Mapping the Climate Suitability of Potato (*Solanum Tuberosum*) using MaxEnt Modeling Approach in Pakistan

TAYYABA KHALIL

(Regn # 00000205230)

Institute of Environmental Sciences and Engineering (IESE)

School of Civil and Environmental Engineering (SCEE)

National University of Science and Technology (NUST)

Islamabad, Pakistan

(2020)

Mapping the Climate Suitability of Potato (*Solanum Tuberosum*) using MaxEnt Modeling Approach in Pakistan

TAYYABA KHALIL

Regn Number

00000205230

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

In

Environmental Science

Institute of Environmental Science and Engineering (IESE)

School of Civil and Environmental Engineering (SCEE)

National University of Sciences and Technology (NUST)

Islamabad, Pakistan

(2020)

Certificate

This dissertation submitted by Ms. Tayyaba Khalil is accepted in its present form, by the Institute of Environmental Sciences and Engineering (IESE), School of Civil and Environmental Engineering (SCEE), National University of Sciences and Technology (NUST), Islamabad, Pakistan as satisfying the partial requirement for the degree of Master of Science in Environmental Science.

Supervisor____________

Dr. Sofia Baig

Professor

IESE, SCEE, NUST

GEC Member ____________

Dr. Salman Atif

Associate Professor

IGIS, NUST

GEC Member ____________

Dr. Muhammad Arshad

Assistant Professor

IESE, SCEE, NUST

Co-Supervisor ____________

Dr. Saeed Asad

CCRD, COMSATS University Islamabad

THESIS ACCEPTANCE CERTIFICATE

It is certified that the contents and forms of the thesis entitled "**Mapping the Climate Suitability of Potato (Solanum Tuberosum) using MaxEnt Modeling Approach in Pakistan**" submitted by Ms. Tayyaba Khalil, Registration # 00000205230 has been satisfactory, free of errors and plagiarism for the requirements of the degree of Master of Science in Environmental Science. It is further certified that necessary amendments as pointed out by the GEC members of the scholar have also been incorporated in the said thesis.

Signature: ____________________________

Name of the Supervisor: Dr. Sofia Baig

Dated: _______________________________

Head of Department: _____________________

Dated: _________________________________

Principal: _______________________________

Dated: _________________________________

Acknowledgements

I am thankful to my Creator Allah Subhana-Watala to have guided me throughout this work at every step and for every new thought which You setup in my mind to improve it. Indeed I could have done nothing without Your priceless help and guidance. Whosoever helped me throughout the course of my thesis, whether my parents or any other individual was Your will, so indeed none be worthy of praise but You.

I am profusely thankful to my beloved parents, specially my father who raised me when I was not capable of walking and who struggled his whole life to make me what I am today. He continued to support me throughout in every department of my life.

I would also like to express special thanks to my supervisor Dr. Sofia Baig for her help throughout my thesis and also for teaching me R software for which I had no previous background.

I would also like to pay special thanks to Sakina Fatima and my friend Nusaiba Khubaib for their tremendous support and cooperation. Each time I got stuck in something, they came up with the solution. Without their help I wouldn't have been able to complete my thesis. I appreciate their patience and guidance throughout the whole thesis.

I would also like to thank Dr. Saeed Asad, Dr. Muhammad Arshad and Dr. Salman Atif for being on my thesis guidance and evaluation committee and express my special thanks to Dr. Salman for his help and cooperation.

Finally, I would like to express my gratitude to all the individuals who have rendered valuable assistance to my study.

Dedicated to my exceptional parents and adored siblings whose tremendous support and cooperation led me to this wonderful accomplishment.

Abstract

Impacts of climate change are becoming vibrant and clear and one of the indirect impacts of climate change includes the subject of food security. Pakistan is included among one of the most vulnerable nations to climate change. In the 21st century Pakistan is already experiencing visible adverse impacts of climatic changes leading to food insecurity. Potato is a food security crop declared by FAO as it is rising as a big source against hunger. To assess the impact of climate change on potato crop in Pakistan, it is essential to analyze its distribution in future climate change scenarios. Maximum Entropy model has been used in this study to predict the spatial distribution of Potato in 2070 using two CMIP5 models for two climate change scenarios (RCP 4.5 and RCP 8.5). The data used for this study was obtained mainly from Pakistan Bureau of Statistics and was processed in the form of presence locations. R-Software was used in this research and there were 19 Bioclim variables incorporated in this research along with irrigation, soil type and elevation. Climate suitability maps were generated for current and future scenarios for two models and the results indicated drastic decrease in the suitable area for potato growth in RCP 8.5 and slight decrease in suitable area in RCP 4.5 with both the models. The Jackknife test predicted the most important contributors in current climate distribution of potato were isothermality, precipitation of wettest quarter and annual precipitation. In model 1 with RCP 4.5, the predictors were annual precipitation and precipitation of coldest quarter and with RCP 8.5 the variables were precipitation of wettest month and mean temperature of warmest quarter and In model 2 with RCP 4.5 important environmental factors included annual mean temperature and precipitation of coldest quarter and with RCP 8.5 the variables were precipitation of coldest quarter and irrigation.

Key Words: *Maxent, Food security, Potato, Representation Concentration Pathway*

Table of Contents

List of Figures

List of Tables

CHAPTER 1: INTRODUCTION

1.1 Overview

Recently, the subject of climate change is the cause of concern globally as its occurrence frequency is unprecedented. Average temperatures have undergone an amplification of 0.85℃ in the last century and it will increase to maximum 8℃ by 2100 according to the predictions by IPCC (Wei et al., 2018). The scientists' community and ruling authorities of various nations have shown their interest in resolving the problem of climate change (Gao et al., 2016). It is observed that the overall global climate change, mainly an increase in temperature, has negatively influenced the natural ecosystems and agriculture (Streck 2005; Field et al., 2014; Ochieng et al. 2016; Kukal et al. 2018). Climate has of various areas has affected the reproduction of plant species and contributes to the determination of their geographical distribution (Wang et al., 2018). In coming years, the most suffering communities from climate change will be the ones in Global South, especially sub-Saharan Africa and South and Southeast Asia (IPCC, 2014). Out of all the other sectors agriculture is the directly affected by global warming and climatic changes as there are altered weather conditions biophysical effects (Rosenzweig et al, 2014).

There are complex means in which climate change influences the agriculture and food production. Alteration in agro-ecological conditions due to changes in climate are direct responsible for the change in food production patterns. There are also changes in growth and distribution of incomes along with increased demand for agriculture (Schmidhuber et al., 2007).

The United Nations Sustainable Development Goals (SDGs) highlights the crucial requirement for the transformation of agriculture to meet the food requirements, sustainability of various ecosystems, prosper economy and social impartiality over the years ahead. In future, food demand is predicted to be grown by 70-85% with the increase in population (about 9 billion people) in 2050 (FAO, 2017).

