the geographical area determines the habitat used by Asian Elephants (ten Velde, 1997). The SRTM 90m elevation data produced by NASA was used to derive two topographical factors – elevation and slope.

The availability and distribution of vegetation, including forest and grassland, largely influence the habitat use and home range of elephants (Sukumar, 1989; ten Velde, 1997). The vegetation indices are widely used remote sensing products used to represent the greenness, variation, and primary productivity of the

[1] https://www.gbif.org/

vegetation. The study area is characterised mainly by tropical habitats, and therefore, Moderate Resolution Imaging Spectro-radiometer (MODIS) enhanced vegetation index (EVI) data was used, which is known to have improved sensitivity over high biomass. Three 12-months MODIS (MOD13Q1 version 6.1) time series of 16 days composite with 23 granules each at 250 m spatial resolution for 2017-2019 were created for the study area. Two tiles of MODIS data were required to cover the study area. From each downloaded MODIS granule, EVI data was extracted by tiles, mosaicked, reprojected from sinusoidal to EPSG:32644 UTM zone 44N and subset to the study area. The Savitzky-Golay filtering was used to smoothen the observed anomalies in the temporal profiles of EVI time series data. Four EVI indices – annual mean, annual minimum, annual maximum, and annual standard deviations EVI were derived in ENVI Classic 5.6.1. Vegetation vertical structure also influences the distribution of species. It is the proxy for vegetation productivity and modification in the habitat type. The forest canopy height developed by integrating NASA GEDI lidar forest structure measurements and Landsat analysis-ready data series by the Global Land Analysis and Discovery lab at the University of Maryland (Potapov et al., 2021) was used to represent the vertical vegetation structure.

The OpenStreetMap (OSM) data provides free access to the geographic database of the world. The QuickOSM plugin in the Quantum-GIS (Q-GIS) software enables the extraction of OSM spatial data through overpass API. The extraction process requires a query specifying the "key" and "value" based on the OSM map features to cull and organise the required data. The layer of water bodies, including rivers and reservoirs, were generated using OSM data through overpass API in Q-GIS 3.22.2. The "key = natural" and "value=water" was used to query the available waterbodies layer from the OSM data for the given extent of the study area. The final water bodies layer was filtered based on the water presence for more than eight months and local knowledge on which water sources are used mainly by elephants. The Euclidean distance to water bodies was calculated using the Spatial Analyst toolbox in ArcMap 10.8.2.

Table 1: Predictor variables for elephant habitat suitability modelling

Category	**Variables**	**Unit**	**Abbreviation**	**Data source**
Bio-Climatic	Annual Mean Temperature	°C	bio1	WorldClim
	Mean Diurnal Range	°C	bio2	
	Isothermality	Unitless	bio3	
	Temperature Seasonality	°C	bio4	
	Max Temperature of Warmest Month	°C	bio5	
	Min Temperature of Coldest Month	°C	bio6	
	Temperature Annual Range	°C	bio7	
	Mean Temperature of Wettest Quarter	°C	bio8	
	Mean Temperature of Driest Quarter	°C	bio9	
	Mean Temperature of the Warmest Quarter	°C	bio10	
	Mean Temperature of Coldest Quarter	°C	bio11	
	Annual Precipitation	mm	bio12	
	Precipitation of Wettest Month	mm	bio13	
	Precipitation of Driest Month	mm	bio14	
	Precipitation Seasonality	Unitless	bio15	

Category	Variables	Unit	Abbreviation	Data source
	Precipitation of Wettest Quarter	mm	bio16	
	Precipitation of Driest Quarter	mm	bio17	
	Precipitation of Warmest Quarter	mm	bio18	
	Precipitation of Coldest Quarter	mm	bio19	
Topographic	Elevation	m	elev	SRTM/NASA
	Slope	Degree	slope	
Vegetation	Forest Canopy Height	m	forHeight	UMD GLAD
	Annual maximum EVI	Unitless	evi_max	NASA/EARTHDATA
	Annual minimum EVI	Unitless	evi_min	
	Annual mean EVI	Unitless	evi_mean	
	Standard deviation EVI	Unitless	evi_stddev	
Water	Distance from water bodies	m	dist2water	OSM

2.2.4. Predictor variables for human-elephant conflict risk prediction

HEC is attributed chiefly to anthropogenic factors. The nine variables used to predict HEC risk are habitat suitability map, forest fragmentation metrics, settlement density, population density, distance to the protected area, distance to road and distance to the forest (Table 2). The variables were resampled to ~1km using the bilinear method, reprojected to EPSG: 32644 – WGS 84/UTM 44N zone and clipped to the study area for further analysis.

Understanding the spatial extent of suitable habitats for elephants in human-dominated landscapes is essential to draw insights into the dimensions of HEC and risk areas. The development of the habitat suitability layer in this research is the intermediate step which is used as one of the predictor variables in HEC risk prediction.

Forest fragmentation was taken as a proxy for habitat fragmentation. The forest/non-forest raster layer available at 30 m resolution for 2017 was downloaded from ALOS PALSAR[2] and used to calculate forest fragmentation metrics. The raster version of FRAGSTATS 4.2 computes continuous fragmentation metrics for statistical analysis using a circular moving window sampling strategy (Mcgarigal, 2015). For this study, the choice of appropriate scale (radius of a circular moving window) was based on the elephant's average home range in the study area. The average size of the home range of elephants was assumed to be circular and taken as 300 km^2 which gives a radius scale of 9.7 km for calculating fragmentation metrics at a landscape level. The binary forest/non-forest layer was coded as 1-non-forest and 2-forest area, and three forest fragmentation metrics— landscape division index, splitting index, and effective mesh size were calculated.

The landscape division index is based on the degree of cohesion, defined as the probability of two animals finding each other when they are present in a different area within a given landscape (Jaeger, 2000). The degree of landscape division indicates that two forest patches are not situated in the same undissected

[2] https://www.eorc.jaxa.jp/ALOS/en/index_e.htm

patch in a given landscape (Mcgarigal, 2015). The value of the landscape division index ranges from 0 ≤ Division ≤ 1. The highest value means that the patch type consists of a single small patch of one cell in the area. A division equal to one indicates a highly fragmented landscape (Mcgarigal, 2015).

The splitting index is defined as the resulting number of patches when the entire landscape is divided into equal-sized patches so that the new configuration leads to the same degree of landscape division as obtained for the observed cumulative area distribution (Jaeger, 2000). The splitting index value ranges from 1≤ Split ≤ number of cells in the landscape area squared. The value of split is 1 when the landscape consists of single continuous patches. The upper limit increment corresponds to an increase in the number of single small patches indicating forest fragmentation (Mcgarigal, 2015). The value of SPLIT is unitless.

The effective mesh size indicates the size of the patches one gets when the landscape is divided into S areas of the same size with the same degree of landscape division as obtained for the observed cumulative area distribution (Jaeger, 2000). The range of effective mesh size is the ratio of cell size to landscape area ≤ Mesh ≤ total landscape area, where the lower limit corresponds to a single one-pixel patch meaning highly fragmented. It is maximum when the landscape is a single continuous patch (Mcgarigal, 2015).

The Euclidean distance to the forest was calculated using the Spatial Analyst toolbox in ArcMap 10.8.2 from the same binary forest/non-forest raster layer used to calculate the forest fragmentation metrics.

The human settlement density layer (GHS-BUILT-S2 R2020A) was downloaded from the Global Human Settlement Layer (GHSL) datasets at 10 m spatial resolution for the year 2017-2018 provided by the Joint Research Centre (JRC), European Commission. The GHS-BUILT-S2 layer is derived from the Sentinel-2 global image composite for the reference year 2018 using Convolutional Neural Networks (Corbane et al., 2021). The GHS-BUILT-S2 layer grid corresponds to the built-up area probability percentage values (Corbane et al., 2021).

The population density layer (GHS_POP_E2015_GLOBE_R2019A_4326_9SS_V1_0) was downloaded from the GHSL dataset provided by the JRC at 250 m spatial resolution. The GHS population data is derived from the Gridded Population of the World collection, fourth version (GPW4.1), and Landsat-based GHS-BUILT were used as target layers for disaggregation of population estimates for the reference year 2015 (Marcello et al., 2019). The value represents the number of people per cell.

The polygon shapefile of protected areas for Nepal was downloaded from NTNC[3]. In the case of protected areas for India, the layer was generated from OSM data through overpass API in Q-GIS 3.22.2 using "key=boundary" and "value=protected_area" for the query. The protected area layer consists of national parks and wildlife reserves/sanctuaries for the study area. The Euclidean distance to the protected area was calculated using the Spatial Analyst/Euclidean distance toolbox in ArcMap 10.8.2.

The linear shapefile of the road layer was generated from OSM data using overpass API in QGIS 3.22.2 using "key=highway" and "value=trunk", "value=primary" and "value=secondary" for the query. The road layer consists of major highways and sub-highways. The Euclidean distance to the road was calculated using the Spatial Analyst/Euclidean distance toolbox in ArcMap 10.8.2.

[3] http://geoportal.ntnc.org.np/

Table 2: Predictor variables for human-elephant conflict risk prediction

Variables	Unit	Abbreviation	Data source
Habitat Suitability	Unitless	hab_suit	
Distance from forest	m	dist2forest	OSM
Human Population Density	Number of people/sq. km	pop_den	EC JRC
Settlement Density	Proportion	settle_den	EC JRC
Distance to roads	m	dist2road	OSM
Distance to the protected area	m	dist2PA	OSM
Landscape Division Index	Proportion	division	
Splitting Index	Unitless	split	
Effective Mesh Size	ha	mesh	

2.3. Multi-collinearity analysis of predictor variables

The issue of multi-collinearity occurs when two or more explanatory variables are intercorrelated so that the phenomena of one variable can be predicted by others to a great extent. The effect of collinearity in SDMs can be observed while training the model and extrapolating the model prediction to unfamiliar environmental conditions (Dormann et al., 2013), thus producing uncertainty in model results. There is a growing consensus that machine learning algorithms such as Maximum Entropy might take advantage of existing collinearity in prediction capability (De Marco and Nóbrega, 2018; Feng et al., 2019); nevertheless, the multi-collinearity might pose a severe problem while extrapolating beyond the environmental range of sampled data (Dormann et al., 2013; Feng et al., 2019) and change the variable importance in the models (Genuer et al., 2010). The Variance Inflation Factor (VIF) and pairwise correlation coefficient (r) were used to diagnose the collinear variables. The variables with VIF > 10 and $|r| > 0.7$ have the multi-collinearity issue (Dormann et al., 2013) and, thus, are removed from model training.

2.3.1. Multi-collinearity analysis of predictor variables for elephant habitat suitability modelling

Most bioclimatic variables had a VIF value greater than ten and were highly correlated to each other and the topographic factors (Annex 1). Out of twenty-seven predictor variables, nine were selected for habitat suitability modelling. The final selection of predictor variables had VIF < 10 (Table 3) and $|r| < 0.7$ (Figure 4), which reflects the relative importance of habitat suitability modelling of Asian elephants in the study area.

Table 3: VIF values of the selected variables for elephant habitat suitability modelling

S.N.	Variables	VIF
1	bio5	8.18
2	elev	7.70
3	bio6	3.19
4	bio3	5.71
5	bio14	2.58
6	bio13	1.92
7	evi_min	1.77
8	evi_stddev	1.35
9	dist2water	1.26

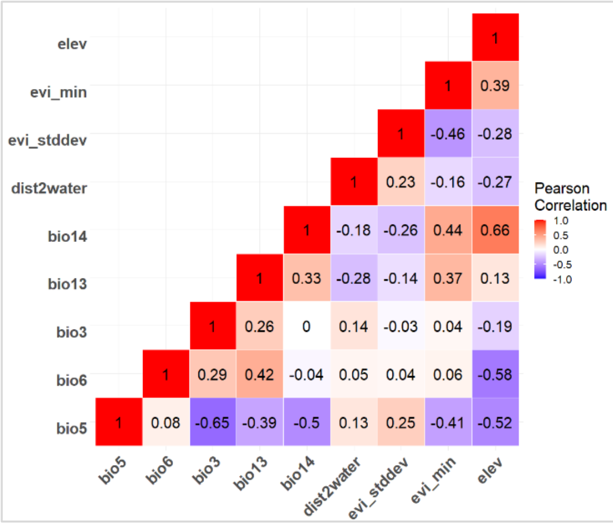


Figure 4: Pearson's correlation coefficient of predictor variables selected for elephant habitat suitability modelling

2.3.2. Multi-collinearity analysis of predictor variables for human-elephant conflict risk prediction

None of the identified predictor variables for human-elephant conflict risk showed an issue of multi-collinearity. All nine variables had a VIF value of less than ten (Table 4) and a Pearson's correlation coefficient value of less than 0.7 (Figure 5).

Table 4: VIF values of the selected variables for human-elephant conflict risk prediction

S.N.	Variables	VIF
1	dist2forest	1.58
2	settle_den	1.56
3	division	1.44
4	pop_den	1.41
5	hab_suit	1.36
6	mesh	1.32
7	dist2PA	1.22
8	split	1.13
9	dist2road	1.08

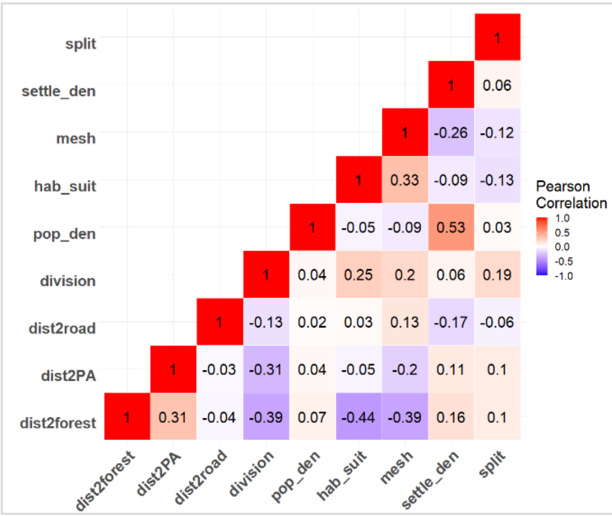


Figure 5: Pearson's correlation coefficient of predictor variables selected for human-elephant conflict risk prediction

2.4. Methods

2.4.1. Machine learning models for habitat suitability modelling

The modelling of habitat suitability to produce a final habitat suitable map was done in two steps. First, three individual Boosted Regression Trees (BRT), Maximum Entropy (MaxEnt) models, and Random Forest (RF) were built as base learners. Second, the habitat suitability predictions from each base learner were combined by meta learners to produce a final habitat suitable map. The process is known as ensemble stacking, and the general structure is shown in Figure 6. Logistic Regression (LR) was taken as a meta-learner as suggested and applied in previous studies (Hu et al., 2020; Ting and Witten, 1999; Wang et al., 2020).