1.1.1 Food Security

According to Food and Agriculture Organisation, when people having limited physical, social, and economic access to sufficient, safe, and nutritious food to fulfill their dietary needs and food preferences to spend a healthy life is a situation known as food security. The definition has four important dimensions of food supply including availability, stability, access, and utilization. Pakistan has been fighting against hunger and agriculture inefficiency for long time in the past (Chaudhri, 2017). Though it is an agriculture economy and has means for the production of sufficient food in accordance with the population needs, but it also lacks in infrastructure leading to difficulties in management as well as distribution of food. There can be several reasons including poor research and development programs that generate good variety of seeds, poor planning and marketing and lack of governmental interventions for agricultural economy (Asim and Akbar, 2019). Lack in agriculture performance has aggravated food insecurity risk and this might turn on the situation of hazardous health and economic conditions in future (Baig et al., 2019). Out of 119 countries Pakistan is ranked at 106 according to Global Hunger Index (2017) having an alarming score of 32.6 (Grebmer et al., 2017). This indicates that with other development issues, food security is one of the major concern of the country. A report by WFP (2017) states that overall 60% of the Pakistan's population has food insecurity problem and for this 60%, there is always a shortage of food and thus malnutrition is widespread in the country.

Pakistan is a country that extends over an area of 796,000 Km^2 with a diverse climate. The eastern part of the southern half mainly get precipitation through summer monsoon rains, whereas the north and west of the southern part of Pakistan get rainfall mainly in winter (from December to March). Out of overall precipitation summer monsoon contributes around 60% of total rainfall.. The climate varies from arid to semiarid region. The temperature drop in the area during winter season to –50°C and the in warmest season temperature remains around 15°C (ADB, 2008).

1.1.2 Potato Cultivation in Pakistan

Potato *(Solanum tuberosum)* is a crop which is consumed all over the world. Potato crop has been orginated mainly from South America and it has been transported and exported from Peru to the other countries (Spooner et al., 2005).

Potato crop can be grown in warmer as well as cooler climates, however it cannot thrive in harsh climatic conditions such drought, high temperature and high humidity (Zhao et al., 2016). Climate change is responsible for acute effects in food supply Thus; the major challange for modern agriculture is the development of strategies that can cope with possible negative impacts climate change in the future to ensure food security by 2050 and beyond.

Usually potatoes are the key source of income and attaining the food security in developing nations (Raymundo et al., 2014). Potato is very important food crop extensively cultivated in various agro-climatic conditions in the country and all over the globe. It is amongst the four major staples which contributes to the consumption and food requirements of the nation. Regardless of the ease in the agriculture, potato production in our country is still not that promising in comparison to the countries like India and Bangladesh. Extreme temperature along with other factors like nutrient deficit soil, drought, salinity, poor irrigation water, lack of quality seeds are among the abiotic stresses which pose challenges to potato productivity. In Pakistan there is low allocation of land for the cultivation of potatoes than the other crops thus leading to low yields (Majeed et al., 2018). According to the latest agriculture statistics conducted in 2017, area under production for potato is 170,300 hectares with production of about 4 Million Tons (PBS, 2017). Pakistan currently relies on locally produced potato crop and is self-sufficient in its cultivation as far as household consumption is concerned (Pakissan, 2018). The crop is cultivated in Pakistan wherever there is availability of irrigation, but largely they are grown in the central and northern plains of Punjab, especially the districts of Okara, Sahiwal, Kasur, Sialkot, Sheikhupura, Jhang, Lahore, Narowal, Pak pattan, Gujranwala, T.T. Singh and Khanewal. In KPK they are grown in Nowshera, Dir, Swat, Balakot, Sakardu and Mansehra districts. Pishin, Killa Saifulla and Kalat districts in Balochistan and Gilgit district in Gilgit-Baltistan is where the potato is grown (Arain, 2018).

1.1.3 Species Distribution Model - Maxent

There has been an increased utilization of Species distribution Models (SDMs) for the management of environment which basically include managing the endangered species. There use has also been more common in habitat suitability studies, and for the changes in environment and various other impacts (Ghanbarian et al., 2019). Species distribution modelling (SDM) is known by different other names like environmental (or ecological) niche modelling (ENM), or habitat modelling.

SDMs demonstrate the interaction among the presence of species and environmental variables (Elith et al., 2006). SDMs are governed by the interrelation of bioclimatic or environmental variables and species presence records as they determine habitat suitability on a wide scale by taking a look at ecological drivers (Elith & Leathwick, 2009). There have been developed a diverse range of modelling methods that typically range from instructions-based models to machine learning models. The quality and quantity of input information defines the accuracy of the model. The information can be random occurrence data sampling as well as more precise presence-absence records (Kramer-schatd et al., 2013). There are methods which use environmental data to draw correlation, making a model for environmental conditions meeting ecological requirement and predict the suitability of a species.

There have been four means in which ENMs are used such as estimation of habitat suitability where species are known to be present, estimation of suitability of a habitat in areas which are not known to be occupied with that particular species, it is also used to find changes in habitat suitability with time by using different environmental scenarios and it is known to be used in the estimation of the species niche. (Warren et al., 2011).

Maxent software utilizes the principle of maximum entropy on presence only data for the estimation of set of functions relating the environmental variables with habitat suitability for approximation of the species potential geographic distribution (Philips et al., 2008). The basic objective of using the Maxent software is the estimation of a specific probability distribution by finding out probability distribution of maximum entropy. While using Maxent model, all the pixels in defined area of research makes up a space where probability distribution is defined. Pixels with known species occurrence records constitute the sample points. Maxent model has various advantages including (1) requirement of only presence data along with the knowledge of environmental factors in the considered study area. (2) Utilization of both the continuous and categorical variables, and incorporation of interactions between these variables. (3) There are algorithms developed which support the optimal probability distribution. (5) The results of the model are continuous which allow the makeup of fine divisions between the modeled suitability of different areas. It also allows the predictions in binary form if required.

For Maxent different variables are required for the model input which are described below.

1.1.4. Bioclimatic variables

Bioclimatic variables are derived from the monthly temperature and rainfall values for the generation of more meaningful variables which are used un specie distribution modeling and ecological niche modeling.

They are coded as follows:

BIO1 = Annual Mean Temperature

BIO2 = Mean Diurnal Range (Mean of monthly (max temp - min temp))

BIO3 = Isothermality (BIO2/BIO7) (×100)

BIO4 = Temperature Seasonality (standard deviation ×100)

BIO5 = Max Temperature of Warmest Month

BIO6 = Min Temperature of Coldest Month

BIO7 = Temperature Annual Range (BIO5-BIO6)

BIO8 = Mean Temperature of Wettest Quarter

BIO9 = Mean Temperature of Driest Quarter

BIO10 = Mean Temperature of Warmest Quarter

BIO11 = Mean Temperature of Coldest Quarter

BIO12 = Annual Precipitation

BIO13 = Precipitation of Wettest Month

BIO14 = Precipitation of Driest Month

BIO15 = Precipitation Seasonality (Coefficient of Variation)

BIO16 = Precipitation of Wettest Quarter

BIO17 = Precipitation of Driest Quarter

BIO18 = Precipitation of Warmest Quarter

BIO19 = Precipitation of Coldest Quarter

(Source: www.worldclim.org)

1.1.4 CMIP5 models

Coupled Model Intercomparison Project Phase 5 (CMIP5) is an initiative of World Climate Research Programme (WCRP) as it provides IPCC AR5 (Fifth Assessment Report, IPCC 2013) with the most important environmental variables. A huge network of scientists joined hands for the development of the models and contribution to this work (Taylor et al., 2012). A source of output variables and a time series was generated by taking the initiative, and thus their accces was given to the scientific community for the research purposes, and can be accessed from the ESGF portal.

There is a database for CMIP5 models which permits the access to the variable output from the individual models that are available in the form of time series from 2006 to 2100 except for some models that predict till 2300. The computing for these projections is done via different representative pathways such as RCP 2.5, 4.5, 6 and 8.5. every representative pathway comprose the same category of the data with different values projecting different levels of carbon emissions over a time as a result of human activities (Combal and Caumont, 2016).

1.1.5 Representative Concentration Pathways (RCPs)

The response of earth and ecosystem is not enough to cope with the climatic changes but also the reaction of humans to these changes by changing the lifestyles and living. As the responses are undefined, thus, scenarios with future climate are being utilized to discover the consequences of diverse options. Decision making has become easy for governments and institutions with these scenarios. Policy decisions linked to risk factor and values will help determine the pathway followed. The Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (AR5) has presented a novel method of developing scenarios. These scenarios span the range of possible radiative forcing scenarios, and are known as representative concentration pathways (RCPs). They are suggested paths for GHGs and aerosol concentrations and land use change which are consistent with a range of broad climate outcomes used by scientists. The pathways are characterized by the amount of radiative forcing which is the additional heat trapped by the lower atmosphere as a result of surplus GHGs emission. Unit used to measure the radiative forcing is Watts per square metre (W/m^2).

There are four representative pathways listed below.

Table 1.1: Representative Concentration Pathways

Radiative forcing pathways	**Atmospheric CO_2 equivalent (parts per million)**	**When**
8.5	>1370	Rising by 2100
6	850	Stabilization after 2100
4.5	650	Stabilization after 2100
2.6	490	Peak before 2100 then decline

(Jubb et al., 2013)

The RCPs considered for this research are RCP 4.5 and RCP 8.5.

1.1.5 Significance of the study

Pakistan is already experiencing visible adverse impacts climatic changes regarding glacial melt, droughts, intense flooding and heat waves along with early summer season and hot winters. All of these factors are negatively impacting the plant species and agriculture of the country. Thus it is essential to predict the future distribution patterns of major crops of Pakistan such as potato in order to cope with food security issues. For the assesement of climate change impacts on potato crop in Pakistan especially in the districts which are rainfed, various modeling techniques can be used (Nabout et al., 2011). Species Distribution Models (SDMs) are getting popular for their use in empirical studies. These SDMs use a combination of numerical tools along with species presence data along with various environmental factors. MaxEnt is one of the model which forecasts the geographical distribution about a particular species using only its presence data (Byeon etal., 2018). The model has been widely used all across the world to predict the changes in geographical distribution of species under the climate change (Guo et al., 2017).

1.2 Study Objectives

The objectives include:

1 Prediction and comparison of effects of different Climate Change scenarios on potato distribution in Pakistan.

2 Identification of environmental variables that are highly correlated with spatial distribution of potato in Pakistan.

CHAPTER 2: LITERATURE REVIEW

This chapter focuses on the previous information about geographical distribution of Species and relevant studies all over the globe and in Pakistan using Species Distribution Models specifically Maxent Species Distribution Model, its evaluation and outcomes.

2.1 Species Distribution Models (SDMs) and their importance

The idea behind the inititative of Species Distribution Modeling is to predict the geographical distruibution of species with the help of various environmental and climatic factors. For robust and reliable modeling of Species, there is immense need for the understanding of ecological principles and uncertainties in the model. (Hallston, 2011). The main species on which the models apply include animals, plants and microorganisms. These models are utilized in environmental management including the management of species that are endangered and for the studies where habitat suitability of the species is concerned. (Ghanbarian et al., 2019). As far as non living factors are concerned, the most influencing factors in limiting the distribution of plants are climate and soil characterstics. Biotic components also influence the distribution of species and may result in the spatial changes in communities of plants. (Chahouki et al., 2016). Species Distribution Models (SDMs) use a combination of numerical tools along with species presence data along with various environmental factors.

One type of SDM is Random Forest Algorithem that uses more than one modelof several Decision Trees to have a better prediction. It generates many classification trees and a bootstrap sample technique is applied to train each tree from the set of training data. A study on Mountainous rangelands in Iran was conducted by Sahragard et al (2018) as there is huge role of mountainous rangelands in providing food and fodder for livestock mostly in summers. They also are vital in maintaining the ecological balance. (RFM) was used to identify the environmental variables that influenced all the plant species distribution in rangeland. The study was carried out in rangelands of the Taftan Mountain, Sistan and Baluchestan Province in southeastern Iran. Systematic random sampling was used to determine the potential distribution and 90 presence points records were used for the purpose.. The methods used for the preparation of environmental variables layers were Kriging interpolation method and Geographic Information System facilities. AUC (Area under curve) showed good results of predictive modeling and thus it is concluded that RFM is a strong modeling approach to examine the relationship of species distribution and environmental variables.