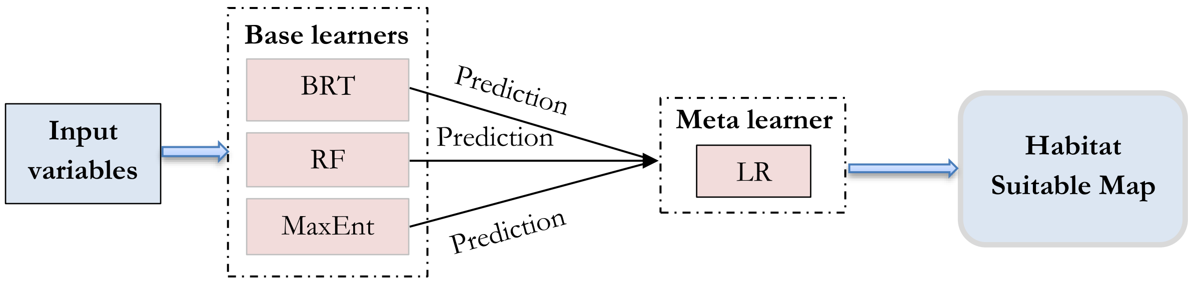


Figure 6: Ensemble stacking process for habitat suitability mapping

2.4.1.1. Boosted Regression Trees

The Boosted Regression Trees (BRT), an extended form of decision trees, is an ensemble method that uses a single base learning algorithm to produce homogenous base learners (Zhou, 2009). It is a non-parametric regression technique that uses two algorithms; first, building regression trees and second, boosting to build the regression trees and combining a collection of results. Boosting is a form of functional gradient descent opted as a numerical optimisation technique to minimise the loss (loss in predictive performance) function by adding a new tree at each step. The existing trees are left unchanged as the new trees are added based on loss functions like deviance. Usually, the linear combinations of hundreds to thousands of such trees give the BRT model (Elith et al., 2008). Decision trees are greedy learners, and there are three main factors for regularising BRT and constraining overfitting. One is the learning rate, also known as the shrinkage parameter, which indicates the contribution and strength of each tree to the final model. The second factor is tree complexity, which determines the number of trees splits to check the fitted interactions. The third factor is the number of trees, determined based on learning rate and tree complexity, to find the optimal prediction (Elith et al., 2008).

The particular advantages of BRT include the ability to deal with different types of response variables (Elith et al., 2008), predictive accuracy, accommodating the ability of missing data and the ability to describe complex non-linearities and interactions between variables (Colin et al., 2017). However, challenges can arise in remote sensing to accommodate large data sizes and incorporate spatial information with disjoint geographic areas while using BRT (Colin et al., 2017).

2.4.1.2. Maximum Entropy

The Maximum Entropy (MaxEnt) is a widely accepted machine learning technique that originated from statistical mechanics. MaxEnt calculates the distribution of target species over geographical spaces based on maximum entropy (closest to uniform) under an assumption that the expected value of each feature matches its empirical average under the estimated distribution (Phillips et al., 2004). It has the advantage of requiring the presence-only data to produce a probability distribution of species based on an informed set of climatic/environmental features, including continuous and categorical variables (Phillips et al., 2006). The modelling of habitat suitability of plant species like *Juniperous* spp. (Boogar et al., 2019), *Aedes albopictus* using Bioclim variables (Ibáñez-Justicia et al., 2020), *Panthera uncia* (Bai et al., 2018) and *Elephas maximus* (Asian elephant) (Huang et al., 2019) using MaxEnt suggests its frequent and comprehensive use.

Maxent is known to perform well with sparse samples (Elith and Graham, 2009) and samples with spatial error (Baldwin, 2009). However, MaxEnt is prone to overfitting if the regularisation parameter is not fine-tuned properly.

2.4.1.3. Random Forest

Random Forest (RF), a non-parametric classifier, is an extension of decision trees that uses ensemble bagging approaches where each model/classifier is trained on a random subset of training data through replacement, also called bootstrap samples (Belgiu and Drăgu, 2016). Bootstrapping allows the selection of two-thirds of total samples, also known as in-bag samples, to be used for training the classifier, whereas one-third of total samples, also known as out-of-bag (OOB) samples, are used as internal cross-validation on a model developed by a classifier to assess the performance of the model (Breiman, 2001).

In RF, there are two parameters to consider for fine-tuning the model to avoid overfitting. The 'mtry' is the number of randomly drawn features, and the best split is selected within the subset of features allowing the tree to grow without pruning (Breiman, 2001; Probst et al., 2019). The 'ntree' defines the number of trees to be grown. Both parameters are user-defined and used to develop trees with a reduced bias (Breiman, 2001).

RF is relatively robust to outliers and noise and provides an internal estimate of error and variable importance (Breiman, 2001). RF can deal with multi-dimensionality and mislabelled data in remote sensing; however, it is sensitive to spatial autocorrelation of the training classes (Belgiu and Drăgu, 2016).

2.4.1.4. Training machine learning models and ensemble stacking

The elephant's occurrence dataset contained only presence data; therefore, background or pseudo absence data were generated to create a binary dataset for training the models. Generally, it is recommended to generate large background samples representing the wider range of environmental conditions being considered for habitat suitability modelling (Barbet-Massin et al., 2012). A total of 5000 background points were generated randomly across the study area to produce a dataset with presence and pseudo-absence points for habitat suitability modelling. All the values from the layer of predictor variables were extracted for the given presence and pseudo-absence points and randomly divided into 70% for training and the remaining 30% for testing the model for the performance assessment.

For training the three base learners, the 70 per cent training dataset was divided randomly into five equal subsets, of which four sets were used for training, and one set was used for predictions for each model. The process was repeated five times, and the predictions were used to train the meta learner – LR model for producing the final predicted habitat suitability map. The ensemble stacking training process is shown in Figure 7.

All the models were fitted in R version 4.1.2., using different R-based packages to facilitate the model fitting. The BRT model was fitted using *dismo* package version 1.3-5 in the R environment. The BRT model was regularised by trying the different values of regularisation parameters: learning rate (lr), tree complexity (tc) and number of trees (nt). The number of trees has an inverse relationship with learning rate and tree complexity (Elith et al., 2008). It is preferable to have a minimum of 1000 trees and slower learning rate values because the contribution of each tree is shrunk more, which helps reduce error and produce a reliable estimate of the response by the model (Elith et al., 2008). The BRT model was fitted with the value lr = 0.001 and tc =3 such that the number of trees is more than 1000, and the fitted model was used to predict the suitability of the habitat in the study area.

The RF model was fitted using *randomForest* package version 4.7-1 in the R environment. In the remote sensing community, it is recommended to set ntree = 500 and mtry to the square root of the number of input variables (Belgiu and Drăgu, 2016). When implementing RF for species distribution modelling with a large number of background samples compared to presence samples, it is known that the performance of the RF model is compromised and prone to biased classification because of the class imbalance and overlap between the number of presence and background samples (Valavi et al., 2021). This is due to over-representation of majority class –background samples, leading to type II error causing under-prediction of minority class – presence samples. A down-sampling approach is one of the effective ways to address the class imbalance issue and is known to improve the model's predictive power substantially (Valavi et al., 2021). The down-sampling approach in the classification-RF utilises the same number of background samples in each classification tree by sampling with replacement from the complete background set (Valavi et al., 2021). The RF model was fitted by down-sampling the majority class and setting the value of ntree =500 and mtry =3. The probability of presence samples was computed for habitat suitability prediction in the study area.

The MaxEnt model was fitted using *dismo* package version 1.3-5 in the R environment. The model was fitted using the same presence and background training samples as for BRT and RF. Finally, the predictions from each base learner model were compiled and used as training data for fitting the LR model producing the final habitat suitability map for the study area.

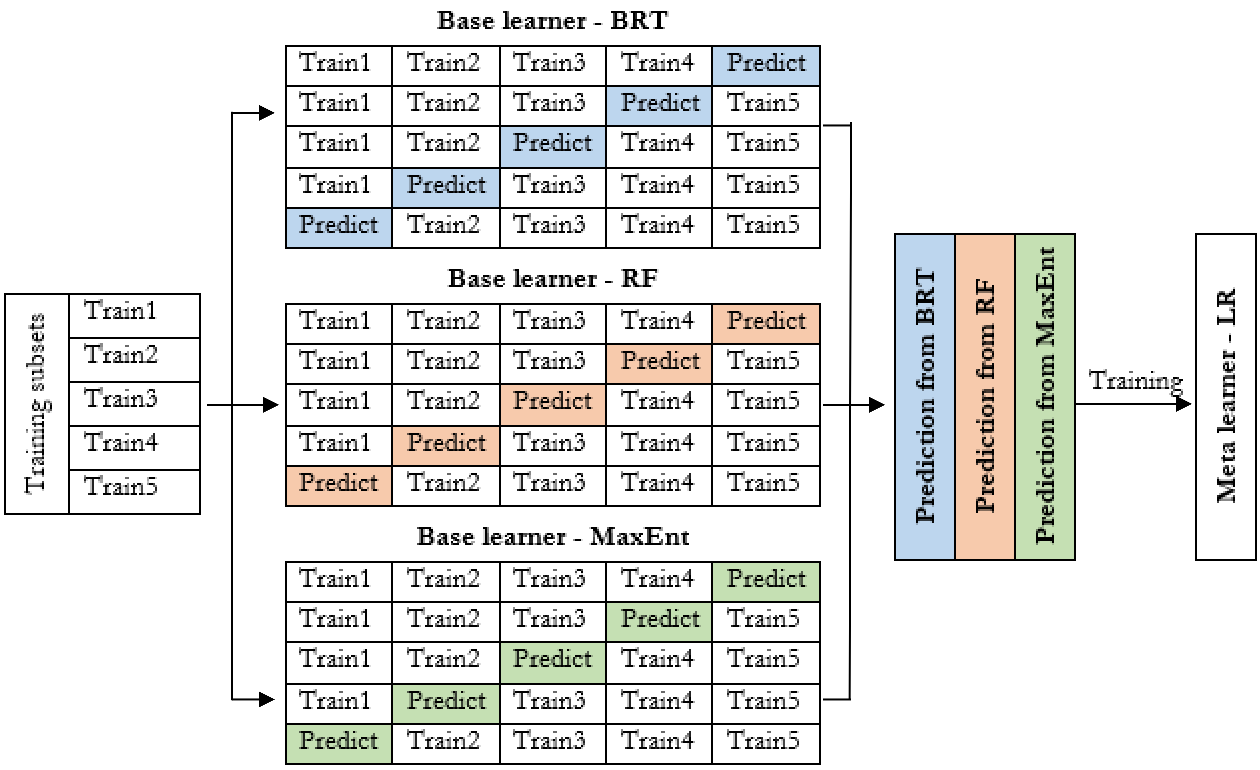


Figure 7: The ensemble stacking training process for elephant habitat suitability modelling

2.4.2. Machine learning models for human-elephant conflict risk prediction

The HEC incidence dataset contained only presence data, and therefore, background or pseudo absence data were generated to create a binary dataset to train the RF model for HEC conflict risk prediction. A total of 5000 background points were generated randomly across the study area representing the range of and extent of predictor variables identified for HEC risk prediction. All the values from the layer of predictor variables were extracted for the given presence and pseudo-absence points and randomly divided into 70% for training and the remaining 30% for testing the model for the performance assessment.

The RF model was fitted using *randomForest* package version 4.7-1 in the R environment. Most samples were down-sampled to address the issue of class imbalance, and overlap and parameters were set to ntree=500 and mtry = 3 for model fitting. The process of RF model fitting was repeated 15 times, and the model with the lowest OOB error rate was chosen for further analysis. The chosen model was used to predict the HEC risk probability across the study area.

2.4.3. Variable importance assessment for elephant habitat suitability modelling

The variable importance assessment was done to understand the key factors influencing the habitat suitability of Asian elephants in the study area. The variable importance output from the MaxEnt gives the estimate of the relative importance of the predictors in the modelling. The Jackknife approach followed in MaxEnt for variable importance excludes one variable at a time during the training process (Baldwin, 2009). The training gain based on the amount of variations on a model with or without the variable is assessed (Bradie and Leung, 2017). This approach provides information on the contribution of each variable and the unique information the variable has to offer in the model, which allows for determining important variables in the model (Baldwin, 2009). In addition to the variable importance information, the response curves from MaxEnt show the logistic predictions change as the value of each predictor variable varies.

2.4.4. Variable importance assessment for human-elephant conflict risk prediction

The variable importance assessment identifies the key factors shaping the spatial distribution of the human-elephant conflict risk. The importance of the variable can be estimated from the permutation importance (Mean Decrease Accuracy) from the RF model. The permutation importance of predictor variable x is estimated by randomly permuting the predictor variable x in the OOB samples in each classifier and subtracting the predictive accuracy of each classifier before and after the permutation of variable x, i.e. with or without the variable (Strobl et al., 2008). The predictor variable x holds high importance if the mean decrease value is high for that variable.

Similarly, the partial dependence plot, one of the output components of the RF model, shows the response of the predicted probability of HEC risk with respect to each predictor variable.

2.4.5. Model performance assessment

Model performance assessment is crucial for any machine learning algorithms and modelling process, as it evaluates the model's predictive capacity (Guisan et al., 2017). The predictive performance of all trained machine learning algorithms for both habitat suitability modelling and human-elephant conflict risk prediction was assessed by the 'area under the curve' (AUC) of the receiver operating characteristics (ROC) function (Elith et al., 2006) and True skill statistics (TSS) (Allouche et al., 2006) using an independent testing dataset. AUC and TSS are commonly used accuracy assessment metrics in species distribution modelling.

The AUC is a threshold independent accuracy metric that assesses the model's capability to discriminate species' presence from absence (Elith et al., 2006; Guisan et al., 2017). The value of AUC lies between 0 and 1, indicating 1 gives the perfect discriminating accuracy, 0.5 meaning the predictive accuracy is no better than random. At the same time, less than 0.5 implies the model's performance is worse than random (Elith et al., 2006). Even though widely used in species distribution modelling, AUC is also criticised for the measure of accuracy (Lobo et al., 2008).

Another model performance assessment metric, TSS, was used to test the accuracy of the model. TSS is not affected by the prevalence and accounts for both omission and commission errors (Allouche et al., 2006). The value of TSS ranges from +1 to -1, where +1 indicates perfect performance and zero or less than zero indicates the performance of classifiers no better than random (Allouche et al., 2006). Because TSS is threshold dependent metric, the maximum sum of sensitivity and specificity threshold value was used as recommended by a previous study (Liu et al., 2013). The *PresenceAbsence* package version 1.1.10 was used in the R environment for model performance assessment.

3. RESULTS

3.1. Elephant habitat suitability modelling

3.1.1. Model performance assessment

The model performance assessment on the testing dataset indicates that all four machine learning models were successful at discriminating species' presence from absence (AUC > 0.87) within the study area (Table 6). The AUC value for all the models did not differ vastly, even though ensemble stacking had the highest AUC (0.90) value, followed by MaxEnt (0.89), RF (0.88) and BRT (0.88).

The calculated TSS value (Table 6) shows that the ensemble stacking (TSS=0.65) depicted good predictive accuracy, although the performance of the RF model was relatively better. However, the ensemble model combines and generalises all the predictions of individual base models.

Table 5: Performance of four different models on elephant habitat suitability modelling, showing the threshold dependent and independent model evaluation results

Accuracy metrics	**Machine learning models**			
	BRT	**RF**	**MaxEnt**	**Ensemble stacking**
AUC	0.88	0.88	0.89	0.90
TSS	0.65	0.66	0.62	0.65

3.1.2. Spatial distribution of suitable habitat for Asian elephant

The threshold that maximised sensitivity plus specificity value (threshold =0.34) was chosen to convert the habitat suitability map produced from the ensemble stacking model into a binary suitable/non-suitable map. The discrete suitable/non-suitable map allowed us to calculate the total suitable area for Asian elephants in the transboundary landscape (Figure 8). The model identified a total of 26,679 km^2 of suitable habitat for the Asian elephant, approximately 36 % of the total area. The overlay of the habitat suitability map with the binary forest map showed that 13,578 km^2 of suitable habitat provides natural forested habitat for the Asian elephant. The predicted suitable habitat is composed of an almost equal proportion of natural forested habitat to non-forest (i.e., closer to 50:50) in the study area. Only 5,366 km^2 of suitable habitat were located inside the fourteen protected areas, which is only 20% of the predicted suitable habitat in the Nepal-India transboundary landscape. The connection of large continuous suitable habitat of the Asian elephant in the Nepal-India transboundary landscape is broken in between, thus divided broadly into two isolated suitable habitats supporting the eastern and western population of elephants.

The suitable habitat map produced from BRT, RF and MaxEnt based on a threshold that maximised sensitivity and specificity are shown in Annex 2, 3 and 4, respectively.

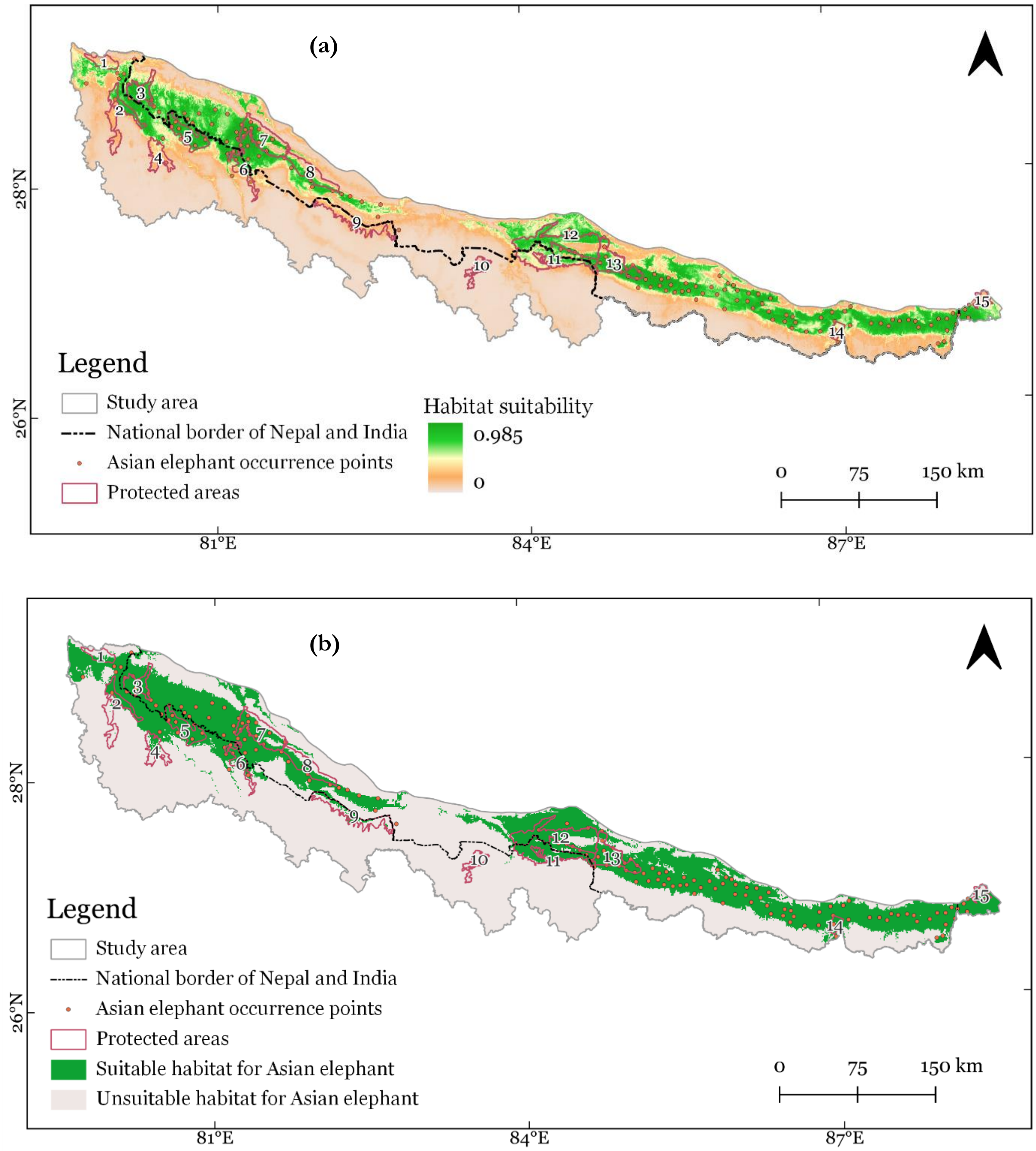


Figure 8: Predicted (a) habitat suitability and (b) suitable habitat for Asian elephant in the Nepal-India transboundary landscape produced from ensemble stacking model. The map shows study area, national border of Nepal and India, Asian elephant occurrence points used in habitat suitability modelling and protected areas. Nine protected areas lie in India: Nandhaur Wildlife Sanctuary(1), Pilibhit Tiger Reserve(2), Kishanpur Wildlife Sanctuary(4), Dudhwa National Park(5), Katarniaghat Wildlife Sanctuary(6), Sohelwa National Park(9), Sohagibarwa Wildlife Sanctuary(10), Valmiki Tiger Reserve(11) and Mahananda Wildlife Sanctuary(15), whereas six PAs lie in Nepal: Suklaphanta National Park(3), Bardiya National Park(7), Banke National Park(8), Chitwan National Park(12), Parsa National Park(13) and Koshi Tappu Wildlife Reserve(14).

The Jackknife test gave information about the contribution and relative importance of predictors variables to generate a MaxEnt model for habitat suitability modelling of Asian elephants in the study area (Figure 9). The variable is considered to have greater predictive ability than others if it offers the highest training gain when used in isolation (with only variable). In addition, a variable contains unique information if the overall training gain is decreased when a variable is omitted (without a variable). The result from Jackknife showed that elevation (elev) followed by precipitation of the driest month (bio14), precipitation of the wettest month (bio13) and temperature of the warmest month (bio5) are the top four variables with the highest contribution to habitat suitability modelling. These four predictor variables together accounted for around 80% contribution to model development, whereas the contribution of elevation alone was around 40%. Elevation showed the highest training gain and contributed substantial unique information alone in model development. Although the maximum temperature of the warmest month (bio5) did not offer significant training gain, it provides significant information that may not be included in other variables, suggesting its substantial contribution to model fitting. Precipitation of the driest month (bio14) offered the second-highest training gain when used in isolation. However, it also showed an overall loss in training gain when it was not included in the model fitting process.

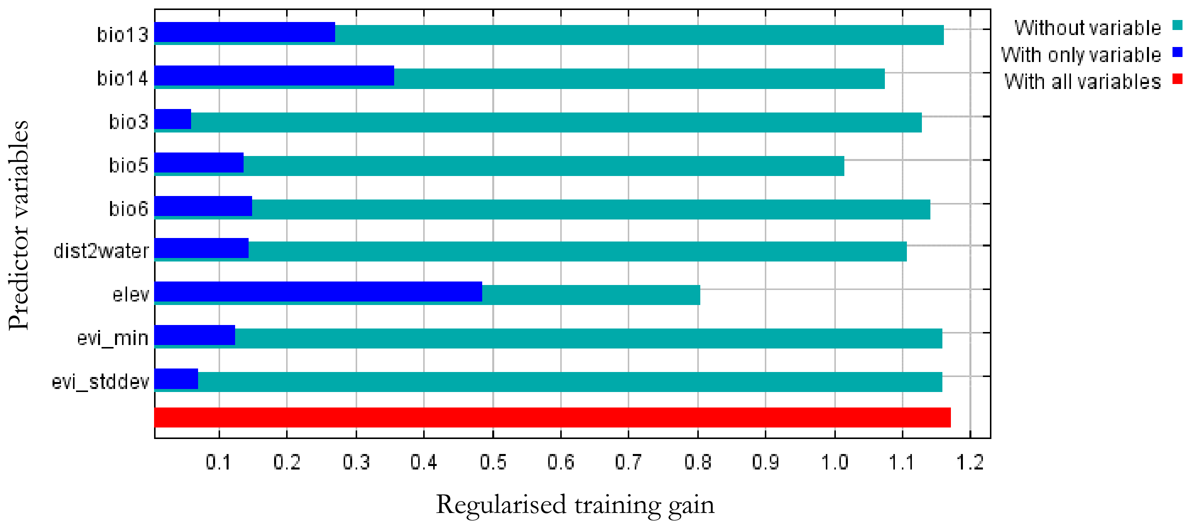


Figure 9: Influence of predictor variables representing climatic and non-climatic factors in modelling habitat suitability of Asian elephant in Nepal-India transboundary landscape. The regularised training gain describes the contribution of each variable "with" and "without" in model fitting process. "With only variable" indicates the result of the training gain when a variable is used in isolation. "With variable" indicates the result of the training gain when a variable is omitted in model fitting. "With all variables" indicates total training gain when all the variables are used in model fitting process.

Response curves are complementary information to the Jackknife test, which facilitates the interpretation of species presence probability and identified predictor variables with the highest contribution and importance in habitat suitability modelling. The response curve shows how habitat suitability logistic prediction changes as each predictor variable varies, keeping all other variables at their average sample check. Only the response curves of the four variables with the highest contribution are shown in Figure 10. According to the response curves, the elephant's occurrence probability was less in both the lowest and highest elevation zone in the study area. This indicates that elephant distribution in the study area is limited to a specific range of elevation, i.e., 100-300 m. The highest habitat suitability responded to the highest precipitation in driest months, lowest precipitation in the wettest month and lowest temperature of the warmest month. This indicates that elephants are most likely to occur in areas with maximum rainfall in the driest month. The scenario changed in the wettest month, where elephants are most likely to favour the region that receives minimum rain (less than 800 mm). The habitat suitability decreases when the temperature of the driest month is beyond 38 °C.

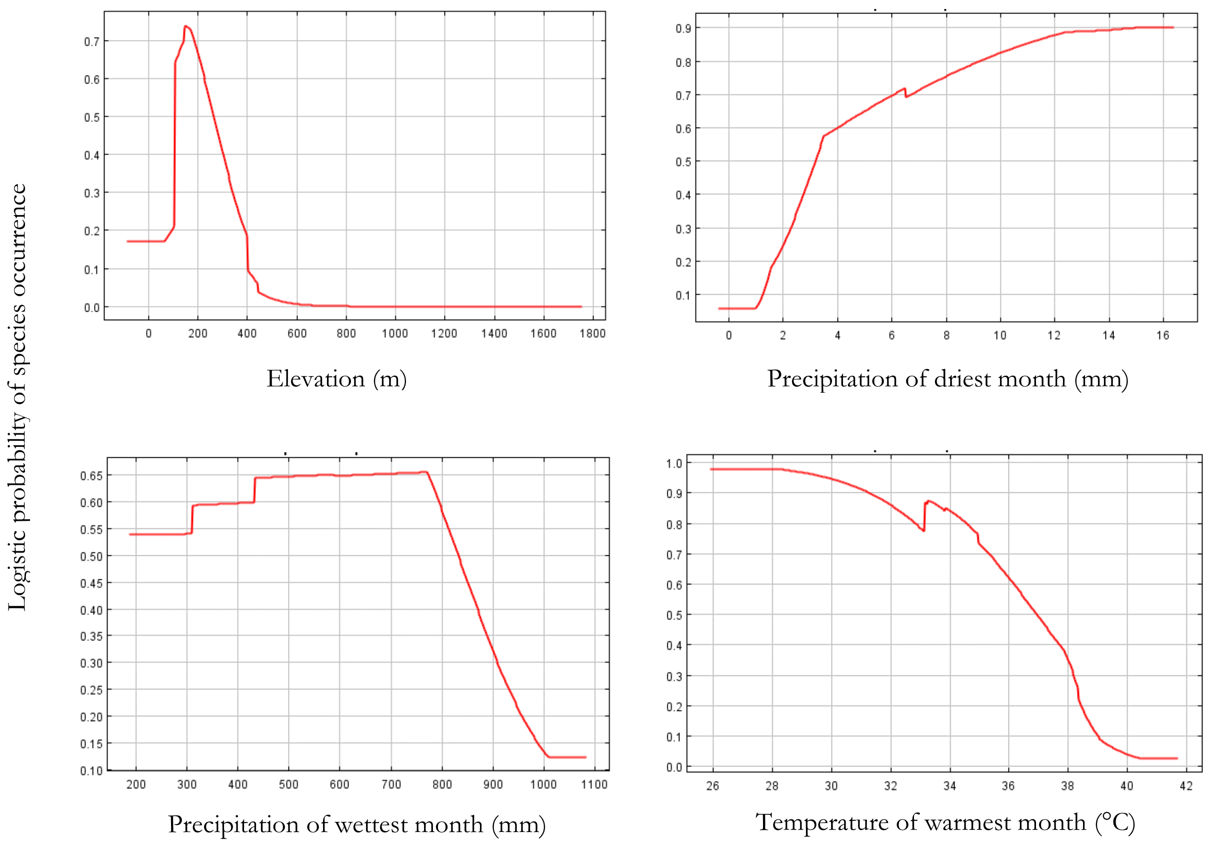


Figure 10: Response curves illustrating the relationship between species occurrence probability and predictor variables used in habitat suitability modelling of Asian elephants in Nepal-India transboundary landscape. The response curve shows how habitat suitability logistic prediction changes as each predictor variable vary, keeping all other predictor variables at their average sample check.

3.2. Human-elephant conflict risk prediction

3.2.1. Model performance assessment

The assessment of the model performance for analysing HEC risk areas was done using two metrics - AUC and TSS. The performance assessment on the testing dataset indicates that the RF model performed well in discriminating HEC occurrence in the study area, with an AUC value of 0.87. Since TSS is the threshold-dependent accuracy assessment metric, the threshold that maximised sensitivity plus specificity value (threshold =0.36) was chosen to calculate TSS. The calculated TSS value was 0.67, indicating good predictive accuracy of the RF model in predicting HEC risk in the Nepal-India transboundary landscape.

3.2.2. Spatial distribution of human-elephant conflict and conflict risk map

The predicted HEC risk map (Figure 11) shows that HEC is concentrated broadly in two regions- eastern and western similar to the distribution of suitable habitats for Asian elephants in the Nepal-India transboundary landscape. The HEC risk map predicted high risk along the eastern and western border between Nepal and India, which extended to Nepal. The transboundary migratory route between Nepal and India on the eastern side observed high conflict risk where elephant migrates from Mahananda WLS in India towards Nepal to connect with the elephant population in Koshi Tappu wildlife reserve and beyond in Nepal. The predicted HEC risk is also high on the migratory route between Nepal and India on the western side of Nepal. On the western side, elephant migrates from Katarniaghat WS and Dhudhwa NP in India towards Bardiya NP in Nepal and from Pilibhit WS in India to Suklaphanta NP in Nepal.