Maxent (Maximum Entropy) is another species distribution model which forecasts the distribution of a given species using only presence data (Byeon etal., 2018) and provides with most precise distribution function established on best entropy (Reddy et al., 2015). The use of model for the prediction of changes in the geographical distribution of species under the climate change is wide. (Guo et al., 2017) where, Singh (2020) has used Maximum Entropy Modeling approach to evaluate the changes in the forest cover over time and provided the coverage for the habitats of threatened species in Southeast Asia. Model is discussed in detail with more examples as under:

2.2 Maxent

Maxent is an SDM that follows a general-purpose approach developing predictions that utilize information which is only partially available. It is based on statistical mechanics (Jaynes, 1957), and supports a wide array of research via a conference titled "Maximum Entropy and Bayesian Methods" that is held annually. Maxent can also be termed as a general-purpose method which can be used for presence only modeling of distribution of a given species. Thus, it is suitable for all such applications that involve presence only datasets. Maxent works on predicting a specific targeted probability distribution by determining the probability distribution of maximum entropy for that species, which means the distribution that is most spread out or even, while dealing with a set of limitations i.e. availability of only partial information about the distribution of target species. This incomplete information of the target species distribution mostly is present in the form of real-valued variables (features) while the limitations are the anticipated values of every feature should be comparable to the average of all the sample points taken from the distribution of target species. When we use Maxent for species distribution modelling using presence-only approach, the coordinates of the study area define the Maxent probability distribution, coordinates that have known species presence records make up the sample points while the features consist of climatic conditions like environmental or bioclimatic variables, elevation, soil or vegetation type and their functions (Phillips et al., 2006). The prediction can be of both scenarios; current and future. While the current prediction of suitable habitat of a species shown by MaxEnt only utilizes the current (present environment) environmental and bioclimatic variables, the future scenarios typically include the environmental variables that are derived for future years (e.g. 2050 or 2070) using various global circulation models that predict different representative concentration pathways (RCPs). Environmental or bioclimatic variables will behave differently under different RCPs.

2.3 Maxent modeling for Animals and Microorganisms

Maxent is being used in various studies for the prediction of suitable habitat for the organisms including plants, animals and even microorganisms. A study was carried out on *Xylella fastidiosa* which is xylem-limited Gram-negative bacterium, a plant pathogen responsible for causing disease in plants. Maxent was developed for the particular specie present in Italy with the objective to analyze the geographical distribution of *Xylella fastidiosa*. The most contributing variables include precipitation of the driest (40.3%) as well as wettest (30.4%) months. According to the results of Species distribution models high occurance probability is seen for X. fastidiosa occurrence in most of the areas of Apulia, Calabria, Basilicata, Sicily, Sardinia and coastal areas of Campania, Lazio and south Tuscany. The Area under the curve with a value more than 0.8 showed that the results of the model are accurate. This study showed that the targeted specie has potential to overcome the current boundaries of distribution and affect areas of Italy outside Apulia (Bosso et al., 2016).

Maxent was used in another study to calculate the habitat suitability index for Common Otter in its area of occurrence by incorporating topographic and environmental layers along with species sighting data. Eight potential predictor variables were identified and selected on the basis of their significance to target species distribution and habitat evaluation. Land cover map of the study area was developed by using LANDSAT satellite data of 30 m spectral resolution. According to results (BIO 6) Minimum Temperature of Coldest Month and (BIO 5) Maximum Temperature of Warmest Month had highest percentage contribution with 45% and 29.1% respectively in the model. Areas of high habitat suitability accounted for only small percentage (0.79%) of total area whereas unsuitable habitat area accounted for (94.39%) which clearly depicts the declining Otter population in the area. Maxent Modelling performance calculated by AUC value (0.988) had evaluated the model with high performance. The outcome of the study draws for an immediate attention towards conservation and management of Eurasian Otter in its habitat to protect it from further getting into Endangered (Ahmed et al., 2017).

2.4 Maxent Modeling for Plants

Use of Maxent for the habitat identification and prediction for the plants has become common in recent years. Few examples for the habitat suitability of plants are as follows.

Garcia and his coworkers (2013) used Maxent approach to perform their analysis on 14 threatened forest tree species for the estimation of climatic impacts on their distribution in Philippines. Occurrence records required bioclimatic variables which were incorporated in the software for the prediction of the impact for current and future climate conditions on the distribution of species. ROC and AUC revealed good results of the model ranging from 0.7 to 0.97. Half of the species were found to be benefitted from future climate scenarios as their habitat will increase, while half of the species would get their suitable habitat decreased in area.

Another study was conducted on maxent probability distribution in Chaharbagh rangeland of Golestan in Province of Iran. The species used for the research were Artemisia aucheri and Bromus tomentellus-Festuca. The method utilized was vegetation sampling using random systematic sampling method where 120 plots in the study area were placed. Soil sampling was done from the depth of soil up to 30cm below the ground due to variation in the terrain along deep roots under the soil. The soil properties taken into the consideration were texture of the soil, organic carbon content in the soil, lime, pH, electrical conductivity, and Nitrogen content. Topographical data obtained from a DEM map including elevation, slope, and aspect was also taken into the account for the better predictions of the model. The first step was to obtain the map of soil factors in the GIS software and then analysis was performed using maxent software. Jackknife testing and response curves were used to analyze the important variable contributing to the prediction of species. According to the findings, Nitrogen, sand, and clay impacted the most on A. aucheri and the most influencing factors on on the presence of B. tomentellus-F. ovina were found to be Nitrogen, sand, silt, clay, and lime in soil.

In a study by Wei et al., (2018), current and future trends for *C. tinctorius* were modeled with three RCPs 2.6, 4.5 and 8.5 for 2050s and 2070s using a maximum entropy modeling approach combined with GIS technology in China. 99 occurrence records and 11 Worldclim environmental variables were used to conduct the analysis. Model was evaluated with both ROC and AUC for which the results show the value of 0.9 showing the model fitness. The current suitable area for the species was found to be located in Sichuan (2.74%), Yunnan (2.09%), Shaanxi (1.82%), Hubei (1.76%), Guizhou (1.66%), Henan (1.66%) and Hunan (1.64%). Xinjiang was the one with the largest share of the species presence which was about 9.70%. The maxent results for future distribution (2050 and 2070) shows the increasing trend of safflower for poor and moderate regions under the current scenario. Though, there is decline in highly suitable areas for RCP2.6, RCP4.5 and RCP8.5 in 2050s and two RCPs, RCP2.6 and RCP4.5 in 2070s.

Qin et al., (2017) used Maxent modeling approach to predict the climatic impacts on the geographical distribution of Thuja sutchuenensis Franch, which is known to be an endangered conifer in southwestern China. This research took 3 years to complete with 107 presence locations in Daba Mountains. The model showed high accuracy with an AUC of 0.998. The model showed the location of the potential geographical distribution for the last interglacial period is in southeastern China, having an optimum habitat area of 1666 km^2. In future predictions of the model show optimal habitat area is outside the current distribution.

Another study using the Maximum Entropy approach was conducted to find the future trends in the distribution of plant habitat and effective in central Iran. GIS technique coupled with Geostatistics facilities was utilized to develop the maps about environmental variables. Area under the curve was used for the assessment of the accuracy of the model which indicated good accuracy for the species except Artemisia sieberi with high frequency. The future maps for species like Artemisia aucheri, Scariola orientalis, Astragalus albispinus, A. sieberi2 and A. sieberi, Zygophyllum eurypterum depicted fair arrangement with the corresponding observed maps except for S. orientalis, A. sieberi and Tamarix ramosissima predictive maps which showed low conformity rate of prediction and observed maps. MaxEnt modeling approach was considerably outstanding for the prediction of distribution about the plant species habitat with narrow ecological niches (Chahouki et al., 2016).