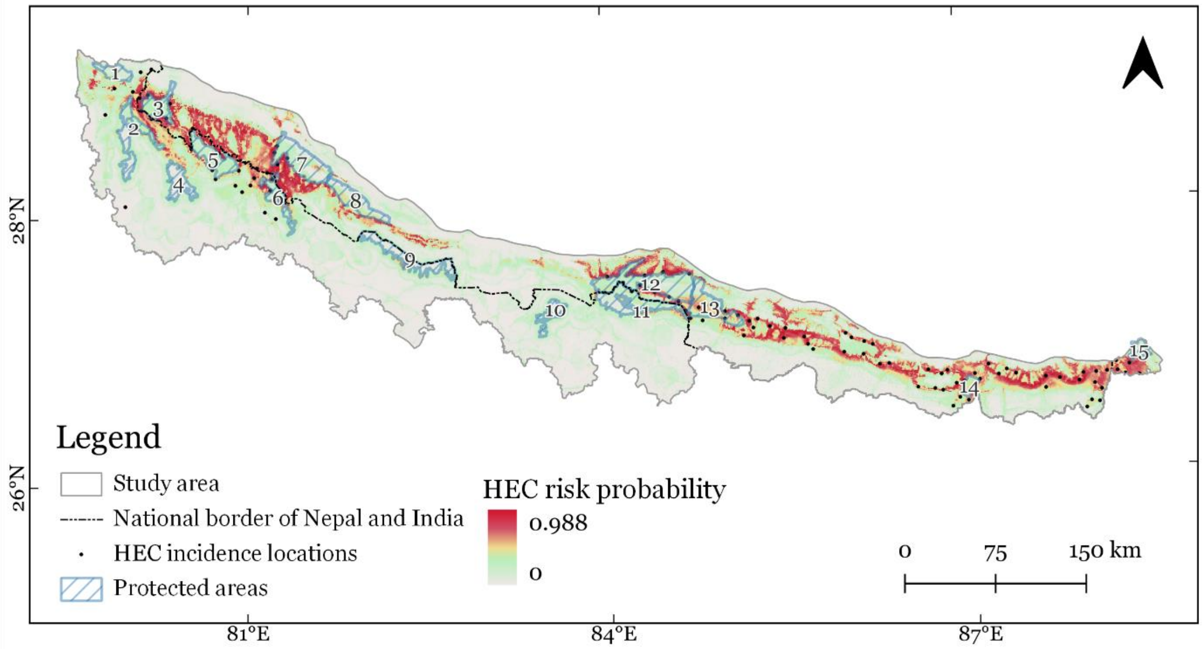


Figure 11: Predicted human-elephant conflict risk probability in Nepal-India transboundary landscape. The probability of HEC risk ranges from 0 to 1. The high value of probability indicates high risk represented by red colour gradient. The map also shows study area, national border of Nepal and India, HEC incidence points used in HEC risk prediction and protected areas. Nine protected areas lie in India: Nandhaur Wildlife Sanctuary(1), Pilibhit Tiger Reserve(2), Kishanpur Wildlife Sanctuary(4), Dudhwa National Park(5), Katarniaghat Wildlife Sanctuary(6), Sohelwa National Park(9), Sohagibarwa Wildlife Sanctuary(10), Valmiki Tiger Reserve(11) and Mahananda Wildlife Sanctuary(15), whereas six PAs lie in Nepal: Suklaphanta National Park(3), Bardiya National Park(7), Banke National Park(8), Chitwan National Park(12), Parsa National Park(13) and Koshi Tappu Wildlife Reserve(14).

3.2.3. Factors determining the risk of human-elephant conflict

The variable importance bar plot based on the "mean decrease accuracy" value from the RF model provides information on the contribution of each variable to the predictive accuracy of a model. The result from the RF model showed that settlement density (settle_den) played a crucial role in predicting HEC conflict risk in Nepal-India transboundary landscape (Figure 12). The decrease in mean accuracy of the model was relatively high for settlement density, which signifies that the variable is most important in predicting HEC risk in the study area. The second significant variable in predicting the HEC risk is the distance to protected areas (dist2PA), followed by habitat suitability (hab_suit) and effective mesh size (mesh).

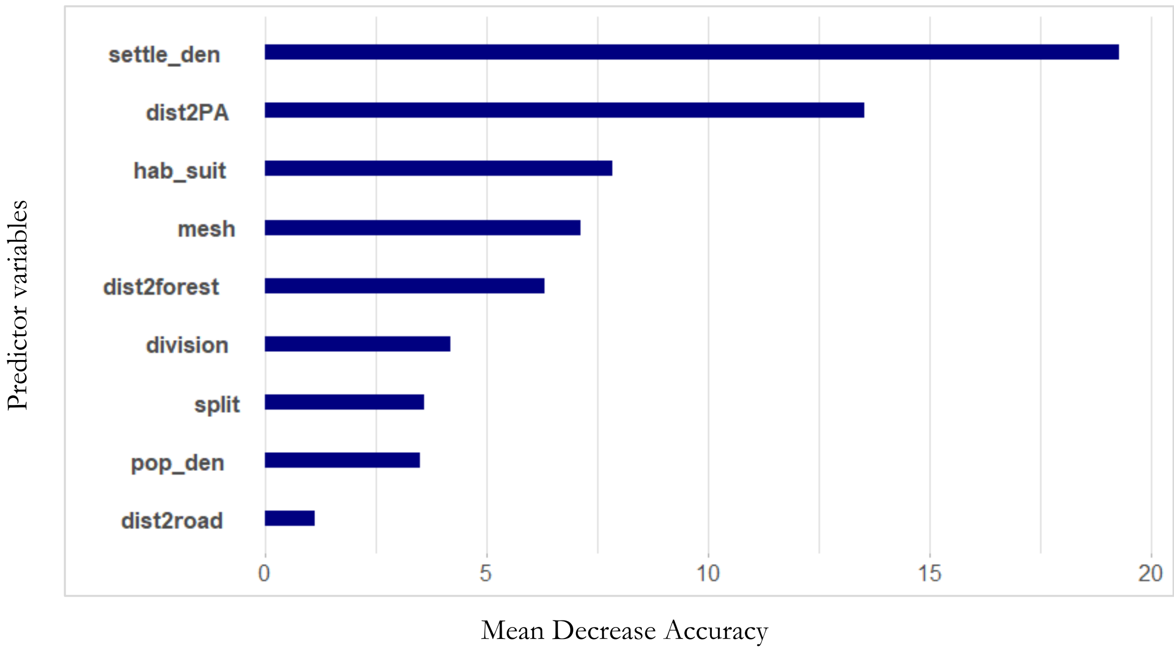


Figure 12: Importance of predictor variables in predictive accuracy of Random Forest model for predicting HEC risk in Nepal-India transboundary landscape. The predictor variable holds high importance if the mean decrease accuracy value is high for that variable.

One of the outputs of RF modelling is the partial dependence plot similar to the response curve output from MaxEnt. The partial dependence plot shows how the logit probability of the class presence of HEC incidence varies with the changing value of predictor variables in HEC risk prediction. Only four partial dependence plots are presented in Figure 13, corresponding to the top four significant contributors to HEC risk prediction. Settlement density had a positive relationship with the likelihood of HEC risk. The probability of conflict risk between humans and elephants increased as the human settlement density increased. The probability of HEC risk was high in the vicinity of protected areas (within ~ 2 km from the protected areas). The village or town near the protected area bore a high chance of experiencing negative interactions with elephants. There was a positive relationship between HEC and habitat suitability. The partial dependence plot of habitat suitability indicated that HEC risk is in the suitable habitat of an elephant. The partial dependence plot of effective mesh size showed a high probability of experiencing HEC risk in the area with a higher effective mesh size value.

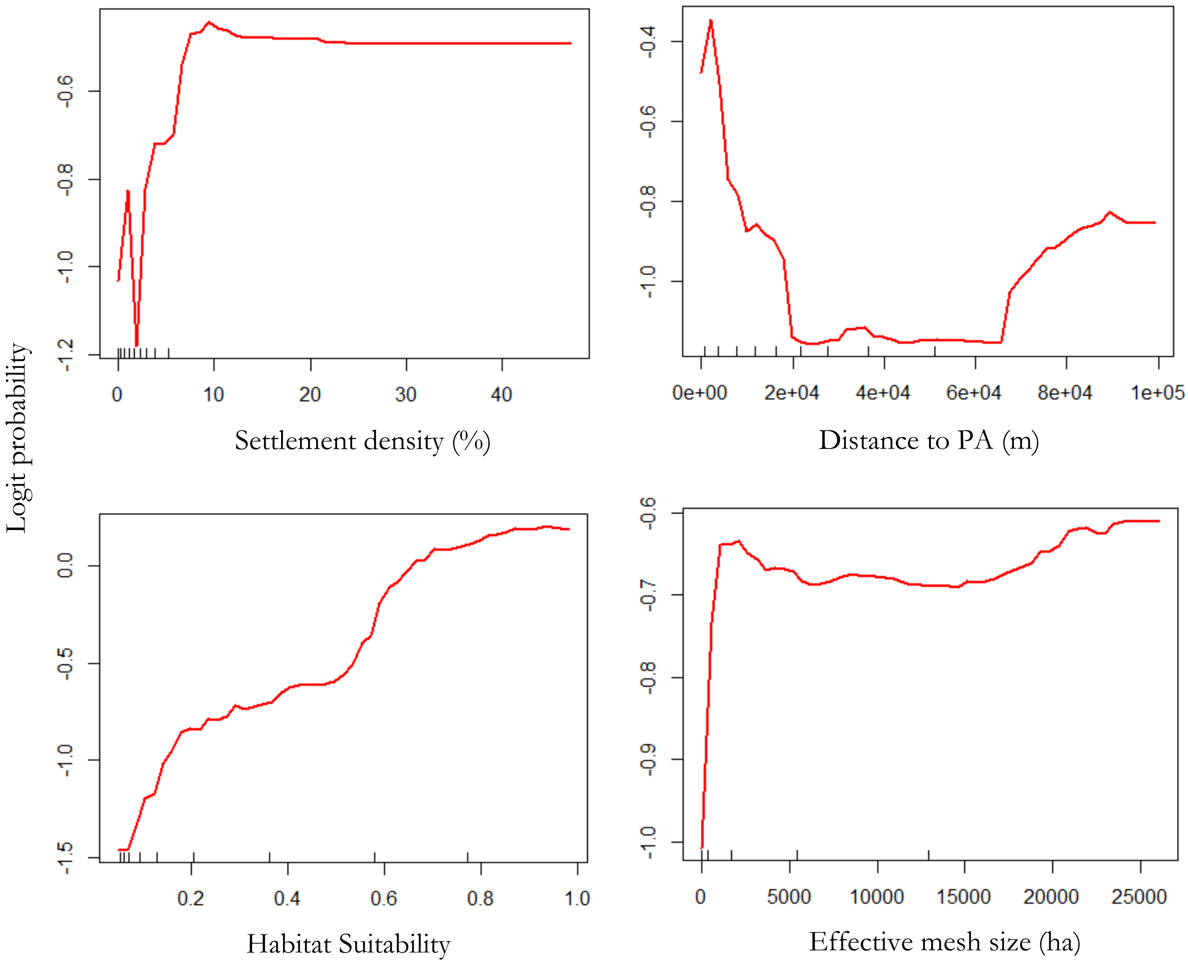


Figure 13: Partial dependence plots illustrate the relationship between probability of human-elephant conflict risk and predictor variables used in Random Forest model. The figure shows partial dependence of only four identified variables that had highest contribution in human-elephant conflict risk prediction in Nepal-India transboundary landscape.

4. DISCUSSION

This study attempted to understand the dynamics of habitat suitability of Asian elephants and human-elephant conflict in the human-dominated transboundary landscape of Nepal and India. The results provide strong evidence that humans are coming in the way of elephants, and the consequence is an increased risk of human-elephant conflict. To my knowledge, this is the first study that integrated the information on habitat suitability to predict human-elephant conflict covering the entire Nepal-India transboundary landscape.

4.1. Suitable habitat for Asian elephant

The study on the habitat suitability of Asian elephants presents the applicability of machine learning-based species distribution models to predict the suitable habitat in the Nepal-India transboundary region. The result of habitat suitability modelling showed that only one-third of the total transboundary landscape area was predicted to be a suitable habitat for Asian elephants. Once distributed throughout the lowland of Nepal in the 1920s (Smith and Mishra, 1992), elephant distribution is now restricted to two geographically distinct suitable habitats. The connectivity of predicted suitable habitats yet remains to be explored. Nonetheless, the forest in the transboundary landscape, which provides the prime natural habitat to elephants, has been lost and fragmented in the last couple of decades, causing significant problems in elephant habitat management. The forest cover within the historical range of Asian elephants decreased by 39.6 % from 1930 to 2013 in India, and forest cover loss counts to 21.5 % in Nepal from 1930 to 2020 (Padalia et al., 2019; Ram et al., 2021b). Albeit the loss of forest cover, elephants have learned to take advantage of non-forest habitats. The predicted suitable habitat is a mix of forest and non-forest (human settlement, farmland, sparse vegetation) in almost equal proportion (ratio closer to 50:50) which suggest a high overlap in space and resource use between elephants and humans. These results resemble with results obtained by other studies on Asian elephant habitat selection and distribution (de la Torre et al., 2021; Kanagaraj et al., 2019; Lamichhane et al., 2017). Even though Nepal and India have been promoting the conservation of migratory species like elephants and their historical habitat in the transboundary landscape, it is surprising that only $1/4^{th}$ of suitable areas are conserved under protected areas as national parks or wildlife sanctuaries in the transboundary landscape. The management of protected areas strictly follows the habitat protection strategy. However, forest outside protected areas is sustaining human pressure. The implication is increased confrontation between elephants and humans which directly reflects on the conservation of elephants and their habitat.

4.2. Factors determining the habitat suitability of Asian elephant

The results of elephant habitat suitability show that the topographic factor, i.e., elevation has the most significant influence in determining the suitable habitat and distribution of elephants in the landscape. The low suitability in the high elevated areas was expected; however, it was surprising to observe low suitability in the low elevated regions. The reason behind the elevation range (neither high nor low) that defines the suitable habitat for Asian elephants may be associated with the distribution of forests in the study area. The preferred elevation range corresponds to the foothills area of Shivalik /Churia hills extending from east to west in the northern part of the study area, which harbours over 75% of the remaining forest of the Terai landscape in Nepal. The flat surface in the lowest elevation area is dominated by agriculture with little or no presence of forest cover (Chanchani et al., 2014). Most forest area is found between 150 to 300 m elevations along the foothills of Shivalik/Churia. The result is similar to the study of the Asian elephant in Peninsular Malaysia, where elephants prefer forested habitats at lower elevations and close to the river (Mohd Taher et al., 2021). Elevation was positively correlated with slope (see Annex 1), which indicates that elephants avoid rugged terrain and steep slopes and favour flat forest plains while migrating from one place to another (de la Torre et al., 2021; ten Velde, 1997). A similar pattern was seen with African elephants in South Africa, where they were found to avoid very flat and very steep slopes, attributed to the availability of nutritious food in an undulating terrain (de Knegt et al., 2011). Low elevation flat land in the study area is also prone to flood in the rainy season, and elephants avoid flooded or flood-prone areas (Kanagaraj et al., 2019).