Species distribution modeling approach was used for the evaluation of distributional range for T. matsutake under current scenario in China. The software used for the purpose was Maxent in which species presence data along with 24 variables was used. This study was done for two future years 2050s and 2070s for all climate change scenarios RCP 8.5, RCP 6, RCP 4.5 and RCP 2.6. This study concluded that the area with suitable habitats would endure moderately minor changes under all RCPs; though, areas with suitable habitat would considerably decrease, and areas with high suitability of the species presence would nearly disappear. This research can be used as a reference for the ecological conservation and management of this species (Guo et al., 2017).

A study was conducted for the mapping of shrinking population for Malabar nut (A medicinal plant) in future due to over population of other species in the area along with high population of humans in the area and urbanization. This research was purposed to get the clear picture of the species in the future. The study area for this research was Lesser Himalayan foothills in India (Dun valley) and the Species distribution modeling software used for the purpose was Maxent. The study area was spread on about 1877 square kilometers for which 46 occurance points were taken. The variables used in this study were slope, elevation, land use and the bioclimatic variables from worldclim. To evaluate and analyze which variables are important, jackknife testing was done and to evaluate the accuracy of the model, AUC was seen which showed the accuracy of the model to the value of about 92.3%. The method of maxent modeling is an efficient tool for the prediction of the distribution of various species including the medicinal plants and it can be proven very effective in terms of conservation of biodiversity (Yang et al., 2013).

2.5 Maxent Modeling for Crops

Species distribution modeling is not only restricted to medicinal plants and animals but also it has played a good role in the prediction of various crops suitable habitat for future. A similar study on Ceylon spinach was conducted in India to generate the climate suitability maps for the species. The SDM used for the purpose was MaxEnt and the variables used in the study were bioclim variables. The results obtained after conducting this study show that the suitable habitat zone for the species is in slightly southward areas which include parts of Tamilnadu, Pondicherry, Maharashtra, Orissa, West Bengal, Bihar, Madhya Pradesh, Uttar Pradesh, Rajasthan and Gujarat. The training and testing data for AUC showed the values of 0.95 and 0.99 depicting the excellency of the model and showing the model suitability in the prediction of this species in India (Reddy et al., 2015).

Previously, published data along with geographical information and national climate was incorporated in Maxent to analyze the distribution of winter wheat in China to see the suitable climate zones for the crop. Maps for winter wheat zones were generated and climate suitability for this species in the area was seen. According to the previous studies for winter wheat in this region, northeast boundary of the winter wheat cultivation zone is the south of Liaoning Province, and our study has shown a different analysis indicating that the suitable area for the wheat cultivation lies in the north of Heilongjiang and the suitable northwest boundary for the winter wheat is north of Xinjiang Uygur Autonomous Region (Song et al., 2012).

Evaluation of the climatic suitability for the distribution of maize crop was done in China and changes were analyzed at 1.5 °C and 2.0 °C temperature rise in future in accordance to the temperature targets in Paris Agreement. A comparison was made from 1971–2000 and there is a shift of patterns for summer maize to east and for both temperatures, the habitat suitability tends to decrease for this crop which indicates that global warming has negative impact on summer maize production in China. Due to the global warming in future, the optimal zones for maize cultivation move northeastward under RCP4.5 and RCP8.5. In general, half-a-degree more global warming drives the cultivable areas of summer maize to shift northward in China, while the west region shows a certain potential for expansion of summer maize cultivation (He at al., 2019).

2.6 Maxent Studies in Pakistan

Use of Maxent for the prediction of suitable habitat is becoming common in developing countries like Pakistan. Fatima el al., (2016) utilized the application of maxent model for the estimation of spatial patterns of dengue (*Aedes aegypti*) in two regions, Lahore and Swat. The dispersal of *Ae. aegypti* was determined through the application of MaxEnt on environmental variables and occurance records. The ecological dependence of dengue mosquito on every single environmental factor was analyzed through response curves. Results suggested the wide distribution of specie in Lahore. The main reasons for its wide population in the area are population density and infrastructure. In the case of Swat, dengue is found in the form of clumps.

Khanum and his coworkers (2013) utilized the Maximum entropy (Maxent) modeling technique to find out the suitable niches for three plants with medicinal properties including Pentatropis spiralis, Tylophora hirsuta, and Vincetoxicum arnottianum. Occurrence records for the study were taken from herbarium specimens present in different herbaria of the country along with field data for year 2010 and 2011. The evaluation of Maxent showed AUC value of 0.74 for P.

spiralis, 0.84 for V. arnottianum, and 0.59 for T. hirsuta. According to the results of the model there would be gain in the habitat for first two species in the provinces of Punjab and Balochistan and there would be loss of habitat in Sindh. The predictions about Vincetoxicum arnottianum as and T. hirsuta show a gain in the habitat in the upper northern parts of Pakistan. Future maps for T. hirsuta show the loss of habitat in northern Punjab and in the areas of lower peaks of Galliat, Zhob, Qalat etc.

The potential habitat distribution of *Olea ferruginea* was predicted by Ashraf et al., 2016 using the method of maxent for current and future climate (2050). Potential distribution *Olea ferruginea*, an economically important plant, was assessd using bioclimatic variables (both current and future climate), digital elevation model (DEM) slope and occurrence location data. Using 219 occurrence points in their study, they achieved an AUC of 0.98 which showed model performed much better than average. The study determined a substantial impact of future climate scenario on *Olea ferruginea* under global climatic changes. A considerable reduction in the suitable areas of *Olea ferruginea* was noticed under current climate but the model predicted an increase in the suitable areas in the future climate at higher altitudes, a phenomenon known as habitat shift. The study recommended to make use of the potential suitable areas of *Olea ferruginea* that are predicted and restoration of deforested lands.

CHAPTER 3: MATERIALS AND METHODS

3.1 Study Area

The study area for this research is Pakistan with latitude 30.3753° N and longitude 69.3451° E. It is located in South Asia with area of about 881,913 Km^2. It is bordered by countries like India, Afghanistan, Iran and China. It also has coastline along Arabian Sea in the South. Pakistan is a country with diversified temperature and precipitation patterns. The climate has got both arid and semi arid regions and most of the country receive the rain less than 250 mm every year except in the southern slopes of Himalaya andnorthern parts of Pakistan, where annual rainfall ranges from 760 mm to 2000mm. There are four distinct seasons in Pakistan, spring, summer, autumn and winter. Pakistan's economy is mostly based on farming but due to climate change is negatively impacting the plant growth and agriculture of the country. In Pakistan, potato is a widely grown crop as it gives high return to the farmers. Potatoes are widely cultivated in Pakistan wherever irrigation is available. Punjab leads in potato production with the share of 83% followed by KPK with the share of 10%. Baluchistan contributes 6% in the potato production whereas only 1% potatoes are cultivated in Sindh.

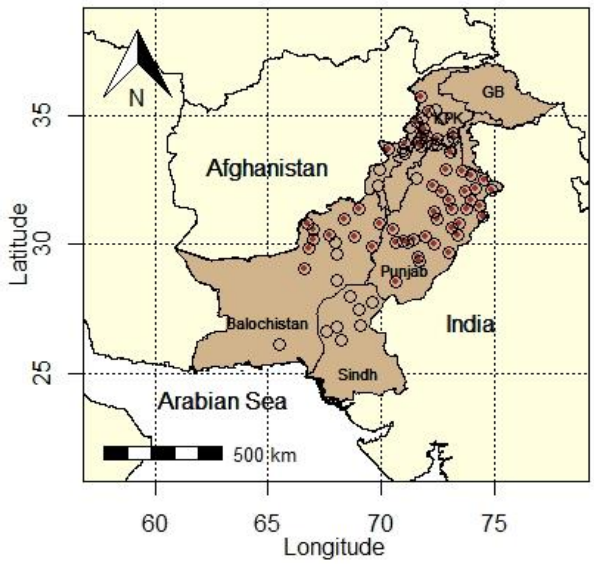


Figure 3.1: Study Area with Potato Presence Locations, where black circles show all the data points and closed brown circles show the final data points after analysis.

3.2 Methodology

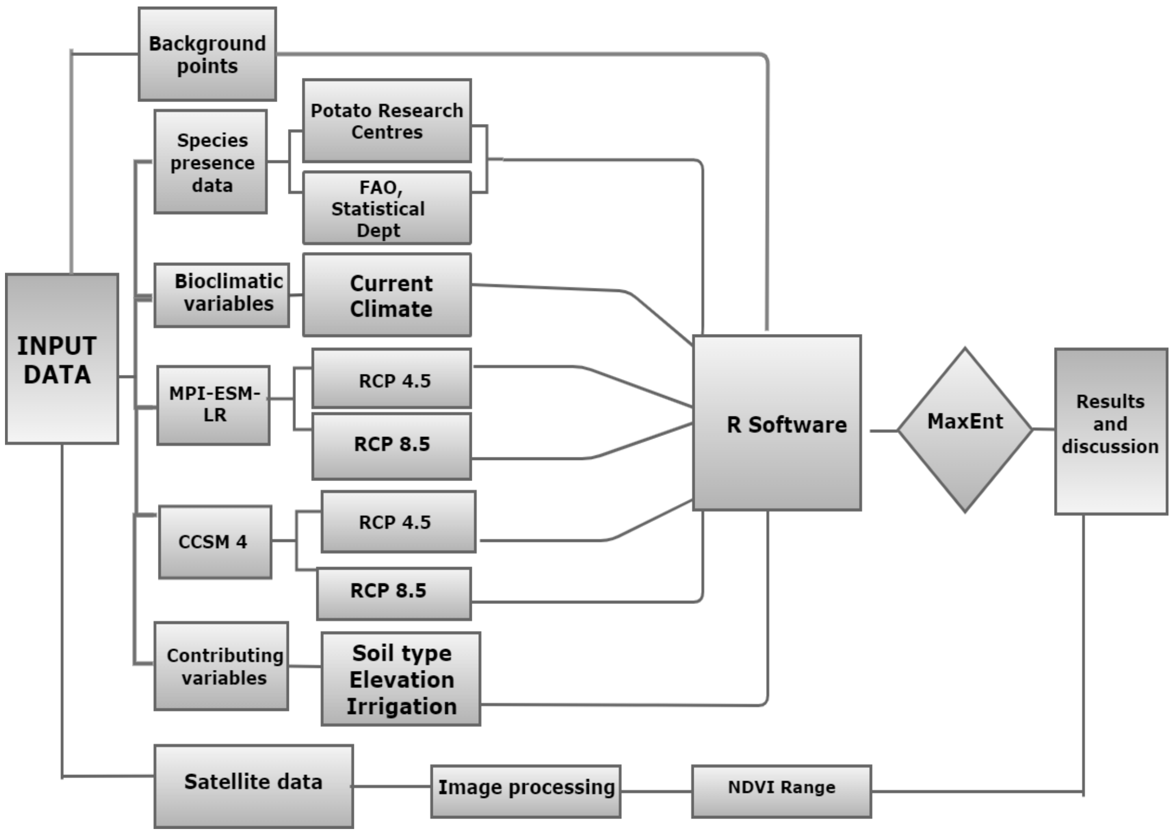


Figure 3.2: Research Design for Maxent Modeling

3.2.1 Occurrence data collection and processing

Collection of Georeferenced occurrence records for *Solanum Tuberosum* was done primarily from data books of Directorate of Agriculture Pakistan (2010-2018), unpublished data have been taken from Potato Research Centers of Sahiwal, Sialkot and Murree. A few GPS points were gathered during field visits to Abbottabad, Mansehra and Feteh Jang in October and

November, 2018. GPS records from potato research centers were checked against high-resolution satellite images using Google Earth Pro. The collected data from literature was in raw form and it was further processed, the areas with above average of last 10 years of yield of potato where potato was grown every year were considered and the areas with low yields were deleted. Average was taken for all the years and production/area above average were considered for modeling. The total occurrence data points were 58 in total as mentioned in the figure 3.1.

3.2.2 Environmental data collection and processing

19 BioClim variable layers (Bio_1 to Bio_19) for two models (CCSM4 and MP-ESM-LR) were obtained from worldClim website (www.worldclim.org). This website provides current (1960-1190), high resolution worldClim climate data and future climate projections from global climate models (GCMs) for four representative concentration pathways (RCPs). The future projections we selected for the analysis along with current bioclimatic layers include two climate change scenarios (RCP 4.5 and RCP 8.5). Each layer had a spatial resolution of 30s and was converted to ASC format in R studio as required by MaxEnt (Phillips et al., 2006). Additional covariates were also downloaded from FAO website (www.fao.org) including Soil type and Irrigation, whereas Elevation was downloaded from USGS website (usgs.gov). These covariates are incorporated in the analysis along with the bioclim layers to evaluate their impact on model outcome.

3.2.3 Satellite data download and processing

Satellite images of Landsat 8, level 1 for October and November, 2018 were downloaded from USGS website. The images downloaded were for whole Pakistan. Total tiles downloaded were 66 in Tif format. The purpose of satellite data was to validate the exact locations of potato production using NDVI range. For this reason, band 4 and band 5 from all the files were extracted and separated. Mosaicking of these layers was done using linux operating system. Cloud cover was removed and further atmospheric correction was done and thus, NDVI for the final product was calculated. The final mosaicked file was further clipped by masking using R software. Classification was done to get the locations of potato using NDVI range extracted from literature review. Map for NDVI range for potato was generated and thus validation was done against the occurrence data.