In addition to the elevation, the precipitation of the driest and wettest months influences the habitat suitability of Asian elephants. The precipitation of the driest month is defined as the lowest cumulative total precipitation during the driest month. In contrast, precipitation of the wettest month is the highest cumulative total precipitation during the wettest month (O'Donnell and Ignizio, 2012). These two bioclimatic variables represent the precipitation extremities. The result suggests that the distribution of elephants is influenced by extreme precipitation conditions, which corroborates with the study (Li et al., 2019) findings on the vast influence of precipitation in the coldest quarter on habitat suitability. The study area exhibits a distinct seasonal pattern, and the dry season limits the forages and water availability in the study area. The seasonal onset and offset of rainfall characterise the availability of different vegetation types and are associated with vegetation productivity. Vegetation productivity is often controlled by interannual variability precipitation, especially dry season precipitation (Murray-Tortarolo et al., 2017). Elephants are known to respond to such seasonal vegetation changes. Elephants occupy riverine vegetation during dry seasons, whereas they move to the tall grass savannah region to take advantage of newly grown grass high in nutrition in the wet season (Sukumar, 1989; ten Velde, 1997). Apart from the influence on vegetation, precipitation affects the water availability in seasonal streams and rivers. Elephant preference for the habitat closer to water resources links to general water requirement for intake and the thermoregulation when the temperature rises (Williams et al., 2008). Elephants are sensitive to high heat and avoid areas with the highest temperature in the warmest months (Kanagaraj et al., 2019). Elephants look for shade in close canopy forests during the daytime (Sitompul et al., 2013), wetting or mud bathing to regulate the ambient heat load when exposed to high environmental temperatures (Mole et al., 2016).

The habitat selection pattern is also affected by the individual behaviour of elephants (de Knegt et al., 2011). Nonetheless, it can be concluded that elephants prefer an area of intermediate elevation with abundant forest, grassland and water availability rather than a flood-prone low elevated area with no forest cover. The movement of large animals like elephants is governed by forage quality than quantity (Bohrer et al., 2014; Owen-Smith, 2014). The topographic and bioclimatic variables influence the habitat suitability of Asian elephants by changing the environment for elephants themselves and changing the habitat in which elephants survive. Depending on the availability of resources like abundant vegetation and fresh biomass in their original habitat, elephants seem to take a risk by travelling longer distances and covering a human-dominated landscape, which might result in frequent conflict with humans.

4.3. Spatial distribution of the risk of human-elephant conflict

The predicted HEC conflict risk map shows that HEC risk is unevenly distributed throughout the Nepal-India transboundary landscape. The risk of HEC is observed within the suitable habitat of Asian elephants. Given the probability of elephant occurrence in the study area, it seems evident that areas experiencing high HEC risk can be divided into two distinct regions: western and eastern.

The Nepal-India transboundary landscape shows the prominent presence of human interference in the composition of suitable remnant habitats for the Asian elephant. Elephants in the transboundary landscape have been migrating from India to Nepal for as long as people can remember. In the landscape where only half of the current predicted suitable habitat is forested, elephants utilise forested and human-dominated landscapes (Lamichhane et al., 2017) while moving through the diverse landscape mosaic. There is historical evidence of clearing forests to establish farmland and human settlement in the study area (Chanchani et al., 2014; Smith and Mishra, 1992). Thus, when elephants move from one protected area to another within or across the country, elephants often come across the cultivated land along the forest fringes and raid crops. This might explain the high probability of HEC risk between and around protected areas, like Baridya NP and Katarniyaghat WS or Mahananda WS and Koshi Tappu WR.

Most people in the study area follow subsistence agriculture and rely on the forest for livestock grazing, forage and fuelwood collection (Chanchani et al., 2014), implying that people like settling near forest areas. Such subsistence farmers establish settlements near water sources where they can easily access water for daily use. Because elephants also prefer areas with available water sources and forest cover, subsistence farmers are likely at high risk of crop-raiding and physical property damage by elephants.

4.4. Factors determining the risk of human-elephant conflict

The result reports that settlement density is the most important factor contributing to HEC. The probability of HEC risk increases in the high settlement as opposed to African elephants in Botswana, where elephants use pathways with low settlement density to capitalise on the opportunity of risk avoidance (Songhurst et al., 2016). However, traditional migratory routes of elephants in the study area landscape are suffering adversely due to human settlement and agricultural expansion (Naha et al., 2019), which brings close contact between humans and elephants. There is a growing problem of expansion of illegal settlements in forests in the study area landscape (Chanchani et al., 2014), which ultimately leads to the permanent conversion of forests into non-forests (Acharya et al., 2011). Nepal and India have witnessed a decrease in forest cover in the last few decades (Padalia et al., 2019; Ram et al., 2021b) for large scale-expansion of agriculture and human settlement (Chanchani et al., 2014). The large continuous patch of forests that once used to be a safe harbour for the passage of elephants are being encroached, and traditional migratory routes of elephants have sustained persistent human disturbances for different land-use activities resulting in habitat loss or fragmentation (Naha et al., 2019; Padalia et al., 2019). Only 20% of the forest in suitable habitat is located in protected areas where wildlife and habitat conservation efforts are practised extensively. The remaining forests are either national or community forests or commercial forest plantations managed with an objective of resource utilisation rather than biodiversity conservation. The national and community forests are under direct pressure from humans and are characterised as areas with low-quality forage availability (Wilson et al., 2013). Such forests are less preferred by elephants (Lamichhane et al., 2017), although they act as a temporary refuge area for the movement of elephants. Elephants are observed to use such forest refuge areas to take shelter and surf around the agricultural field and human settlement areas in proximity to the forest area in search of food resulting in conflict (Fernando et al., 2022).

The probability of experiencing high HEC risk increases in the proximity of protected areas, similar to the finding of other studies (Chen et al., 2016; Fernando et al., 2022; Naha et al., 2019; Sharma et al., 2020a; Tiller et al., 2021; Wilson et al., 2013). The elephant comes out of its forested habitat during the harvesting season of crops like paddy, maize and wheat (Naha et al., 2019; Pant et al., 2016; Ram et al., 2021a) and raids crops on the agricultural field or destroys physical property for harvested grains. Elephants are willing to travel more than 1 km to enter the agricultural field when resources are plentiful (Chen et al., 2016; Wilson et al., 2013). Even though physical barriers like electric fences around protected areas are installed to minimise the crop-raiding, the success of electric fences is highly dependent on regular maintenance and understanding of elephant behaviour (Gunaratne and Premarathne, 2005).

The higher value of mesh means the presence of connected large forest patches and less fragmentation. The study predicts high HEC risk in areas with high mesh value, i.e., in the forest area. This seems unexpected from other research claiming forest fragmentation is causing the escalation of HEC (Chanchani et al., 2014; Fernando et al., 2022; Sharma et al., 2020a). However, the maximum mesh value calculated in the study area (~216 km^2) is less than the average home range assumed in this study. The study result does not contradict the claim that forest fragmentation escalates HEC. The communities with subsistence farming depend on forest resources directly or indirectly and go inside the forest for fuelwood, forage collection, and livestock rearing. The attack on humans by elephants inside the forest happens when elephants encounter people participating in forage collection, livestock rearing, or collecting water (Ram et al., 2021a). Forest loss/fragmentation has a compounding effect on an escalation of HEC by reducing the forest cover required for elephants and increasing the human presence within the suitable habitat of the elephants.

The overall result of the HEC risk prediction in the Nepal-India transboundary region supports the hypothesis that high conflict areas are situated in the suitable habitat of Asian elephants due to the increased human interference. Forest cover loss and fragmentation is causing a frequent confrontation between humans and elephants. The settlements and tea plantations in and around Siliguri in India, settlements in Jhapa and other Terai districts in Nepal are examples of such human interference in the eastern region of the transboundary landscape (connected habitat is Mahananda-KoshiTappu-Chitwan Parsa Complex) (Naha et al., 2019). Similarly, the migratory route and habitat corridors between Katarniaghat/Dudhwa to Bardiya/Suklaphanta complex have been disturbed due to human settlement and agricultural expansion (Chanchani et al., 2014). The conflict risk is high in these areas.

The study investigated HEC by associating the habitat preferences of elephants in the Nepal-India transboundary landscape. However, HEC is also associated with multiple factors like ranging patterns and populations of elephants (Fernando et al., 2008), forage quality available in the forest (Sukumar, 1990), elephant behaviour (Fernando et al., 2022), human tolerance (Neupane et al., 2017) and management practice (Hoare, 2000) which were not considered in the study. Additional investigation on seasonal habitat preferences by elephants will better help to understand the temporal patterns of HEC in the Nepal-India transboundary landscape.

5. CONCLUSION AND RECOMMENDATIONS

This study modelled habitat suitability for Asian elephants in the Nepal-India transboundary landscape. In addition, the study predicted the risk of human-elephant conflict by including habitat suitability, forest fragmentation and other anthropogenic variables. The study further explored the factors driving habitat suitability and human-elephant conflict in the transboundary landscape. The main conclusions that can be drawn from this study are:

a) Approximately one-third of the entire transboundary landscape is predicted to be a suitable habitat for Asian elephants. Only twenty per cent of the predicted suitable habitat is located within the protected areas. The predicted suitable habitat is a mix of forest and non-forest (human settlement, farmland, sparse vegetation) in almost equal proportion (ratio closer to 50:50), suggesting a high overlap in space and resource use between elephants and humans.

b) The habitat suitability of Asian elephants in the Nepal-India transboundary landscape is mainly determined by elevation, precipitation of the driest month and wettest month, and temperature of the warmest month. It suggests that Asian elephants exhibit seasonal preference in habitat use.

c) High human-elephant conflict risk happens in areas of highly suitable elephant habitats suggests strong human interference in the remnant suitable habitats of the Asian elephants in the Nepal-India transboundary landscape.

d) The risk of human-elephant conflict in the Nepal-India transboundary landscape is mainly determined by human settlement density, distance to protected areas, elephant habitat suitability, and forest fragmentation. The human settlements established along the migratory routes and agricultural fields enveloping the protected area are prone to crop-raiding and intense HEC.

The suitable elephant habitat and the human-elephant risk areas identified by this study are important, which could serve as a basis for strengthening the transboundary conservation of elephants and strategies for mitigating man-elephant conflicts. The study revealed that only $1/4^{th}$ of suitable habitat is under legal protection as protected areas. The fragmented forest outside the protected area classified as suitable habitat by this study needs to be protected. The maintenance of ecological corridors that connect fragmented habitats and PAs is necessary to facilitate the migration of elephants. However, in doing so, it is crucial to understand and address how people living in between fragmented habitats and PAs make decisions about current and future resource use (Shaffer et al., 2019). Given that human-elephant conflict in the transboundary landscape is centred around overlapping space and competition for resource use, land use planning like slum resettlement, expansion of farmland and development of infrastructures like airports, irrigation canals or road constructions should be envisaged in a way that does not degrade the elephant's habitat and the migratory route. The study on seasonal habitat preference by Asian elephants might give additional information about temporal patterns of habitat use by elephants which might assist in implementing HEC management effectively.

The study recommends that the transboundary conservation efforts need to be strengthened through regional cooperation and building common strategies. The cross-border cooperation programme like the Terai arc landscape and Kanchenjunga landscape are appreciative efforts to ensure wildlife conservation and promote human well-being. Such collaboration ensures combined synergies and capitalises on opportunities to pave the way for sustainable use of natural resources. However, more study at the transboundary landscape level is required to understand the interplay of ecological, anthropogenic, social, and political-economic scenarios in an escalation of human-elephant conflict. This information will benefit the long-term conservation of Asian elephants and promote human welfare in the Nepal-India transboundary landscape.

LIST OF REFERENCES

Acharya, K.P., Dangi, R.B., Acharya, M., 2011. Understanding forest degradation in Nepal. Unasylva 62, 31–38.

Allouche, O., Tsoar, A., Kadmon, R., 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). J. Appl. Ecol. 43, 1223–1232. https://doi.org/10.1111/j.1365-2664.2006.01214.x

Araujo, M., New, M., 2007. Ensemble forecasting of species distributions. Trends Ecol. Evol. 22, 42–47. https://doi.org/10.1016/j.tree.2006.09.010

Araújo, M.B., Guisan, A., 2006. Five (or so) challenges for species distribution modelling. J. Biogeogr. 33, 1677–1688. https://doi.org/10.1111/j.1365-2699.2006.01584.x

Bai, D.F., Chen, P.J., Atzeni, L., Cering, L., Li, Q., Shi, K., 2018. Assessment of habitat suitability of the Snow leopard (*Panthera uncia*) in Qomolangma National Nature Reserve based on MaxEnt modeling. Zool. Res. 39, 373–386. https://doi.org/10.24272/j.issn.2095-8137.2018.057

Baldwin, R.A., 2009. Use of maximum entropy modeling in wildlife research. Entropy 11, 854–866. https://doi.org/10.3390/e11040854

Bar, S., Parida, B.R., Pandey, A.C., 2020. Landsat-8 and Sentinel-2 based forest fire burn area mapping using machine learning algorithms on GEE cloud platform over Uttarakhand, Western Himalaya. Remote Sens. Appl. Soc. Environ. 18, 100324. https://doi.org/10.1016/j.rsase.2020.100324

Barbet-Massin, M., Jiguet, F., Albert, C.H., Thuiller, W., 2012. Selecting pseudo-absences for species distribution models: how, where and how many? Methods Ecol. Evol. 3, 327–338. https://doi.org/10.1111/j.2041-210X.2011.00172.x

Barua, M., Bhagwat, S.A., Jadhav, S., 2013. The hidden dimensions of human-wildlife conflict: health impacts, opportunity and transaction costs. Biol. Conserv. 157, 309–316. https://doi.org/10.1016/j.biocon.2012.07.014

Baskaran, N., Kannan, G., Anbarasan, U., Thapa, A., Sukumar, R., 2013. A landscape-level assessment of Asian elephant habitat, its population and elephant–human conflict in the Anamalai hill ranges of southern Western Ghats, India. Mamm. Biol. 78, 470–481. https://doi.org/10.1016/j.mambio.2013.04.007

Belgiu, M., Drăgu, L., 2016. Random forest in remote sensing: a review of applications and future directions. ISPRS J. Photogramm. Remote Sens. 114, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Bohrer, G., Beck, P.S.A., Ngene, S.M., Skidmore, A.K., Douglas-Hamilton, I., 2014. Elephant movement closely tracks precipitation-driven vegetation dynamics in a Kenyan forest-savanna landscape. Mov. Ecol. 2, 1–12. https://doi.org/10.1186/2051-3933-2-2

Boogar, A.R., Salehi, H., Pourghasemi, H.R., Blaschke, T., 2019. Predicting habitat suitability and conserving *Juniperus* spp. habitat using SVM and Maximum entropy machine learning techniques. Water 11, 2049. https://doi.org/10.3390/w11102049

Bradie, J., Leung, B., 2017. A quantitative synthesis of the importance of variables used in MaxEnt species distribution models. J. Biogeogr. 44, 1344–1361. https://doi.org/10.1111/jbi.12894

Breiman, L., 2001. Random Forests, in: Machine Learning. pp. 5–32.