3.2.4 Maxent Modeling and Evaluation

Occurrence data file in "txt" format along with 19 current bioclimatic variable and covariates (soil type, irrigation and elevation) ASCII grid files with same resolution were entered in - R with the installation of required packages including maxent package. 58 presence point along with random 1000 background points were added into the software to generate the suitable area maps. For the cross-validation 5-fold method has been used in which the model randomly divides data into five independent sets and repeated the process for five times using a different set for the modeling test and the rest to train it, that results the prediction with average of 5 models resulting predictions being the averages over all 5 models. Most contributing variables in the model were obtained with the help of codes in R and then the maps for suitable habitat for current climatic conditions were generated. Same procedure was repeated for future scenarios includingRCP 4.5 and RCP 8.5 with future bioclimatic variable layers but the covariates were kept constant. Maps for future scenarios were generated and suitable areas for potato production were calculated for current climate and future scenarios using raster package. QGIS was used to extract the shape files for provinces which were then incorporated in R-studio to mask on the generated maps for suitable area and then the codes were applied on these masked maps to calculate the area for individual provinces of Pakistan. Model was evaluated on the base of area under receiver operating characteristic curve (AUC) which is a threshold-independent statistic commonly used to evaluate species distribution models.

Assessment of maxent predictions are done using ROC which is receiver operating curve. This plot comes up after we use the sensitivity values (true positive) on y axis against the equivalent values (false positive) on x-axis. The undelying area in ROC is called AUC which is an important matrix as it provides the accuracy of the model. The value of the AUC ranges between 0.5 and 1.0. The higher the value, the better themodel predicted the distribution of the species (Fielding and Bell, 1997). If AUC> 0.78, the model performs well (Peavey, 2010). Maxent calculated AUC values for both training and test data. AUC values range from 0.5 to 1.0 and model accuracy is classified as suggested by Swets (1988) as AUC > 0.9 = excellent; 0.8 to 0.9 = good; 0.7 to 0.8 = fair; 0.6 to 0.7 = poor; 0.5 to 0.6 = fail.

3.2.5 Variable Importance for Current Climate Distribution

To have a look on the importance of different variables in the study Jacknife test has been used. Number of models are generated with the help of this test. The basic principle of this test is to exclude one variable and creation of model with all the other variables. Then a model is created using each variable in isolation. In addition, a model is created using all variables, as before (Song et al., 2012). The importance of the variables is denoted with pink and blue bars.

3.2.6 Pearson Correlation

Pearson correlation is a the test that measures the association and link among two variables. It also explains the level or magnitude of association and tells if the association is positive or negative. (Kumar et al., 2009). The multicollinearity test was conducted by using Pearson

Correlation Coefficient (r) to examine the correlation of the most important variables with other variables (fig 4.9). After running this test variables correlation with other variables were found that contributed to this study.

CHAPTER 4: RESULTS AND DISCUSSIONS

This chapter is based on the results and the outcomes of Maxent Modeling applied on Potato (*Solanum Tuberosum*) crop in Pakistan.

4.1 Presence Locations of Potato

Potato is widely distributed in Pakistan, the data of 9 years obtained from Pakistan Bureau of Statistics, Food and Agriculture Organization and Potato Research Centre (Sahiwal, Sialkot and Murree) was further processed into average yield (Tonnes/Hectares) and locations (Latitude and Longitude) were extracted for the location where potato production is above the calculated average in order to extract data for areas with rich crop cultivation and production . The presence locations for Potato (*Solanum Tuberosum*) are shown in the following figure 4.1 where black circles show all the data points and closed brown circles show the final data points after selection of areas with maximum yield .

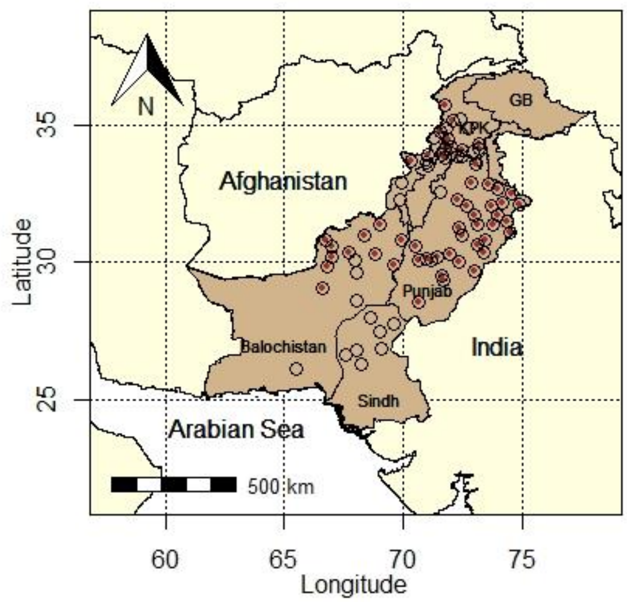


Figure 4.1: Processed presence locations of Potatos where black circles show all the data points and closed brown circles show the final data points after analysis.