Buijs, A., Jacobs, M., 2021. Avoiding negativity bias: towards a positive psychology of human–wildlife relationships. Ambio 50, 281–288. https://doi.org/10.1007/s13280-020-01394-w

Calabrese, A., Calabrese, J.M., Songer, M., Wegmann, M., Hedges, S., Rose, R., Leimgruber, P., 2017. Conservation status of Asian elephants: the influence of habitat and governance. Biodivers. Conserv. 26, 2067–2081. https://doi.org/10.1007/s10531-017-1345-5

Chanchani, P., Lamichhane, B.R., Malla, S., Maurya, K., Busta, A., Warruer, R., Nair, S., Almedia, 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., Borah, J., 2014. Tigers of the transboundary Terai Arc Landscape: status, distribution and movement in the Terai of India and Nepal. Natl. Tiger Conserv. Authority, Gov. India, Dep. Natl. Park Wildl. Conserv. Gov. Nepal 1–98.

Chen, Ying, Marino, J., Chen, Yong, Tao, Q., Sullivan, C.D., Shi, K., Macdonald, D.W., 2016. Predicting hotspots of human-elephant conflict to inform mitigation strategies in Xishuangbanna, Southwest China. PLoS One 11, e0162035. https://doi.org/10.1371/journal.pone.0162035

Colin, B., Clifford, S., Wu, P., Rathmanner, S., Mengersen, K., 2017. Using boosted regression trees and remotely sensed data to drive decision-making. Open J. Stat. 07, 859–875. https://doi.org/10.4236/ojs.2017.75061

Corbane, C., Syrris, V., Sabo, F., Politis, P., Melchiorri, M., Pesaresi, M., Soille, P., Kemper, T., 2021. Convolutional neural networks for global human settlements mapping from Sentinel-2 satellite imagery. Neural Comput. Appl. 33, 6697–6720. https://doi.org/10.1007/S00521-020-05449-7

de Knegt, H.J., van Langevelde, F., Skidmore, A.K., Delsink, A., Slotow, R., Henley, S., Bucini, G., de Boer, W.F., Coughenour, M.B., Grant, C.C., Heitkönig, I.M.A., Henley, M., Knox, N.M., Kohi, E.M., Mwakiwa, E., Page, B.R., Peel, M., Pretorius, Y., van Wieren, S.E., Prins, H.H.T., 2011. The spatial scaling of habitat selection by African elephants. J. Anim. Ecol. 80, 270–281. https://doi.org/10.1111/j.1365-2656.2010.01764.x

de la Torre, J.A., Wong, E.P., Lechner, A.M., Zulaikha, N., Zawawi, A., Abdul-Patah, P., Saaban, S., Goossens, B., Campos-Arceiz, A., 2021. There will be conflict – agricultural landscapes are prime, rather than marginal, habitats for Asian elephants. Anim. Conserv. 1–13. https://doi.org/10.1111/acv.12668

De Marco, P., Nóbrega, C.C., 2018. Evaluating collinearity effects on species distribution models: an approach based on virtual species simulation. PLoS One 13, e0202403. https://doi.org/10.1371/journal.pone.0202403

De, U.S., Dube, R.K., Rao, G.S.P., 2005. Extreme weather events over India in the last 100 years. J. Indian Geophys. Union 9, 173–187.

Dickman, A.J., 2010. Complexities of conflict: the importance of considering social factors for effectively resolving human-wildlife conflict. Anim. Conserv. 13, 458–466. https://doi.org/10.1111/j.1469-1795.2010.00368.x

Distefano, E., 2005. Human-wildlife conflict worldwide: collection of case studies, analysis of management strategies and good practices. SARD Initiative Report, FAO, Rome.

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., Mcclean, C., Osborne, P.E., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D., Lautenbach, S., 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography (Cop.). 36, 27–46. https://doi.org/10.1111/j.1600-0587.2012.07348.x

Eisenberg, J.F., 1980. Ecology and behavior of the Asian elephant. Elephant 1, 36–56. https://doi.org/10.22237/elephant/1521731835

Elith, J., Graham, C.H., 2009. Do they? How do they? WHY do they differ? on finding reasons for differing performances of species distribution models. Ecography (Cop.). 32, 66–77. https://doi.org/10.1111/j.1600-0587.2008.05505.x

Elith, J., H. Graham, C., P. Anderson, R., Dudík, M., Ferrier, S., Guisan, A., J. Hijmans, R., Huettmann, F., R. Leathwick, J., Lehmann, A., Li, J., G. Lohmann, L., A. Loiselle, B., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., McC. M. Overton, J., Townsend Peterson, A., J. Phillips, S., Richardson, K., Scachetti-Pereira, R., E. Schapire, R., Soberón, J., Williams, S., S. Wisz, M., E.

Zimmermann, N., 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography (Cop.). 29, 129–151. https://doi.org/10.1111/j.2006.0906-7590.04596.x

Elith, J., Leathwick, J.R., Hastie, T., 2008. A working guide to boosted regression trees. J. Anim. Ecol. 77, 802–813. https://doi.org/10.1111/j.1365-2656.2008.01390.x

FAO, 2021. Foresty division - wildlife and protected areas management [WWW Document]. Food Agric. Organ. United Nations Rome. URL https://www.fao.org/3/cb2447en/cb2447en.pdf (accessed 6.23.21).

Feng, X., Park, D.S., Liang, Y., Pandey, R., Papeş, M., 2019. Collinearity in ecological niche modeling: confusions and challenges. Ecol. Evol. 9, 10365–10376. https://doi.org/10.1002/ece3.5555

Fernando, C., Weston, M.A., Corea, R., Pahirana, K., Rendall, A.R., 2022. Asian elephant movements between natural and human-dominated landscapes mirror patterns of crop damage in Sri Lanka. Oryx 1–8. https://doi.org/10.1017/S0030605321000971

Fernando, P., Wikramanayake, E.D., Janaka, H.K., Jayasinghe, L.K.A., Gunawardena, M., Kotagama, S.W., Weerakoon, D., Pastorini, J., 2008. Ranging behavior of the Asian elephant in Sri Lanka. Mamm. Biol. 73, 2–13. https://doi.org/10.1016/j.mambio.2007.07.007

Fick, S.E., Hijmans, R.J., 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. Int. J. Climatol. 37, 4302–4315. https://doi.org/10.1002/joc.5086

Fletcher, R., Toncheva, S., 2021. The political economy of human-wildlife conflict and coexistence. Biol. Conserv. 260, 109216. https://doi.org/10.1016/j.biocon.2021.109216

Franklin, J., 2012. Species distribution modeling, in: Mapping Species Distributions: Spatial Inference and Prediction. Cambridge University Press, Cambridge, pp. 3–20. https://doi.org/10.1017/cbo9780511810602.003

Genuer, R., Poggi, J.-M., Tuleau-malot, C., 2010. Variable selection using Random forests. Pattern Recognit. Lett. Elsevier 31, 2225–2236.

Gislason, P.O., Benediktsson, J.A., Sveinsson, J.R., 2006. Random forests for land cover classification. Pattern Recognit. Lett. 27, 294–300. https://doi.org/10.1016/j.patrec.2005.08.011

Gobeyn, S., Mouton, A.M., Cord, A.F., Kaim, A., Volk, M., Goethals, P.L.M., 2019. Evolutionary algorithms for species distribution modelling: a review in the context of machine learning. Ecol. Modell. https://doi.org/10.1016/j.ecolmodel.2018.11.013

Gross, E., Jayasinghe, N., Brooks, A., Polet, G., Wadhwa, R., Hilderink-Koopmans, F., 2021. A future for all: the need for human-wildlife coexistence. WWF, Gland, Switzerland.

Guisan, A., Thuiller, W., 2005. Predicting species distribution: offering more than simple habitat models. Ecol. Lett. 8, 993–1009. https://doi.org/10.1111/j.1461-0248.2005.00792.x

Guisan, A., Thuiller, W., Zimmermann, N.E., 2017. Habitat suitability and distribution models: with applications in R, Cambridge University Press. https://doi.org/10.1017/ 9781139028271

Gunaratne, L.H.P., Premarathne, P.K., 2005. The effectiveness of electric fencing in mitigating human-elephant conflict in Sri Lanka. EEPSEA, IDRC Regional Office for Southeast and East Asia, Sigapore, SG.

Hoare, R., 2000. African elephants and humans in conflict: the outlook for co-existence. Oryx 34, 34–38. https://doi.org/10.1046/j.1365-3008.2000.00092.x

Hu, X., Zhang, H., Mei, H., Xiao, D., Li, Y., Li, M., 2020. Landslide susceptibility mapping using the stacking ensemble machine learning method in Lushui, Southwest China. Appl. Sci. 10. https://doi.org/10.3390/app10114016

Huang, C., Li, X., Khanal, L., Jiang, X., 2019. Habitat suitability and connectivity inform a co-management policy of protected area network for asian elephants in China. PeerJ 7, e6791.

https://doi.org/10.7717/peerj.6791

Ibáñez-Justicia, A., Alcaraz-Hernández, J.D., Van Lammeren, R., Koenraadt, C.J.M., Bergsma, A., Delucchi, L., Rizzoli, A., Takken, W., 2020. Habitat suitability modelling to assess the introductions of *Aedes albopictus* (Diptera: Culicidae) in the Netherlands. Parasites and Vectors 13, 1–13. https://doi.org/10.1186/s13071-020-04077-3

Jaeger, J.A.G., 2000. Landscape division, splitting index, and effective mesh size: new measures of landscape fragmentation. Landsc. Ecol. 15, 115–130. https://doi.org/10.1023/A:1008129329289

Kanagaraj, R., Araujo, M.B., Barman, R., Davidar, P., De, R., Digal, D.K., Gopi, G. V., Johnsingh, A.J.T., Kakati, K., Kramer-Schadt, S., Lamichhane, B.R., Lyngdoh, S., Madhusudan, M.D., Ul Islam Najar, M., Parida, J., Pradhan, N.M.B., Puyravaud, J., Raghunath, R., Rahim, P.P.A., Muthamizh Selvan, K., Subedi, N., Trabucco, A., Udayraj, S., Wiegand, T., Williams, A.C., Goyal, S.P., 2019. Predicting range shifts of Asian elephants under global change. Divers. Distrib. 25, 822–838. https://doi.org/10.1111/ddi.12898

Kitratporn, N., Takeuchi, W., 2019. Spatiotemporal distribution of human–elephant conflict in Eastern Thailand: a model-based assessment using news reports and remotely sensed data. Remote Sens. 12, 90. https://doi.org/10.3390/rs12010090

König, H.J., Kiffner, C., Kramer-Schadt, S., Fürst, C., Keuling, O., Ford, A.T., 2020. Human–wildlife coexistence in a changing world. Conserv. Biol. 34, 786–794. https://doi.org/10.1111/cobi.13513

Krishnan, V., Kumar, M.A., Raghunathan, G., Vijayakrishnan, S., 2019. Distribution and habitat use by Asian elephants (*Elephas maximus*) in a coffee-dominated landscape of Southern India. Trop. Conserv. Sci. 12, 1–12. https://doi.org/10.1177/1940082918822599

Lamarque, F., Anderson, J., Fergusson, R., Lagrange, M., Osei-Owusu, Y., Bakker, L., 2009. Human-wildlife conflict in Africa: causes, consequences and management strategies. Food and Agriculture Organization of the United Nations, Rome.

Lamichhane, B.R., Subedi, N., Pokheral, C.P., Dhakal, M., Acharya, K.P., Pradhan, N.M.B., Smith, J.L.D., Malla, S., Thakuri, B.S., Yackulic, C.B., 2017. Using interviews and biological sign surveys to infer seasonal use of forested and agricultural portions of a human-dominated landscape by Asian elephants in Nepal. Ethol. Ecol. Evol. 30, 6112–6119. https://doi.org/10.1080/03949370.2017.1405847

Leslie, S., Brooks, A., Jayasinghe, N., Koopmans, F., 2019. Human wildlife conflict mitigation: lessons learned from global compensation and insurance schemes. HWC SAFE Series. WWF Tigers Alive.

Li, W., Yu, Y., Liu, P., Tang, R., Dai, Y., Li, L., Zhang, L., 2019. Identifying climate refugia and its potential impact on small population of Asian elephant (*Elephas maximus*) in China. Glob. Ecol. Conserv. 19, e00664. https://doi.org/10.1016/j.gecco.2019.e00664

Liu, C., White, M., Newell, G., 2013. Selecting thresholds for the prediction of species occurrence with presence-only data. J. Biogeogr. 40, 778–789. https://doi.org/10.1111/jbi.12058

Liu, J., Yong, D.L., Choi, C.-Y., Gibson, L., 2020. Transboundary frontiers: an emerging priority for biodiversity conservation. Trends Ecol. Evol. 35, 679–690. https://doi.org/https://doi.org/10.1016/j.tree.2020.03.004

Lobo, J.M., Jiménez-Valverde, A., Real, R., 2008. AUC: a misleading measure of the performance of predictive distribution models. Glob. Ecol. Biogeogr. 17, 145–151. https://doi.org/10.1111/j.1466-8238.2007.00358.x

Marcello, Schiavina, Sergio, K.F., MacManus, 2019. GHS population grid multitemporal (1975, 1990, 2000, 2015) R2019A. European Commission, Joint Research Centre (JRC).

Mason, N., Ward, M., Watson, J.E.M., Venter, O., Runting, R.K., 2020. Global opportunities and challenges for transboundary conservation. Nat. Ecol. Evol. 4, 694–701. https://doi.org/10.1038/s41559-020-1160-3

Mcgarigal, K., 2015. FRAGSTATS help. University of Massachusetts: Amherst, MA, USA.

Menon, V., Tiwari, S.K.R., 2019. Population status of Asian elephants *Elephas maximus* and key threats. Int. Zoo Yearb. 53, 17–30. https://doi.org/10.1111/izy.12247

Ministry of Forests and Soil Conservation, 2016. Conservation landscapes of Nepal. Ministry of Forests and Soil Conservation, Singha Durbar, Kathmandu, Nepal.

Mohd Taher, T., Lihan, T., Tajul Arifin, N.A., Khodri, N.F., Ahmad Mustapha, M., Abdul Patah, P., Razali, S.H.A., Mohd Nor, S., 2021. Characteristic of habitat suitability for the Asian elephant in the fragmented Ulu Jelai Forest Reserve, Peninsular Malaysia. Trop. Ecol. 62, 347–358. https://doi.org/10.1007/s42965-021-00154-5

Mole, M.A., Rodrigues Dáraujo, S., Van Aarde, R.J., Mitchell, D., Fuller, A., Cooke, S., 2016. Coping with heat: behavioural and physiological responses of Savanna elephants in their natural habitat 4. https://doi.org/10.1093/conphys/cow044

Mpakairi, K., Ndaimani, H., Vingi, K., Madiri, T.H., Nekatambe, T., 2018. Ensemble modelling predicts human carnivore conflict for a community adjacent to a protected area in Zimbabwe. Afr. J. Ecol. 56, 957–963. https://doi.org/10.1111/aje.12526

Murray-Tortarolo, G., Jaramillo, V.J., Maass, M., Friedlingstein, P., Sitch, S., 2017. The decreasing range between dry- and wet-season precipitation over land and its effect on vegetation primary productivity. PLoS One 12, e0190304. https://doi.org/10.1371/journal.pone.0190304

Naha, D., Sathyakumar, S., Dash, S., Chettri, A., Rawat, G.S., 2019. Assessment and prediction of spatial patterns of human-elephant conflicts in changing land cover scenarios of a human-dominated landscape in North Bengal. PLoS One 14, 1–19. https://doi.org/10.1371/journal.pone.0210580

Naughton, L., Rose, R.A., Treves, A., 1999. The social dimensions of human-elephant conflict in Africa: a literature review and two case studies from Uganda and Cameroon. A report submitted to the African Elephant Specialist Group, Human-Elephant Task Conflict Task Force, IUCN, Gland, Switzerland.

Neupane, D., Johnson, R.L., Risch, T.S., 2017. How do land-use practices affect human-elephant conflict in Nepal? Wildlife Biol. 2017, 1–9. https://doi.org/10.2981/wlb.00313

Nyhus, P.J., 2016. Human–wildlife conflict and coexistence. Annu. Rev. Environ. Resour. 41, 143–171. https://doi.org/10.1146/annurev-environ-110615-085634

O'Donnell, M.S., Ignizio, D.A., 2012. Bioclimatic predictors for supporting ecological applications in the conterminous United States. U.S Geol. Surv. data Ser. 691.

Olden, J.D., Lawler, J.J., Poff, N.L., 2008. Machine learning methods without tears: a primer for ecologists. Q. Rev. Biol. 83, 171–193. https://doi.org/10.1086/587826

Olivier, R., 1978. Distribution and status of the Asian elephant. Oryx 14, 379–424. https://doi.org/10.1017/S003060530001601X

Onlinekhabar, 2021. Human-elephant conflict in Nepal: 7 things you should know - OnlineKhabar English News [WWW Document]. URL https://english.onlinekhabar.com/human-elephant-conflict-in-nepal-7-things-you-should-know.html (accessed 10.24.21).

Owen-Smith, N., 2014. Spatial ecology of large herbivore populations. Ecography (Cop.). 37, 416–430. https://doi.org/10.1111/j.1600-0587.2013.00613.x

Padalia, H., Ghosh, S., Reddy, C.S., Nandy, S., Singh, S., Kumar, A.S., 2019. Assessment of historical forest cover loss and fragmentation in Asian elephant ranges in India. Environ. Monit. Assess. 191, 1–13. https://doi.org/10.1007/s10661-019-7696-5

Pant, G., Dhakal, M., Pradhan, N.M.B., Leverington, F., Hockings, M., 2016. Nature and extent of human-elephant Elephas maximus conflict in Central Nepal. ORYX. https://doi.org/10.1017/S0030605315000381

Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. Ecol. Modell. 190, 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Phillips, S.J., Dudík, M., E. Schapire, R., 2004. A maximum entropy approach to species distribution modeling, in: Proceedings of the Twenty-First International Conference on Machine Learning. pp. 655–662.

Plumptre, A.J., Kujirakwinja, D., Treves, A., Owiunji, I., Rainer, H., 2007. Transboundary conservation in the greater Virunga landscape: its importance for landscape species. Biol. Conserv. 134, 279–287. https://doi.org/10.1016/j.biocon.2006.08.012

Pradhan, N.M.B., Williams, A.C., Dhakal, M., 2011. Current status of Asian elephants in Nepal. Gajah 35, 87–92.

Prakash, T.G.S.L., Wijeratne, A.W., Fernando, P., 2020. Human-elephant conflict in Sri Lanka: patterns and extent. Gajah 51, 16–25.

Probst, P., Wright, M.N., Boulesteix, A.L., 2019. Hyperparameters and tuning strategies for random forest. Wiley Interdiscip. Rev. Data Min. Knowl. Discov. 9, 1–15. https://doi.org/10.1002/widm.1301

Ram, A.K., 2021. Landscape level modeling of Asian Elephant (*Elephas maximus*) habitat, its population and interaction with humans in Nepal (Doctoral dissertation). Wildlife Institute of India, Dehradun, India and Saurashtra University, Rajkot, India.

Ram, A.K., Mondol, S., Subedi, N., Lamichhane, B.R., Baral, H.S., Natarajan, L., Amin, R., Pandav, B., 2021a. Patterns and determinants of elephant attacks on humans in Nepal. Ecol. Evol. 11, 11639–11650. https://doi.org/10.1002/ece3.7796

Ram, A.K., Yadav, N.K., Kandel, P.N., Mondol, S., Pandav, B., Natarajan, L., Subedi, N., Naha, D., Reddy, C.S., Lamichhane, B.R., 2021b. Tracking forest loss and fragmentation between 1930 and 2020 in Asian elephant (*Elephas maximus*) range in Nepal. Sci. Rep. 11, 1–13. https://doi.org/10.1038/s41598-021-98327-8

Ram, A.K., Yadav, N.K., Subedi, N., Pandav, B., Mondol, S., Khanal, B., Kharal, D.K., Acharya, H.B., Dhakal, B.K., Acharya, K.P., Baral, H.S., Dahal, B.R., Mishra, R., Naha, D., Pradhan, N.M.B., Natarajan, L., Lamihhane, B.R., 2022. Landscape predictors of human elephant conflicts in Chure Terai Madhesh landscape of Nepal. Environ. Challenges 7, 100458. https://doi.org/10.1016/j.envc.2022.100458

Sandwith, T., Shine, C., Hamilton, L., Sheppard, D., 2001. Transboundary protected areas for peace and co-operation. IUCN, Gland, Switzerland and Canbridge, UK.

Secretariat of the Convention in Biological Diversity, 2016. Biodiversity and the 2030 agenda for sustainable development. Technical Note.

Sewell, M., 2008. Ensemble learning. RN 11, 1–34.

Shaffer, L.J., Khadka, K.K., Van Den Hoek, J., Naithani, K.J., 2019. Human-elephant conflict: a review of current management strategies and future directions. Front. Ecol. Evol. 6, 1–12. https://doi.org/10.3389/fevo.2018.00235

Sharma, P., Chettri, N., Uddin, K., Wangchuk, K., Joshi, R., Tandin, T., Pandey, A., Gaira, K.S., Basnet, K., Wangdi, S., Dorji, T., Wangchuk, N., Chitale, V.S., Uprety, Y., Sharma, E., 2020a. Mapping human–wildlife conflict hotspots in a transboundary landscape, Eastern Himalaya. Glob. Ecol. Conserv. 24, e01284. https://doi.org/10.1016/j.gecco.2020.e01284

Sharma, P., Panthi, S., Yadav, S.., Bhatta, M., Karki, A., Duncan, T., Poudel, M., Acharya, K.., 2020b. Suitable habitat of wild Asian elephant in Western Terai of Nepal. Ecol. Evol. 10, 6112–6119. https://doi.org/10.1002/ece3.6356

Simons, R.R.L., Croft, S., Rees, E., Tearne, O., Arnold, M.E., Johnson, N., 2019. Using species distribution models to predict potential hot-spots for Rift Valley Fever establishment in the United

Kingdom. PLoS One 14, e0225250. https://doi.org/10.1371/journal.pone.0225250

Singh, S.K., Jabin, G., Basumatary, T., Bhattarai, G.P., Chandra, K., Thakur, M., 2019. Resolving the trans-boundary dispute of elephant poaching between India and Nepal. Forensic Sci. Int. Synerg. 1, 146–150. https://doi.org/10.1016/j.fsisyn.2019.07.007

Sitompul, A.F., Griffin, C.R., Rayl, N.D., Fuller, T.K., 2013. Spatial and temporal habitat use of an Asian elephant in Sumatra. Animals 3, 670–679. https://doi.org/10.3390/ani3030670

Smith, J.L.D., Mishra, H.R., 1992. Status and distribution of Asian elephants in Central Nepal. Oryx 26, 34–38. https://doi.org/10.1017/S0030605300023206

Songhurst, A., McCulloch, G., Coulson, T., 2016. Finding pathways to human-elephant coexistence: a risky business. Oryx 50, 713–720. https://doi.org/10.1017/S0030605315000344

Strobl, C., Boulesteix, A.L., Kneib, T., Augustin, T., Zeileis, A., 2008. Conditional variable importance for random forests. BMC Bioinformatics 9, 1–11. https://doi.org/10.1186/1471-2105-9-307

Sukumar, R., 2003. The living elephants: evolutionary ecology, behaviour, and conservation, Oxford University Press.

Sukumar, R., 1990. Ecology of the Asian elephant in Southern India. II. Feeding habits and crop raiding patterns. J. Trop. Ecol. 6, 33–53. https://doi.org/10.1017/S0266467400004004

Sukumar, R., 1989. Ecology of the Asian elephant in Southern India. I. Movement and habitat utilization patterns. J. Trop. Ecol. 5, 1–18. https://doi.org/10.1017/S0266467400003175

ten Velde, P., 1997. A status report of Nepal's wild elephant populatio. Unpubl. report, WWF Nepal, Kathmandu.

The Guardian, 2020. Killing of elephant with explosive-laden fruit causes outrage in India | India | The Guardian [WWW Document]. URL https://www.theguardian.com/world/2020/jun/05/killing-of-elephant-with-explosive-laden-fruit-causes-outrage-in-india (accessed 10.24.21).

Tiller, L.N., Humle, T., Amin, R., Deere, N.J., Lago, B.O., Leader-Williams, N., Sinoni, F.K., Sitati, N., Walpole, M., Smith, R.J., 2021. Changing seasonal, temporal and spatial crop-raiding trends over 15 years in a human-elephant conflict hotspot. Biol. Conserv. 254, 108941. https://doi.org/https://doi.org/10.1016/j.biocon.2020.108941

Ting, K.M., Witten, I.H., 1999. Issues in stacked generalization. J. Artif. Intell. Res. 10, 271–289. https://doi.org/10.1613/jair.594

Tiwari, S.K., 2017. Documenting and securing corridors: the approach and process, in: Menon, V., Tiwari, S.K., Ramkumar, K., Kyarong, S., Ganguly, U., Sukumar, R. (Eds.), Right of Passage: Elephant Corridors of India [2nd Edition]. Wildlife Trust of India, New Delhi, pp. 42–63.

Treves, A., Wallace, R.B., Naughton-Treves, L., Morales, A., 2006. Co-managing human–wildlife conflicts: a review. Hum. Dimens. Wildl. 11, 383–396. https://doi.org/10.1080/10871200600984265

Tuia, D., Kellenberger, B., Beery, S., Costelloe, B.R., Zuffi, S., Risse, B., Mathis, A., Mathis, M.W., van Langevelde, F., Burghardt, T., Kays, R., Klinck, H., Wikelski, M., Couzin, I.D., van Horn, G., Crofoot, M.C., Stewart, C. V., Berger-Wolf, T., 2022. Perspectives in machine learning for wildlife conservation. Nat. Commun. 13, 1–15. https://doi.org/10.1038/s41467-022-27980-y

UNEP-WCMC, 2021. Checklist of CITES species, CITES Secretariat, Geneva, Switzerland and UNEP-WCMC, Cambridge, United Kingdom.

Valavi, R., Elith, J., Lahoz-Monfort, J.J., Guillera-Arroita, G., 2021. Modelling species presence-only data with random forests. Ecography (Cop.). 44, 1731–1742. https://doi.org/10.1111/ecog.05615

Vasilijević, M., Zunckel, K., Mckinney, M., Erg, B., Schoon, M., Michel, T.R., 2015. Transboundary conservation: a systematic and integrated approach. Best practice protected area guidelines Series No. 23, Gland, Switzerland: IUCN.

Wang, K., Zheng, X., Wang, G., Liu, D., Cui, N., 2020. A multi-model ensemble approach for gold mineral prospectivity mapping: a case study on the Beishan region, Western China. Minerals 10, 1–19. https://doi.org/10.3390/min10121126

Willi, M., Pitman, R.T., Cardoso, A.W., Locke, C., Swanson, A., Boyer, A., Veldthuis, M., Fortson, L., 2019. Identifying animal species in camera trap images using deep learning and citizen science. Methods Ecol. Evol. 10, 80–91. https://doi.org/10.1111/2041-210X.13099

Williams, A.C., Johnisingh, A.J.T., Krausman, P.R., Qureshi, Q., 2008. Ranging and habitat selection by Asian elephants (*Elephas maximus)* in Rajaji National Park, North West India. J. Bombay Nat. Hist. Soc. 105, 24–33.

Williams, C., Tiwari, S.K., Goswami, V.R., de Silva, S., Kumar, A., Baskaran, N., Yoganand, K., Menon, V., 2020. *Elephas maximus*, The IUCN Red List of Threatened Species 2020: e.T7140A45818198. https://doi.org/https://dx.doi.org/10.2305/IUCN.UK.2020-3.RLTS.T7140A45818198.en

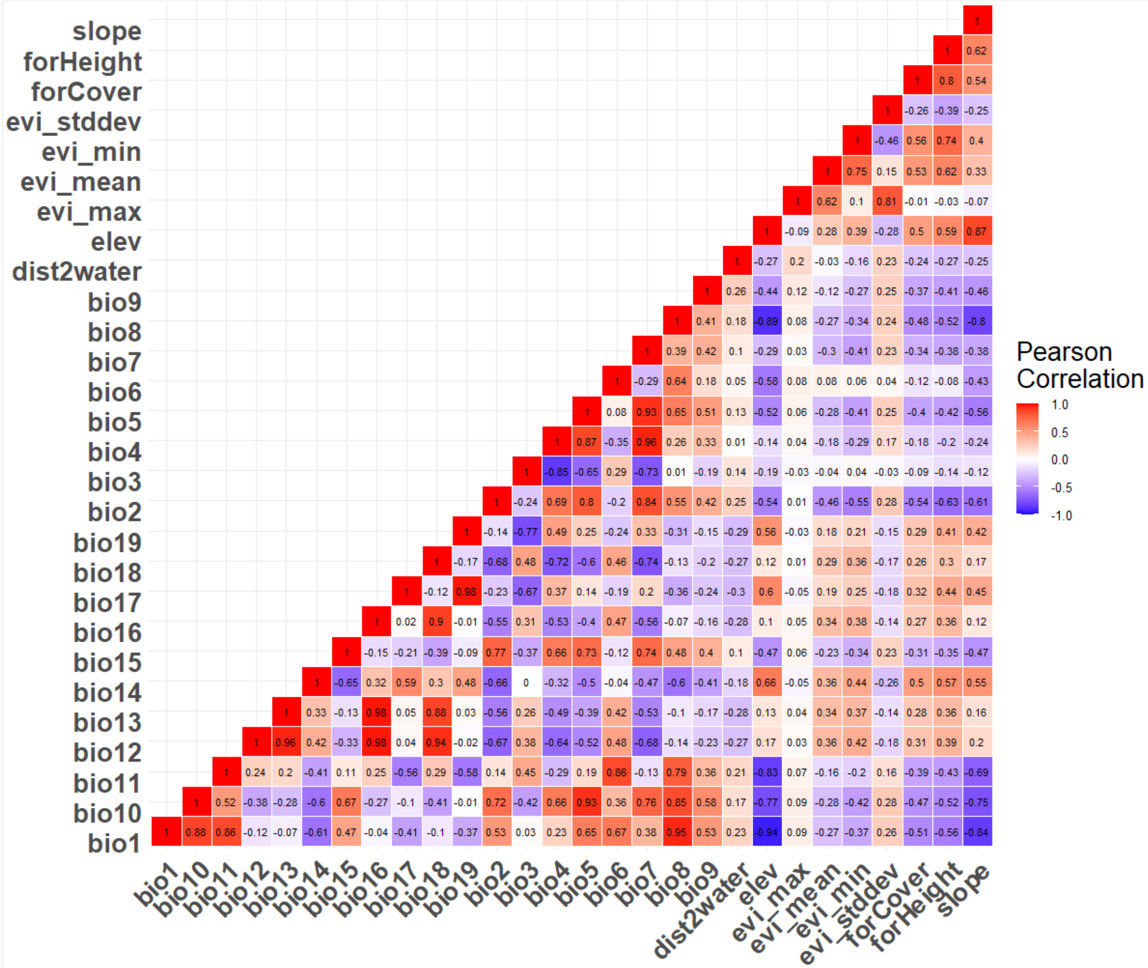
Wilson, S., Davies, T.E., Hazarika, N., Zimmermann, A., 2013. Understanding spatial and temporal patterns of human–elephant conflict in Assam, India. Oryx 49, 140–149. https://doi.org/10.1017/S0030605313000513

Zhou, Z.-H., 2009. Ensemble Learning. Encycl. Biometrics 1, 411–416. https://doi.org/10.1007/978-0-387-73003-5_293

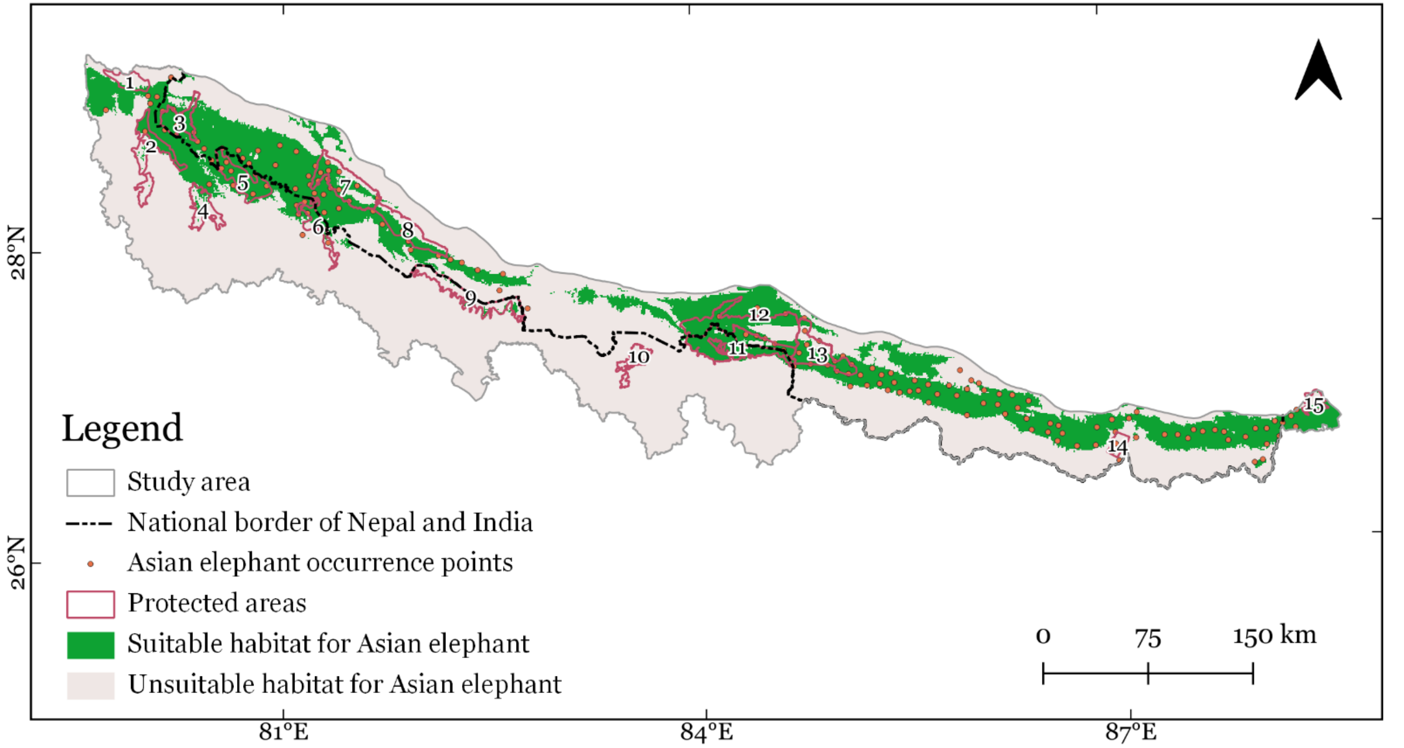
ANNEX

Annex 1: Pearson's correlation coefficient and variance inflation factor (VIF) of all the predictor variables selected for habitat suitability modelling

Variables	VIF
bio7	1.773071e+13
bio5	1.631739e+13
bio6	2.401920e+12
bio11	3.231568e+03
bio10	3.080223e+03
bio4	3.032921e+03
bio1	2.104069e+03
bio2	2.099632e+03
bio3	1.257700e+03
bio12	9.883569e+02
bio16	7.678627e+02
bio19	1.166353e+02
bio17	8.464678e+01
bio13	6.334728e+01
elev	4.766538e+01
bio15	4.590756e+01
evi_stddev	2.862009e+01
evi_max	2.238315e+01
bio8	2.166738e+01
evi_min	2.007296e+01
bio18	1.723491e+01
evi_mean	1.238721e+01
bio14	7.790962e+00
forHeight	7.195824e+00
slope	6.559218e+00
forCover	3.315188e+00
bio9	3.256938e+00
dist2water	1.786454e+00



Annex 2: Predicted suitable habitat from boosted regression tree algorithm for Asian elephants in Nepal-India transboundary landscape

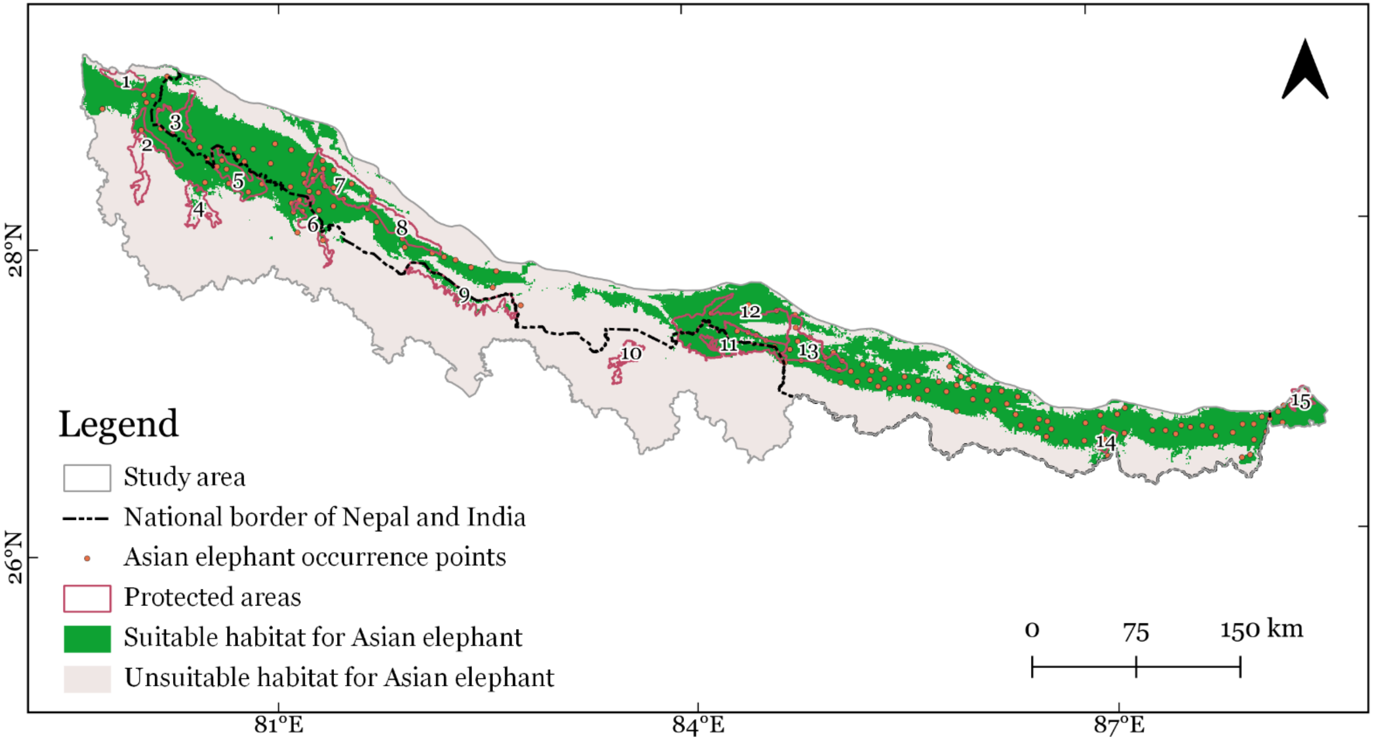


Suitable habitat = 22,487 km²

Unsuitable habitat = 52,321 km²

The map shows study area, national border of Nepal and India, Asian elephant occurrence points used in habitat suitability modelling using boosted regression tree algorithm (BRT) and protected areas. Nine protected areas lie in India: Nandhaur Wildlife Sanctuary(1), Pilibhit Tiger Reserve(2), Kishanpur Wildlife Sanctuary(4), Dudhwa National Park(5), Katarniaghat Wildlife Sanctuary(6), Sohelwa National Park(9), Sohagibarwa Wildlife Sanctuary(10), Valmiki Tiger Reserve(11) and Mahananda Wildlife Sanctuary(15), whereas six PAs lie in Nepal: Suklaphanta National Park(3), Bardiya National Park(7), Banke National Park(8), Chitwan National Park(12), Parsa National Park(13) and Koshi Tappu Wildlife Reserve(14).

Annex 3: Predicted suitable habitat from random forest algorithm for Asian elephants in Nepal-India transboundary landscape

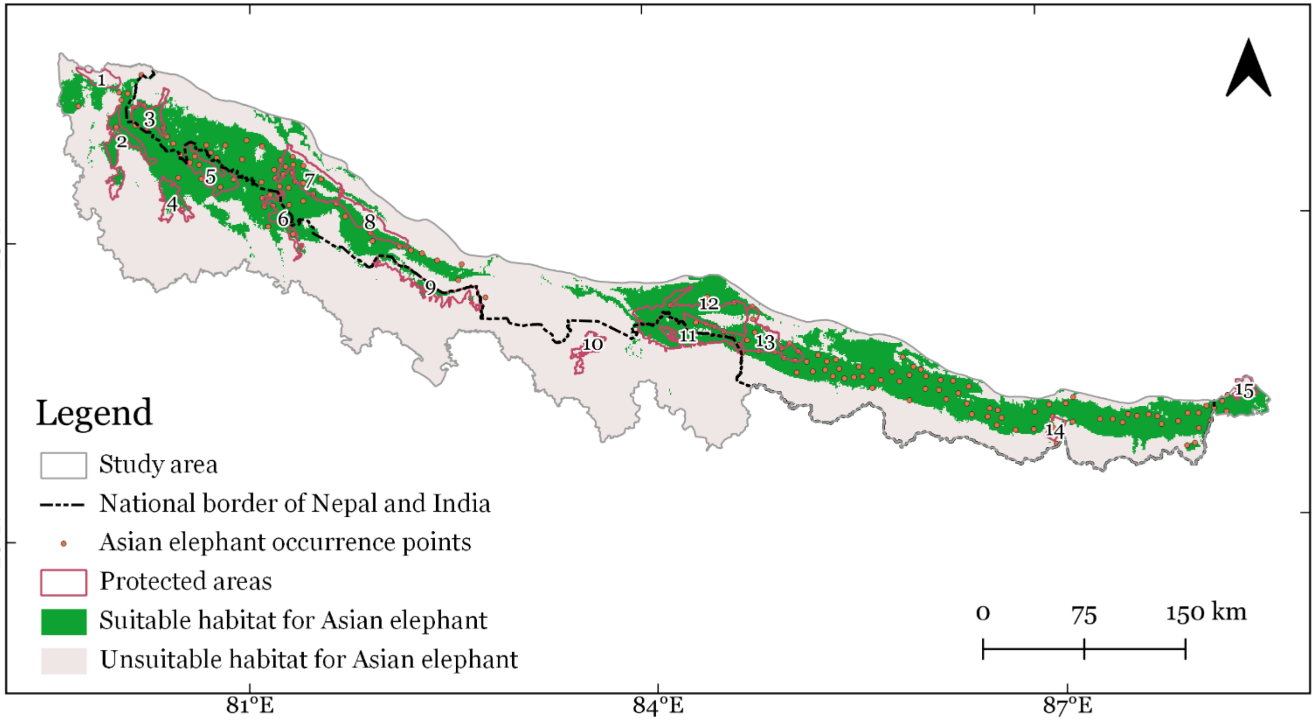


Suitable habitat = 26,783 km²

Unsuitable habitat = 48,025 km²

The map shows study area, national border of Nepal and India, Asian elephant occurrence points used in habitat suitability modelling using random forest algorithm (RF) and protected areas. Nine protected areas lie in India: Nandhaur Wildlife Sanctuary(1), Pilibhit Tiger Reserve(2), Kishanpur Wildlife Sanctuary(4), Dudhwa National Park(5), Katarniaghat Wildlife Sanctuary(6), Sohelwa National Park(9), Sohagibarwa Wildlife Sanctuary(10), Valmiki Tiger Reserve(11) and Mahananda Wildlife Sanctuary(15), whereas six PAs lie in Nepal: Suklaphanta National Park(3), Bardiya National Park(7), Banke National Park(8), Chitwan National Park(12), Parsa National Park(13) and Koshi Tappu Wildlife Reserve(14).

Annex 4: Predicted suitable habitat from maximum entropy model for Asian elephants in Nepal-India transboundary landscape



Suitable habitat = 26,546 km²

Unsuitable habitat = 48,262 km²

The map shows study area, national border of Nepal and India, Asian elephant occurrence points used in habitat suitability modelling using maximum entropy model (MaxEnt) and protected areas. Nine protected areas lie in India: Nandhaur Wildlife Sanctuary(1), Pilibhit Tiger Reserve(2), Kishanpur Wildlife Sanctuary(4), Dudhwa National Park(5), Katarniaghat Wildlife Sanctuary(6), Sohelwa National Park(9), Sohagibarwa Wildlife Sanctuary(10), Valmiki Tiger Reserve(11) and Mahananda Wildlife Sanctuary(15), whereas six PAs lie in Nepal: Suklaphanta National Park(3), Bardiya National Park(7), Banke National Park(8), Chitwan National Park(12), Parsa National Park(13) and Koshi Tappu Wildlife Reserve(14).