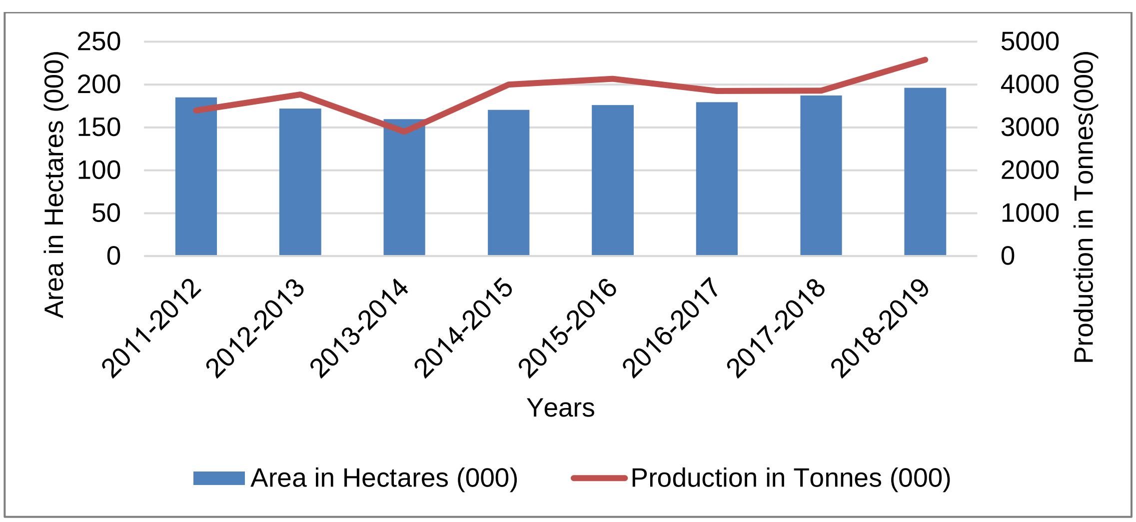


Figure 4.2: Area under potato production for the year 2011-2019

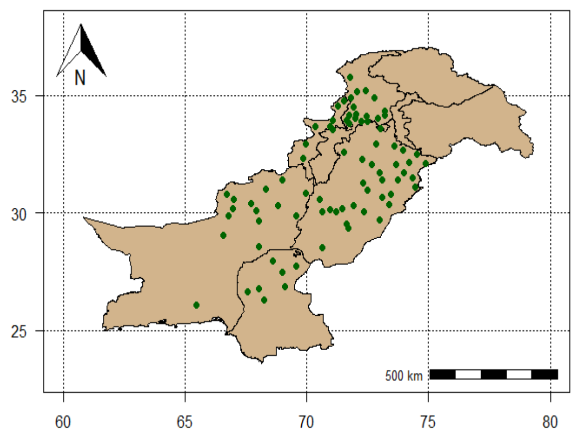
Figure 4.2 presents the graphical representation of the total area under potato production. The figure depicts the fluctuations in the areas and production of potato. Changing climate has a correlation with low yield and poor development of the potato crop (Ahmed et al., 2020). There are abiotic factors that reduce to the production of potato including water scarcity, salinity, heat stress and temperature fluctuations and nutrient deficiency in soils. Moreover, there is less area allocated fo the potato crop harvesting as compared to the other crops which is the major cause of low yields. (Majeed and Zahir, 2018). Thus to meet the growing needs of consumers and growing population demands with such level of production is a major challenge and can lead to food insecurity in the future.

4.1.1 Potato Presence Locations (Field Based and Satellite Based)

Satellite data was downloaded and processed to calculate the NDVI of the crop, the range of NDVI was found to be (0.2-0.8) (Kundu et al., 2018). This aspect was used in the research to validate the presence locations of potato obtained from acquired data from Pakistan Bureau of Statistics.

Figure 4.3 (a) shows the field based locations of Potato and figure 4.3 (b) NDVI based locations of potato crop in Pakistan. NDVI based map is validating the presence locations of the potato in field based map except for some area in Sindh Province. These areas might have some other crops or plants such as Maize (NDVI Range 0.25-0.75) which falls under the same broad range of potato and widely grown in Sindh in the same -climate (Arain, 2013).

(a)



(b)

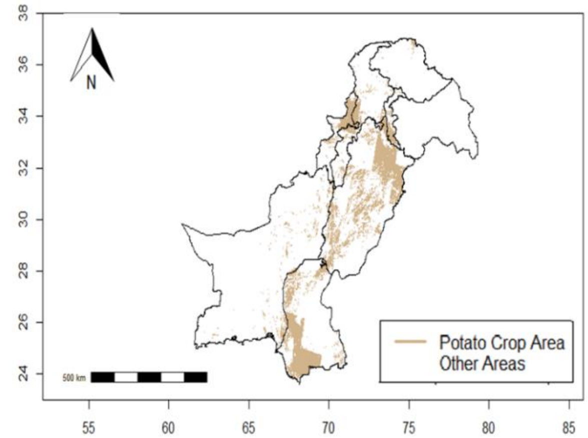


Figure 4.3 (a): Total presence locations for field data of potato crop in Pakistan (b) NDVI based locations for potato crop in Pakistan

4.1.2 Potato Current Distribution using Maxent

The existence probability of Potato was calculated using the Maxent model. According to the accumulative output of original results from the Maxent model, the raw map (figure 4.4 (a) was generated showing the highly suitable, moderately suitable, least suitable and non suitable climate areas for potato production in Pakistan.

(a) (b)

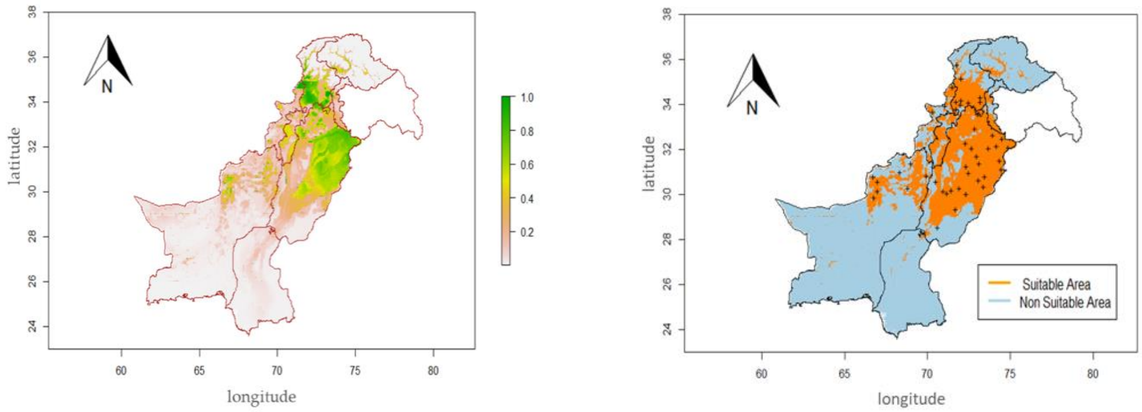


Figure 4.4: (a) Raw Map for Current Climatic Distribution of Potato in Pakistan. (b) Threshold Map for Highly suitable climate regions lie under the range of 0.75-1.0

Another method was used for the evaluation in this study was to define the thresholds. In this method selection of a threshold for the establishment of the sites is done and this the sites are divided into suitable and unsuitable areas for the growth of particular species. These thresholds are established by maximizing sensitivity while minimizing specificity (Phillips et al., 2006). The threshold map was generated to simplify the output of Maxent model and present it in the binary form (Only suitable and Non suitable Climate locations) shown in figure 4.4 (b). Furthermore, Suitable area was calculated based on the threshold raster map. The area function in R under the raster package was used and resulted in a vector of the cell sizes in the raster object (in km^2), which differ from north to south. The entire suitable area size was obtained by adding up all cell sizes of same color or multiplying the median cell size by the number of cells. The area calculated as suitable climate for potato production under current scenario is about 265,016 Km^2.

4.1.3 Provincial Breakdown of Suitable Climate for Potato under Current Climate

Area calculation was done based on the provinces of Pakistan and Punjab was found to be a province with maximum share of suitable area to grow potato. The suitable area for Punjab is about 152411 Km^2, followed by KPK with the area of 56392 Km^2. The share of Balochistan and Sindh is about 42509 Km^2 and1239 Km^2. Suitable area under the territory of Gilgit Baltistan is about 5534 km^2. Out of all the provinces, Sindh has the least suitable climate to grow potato as the temperature and climate of Sindh even in winters is not suitable for Potato cultivation shown in figure 4.5.

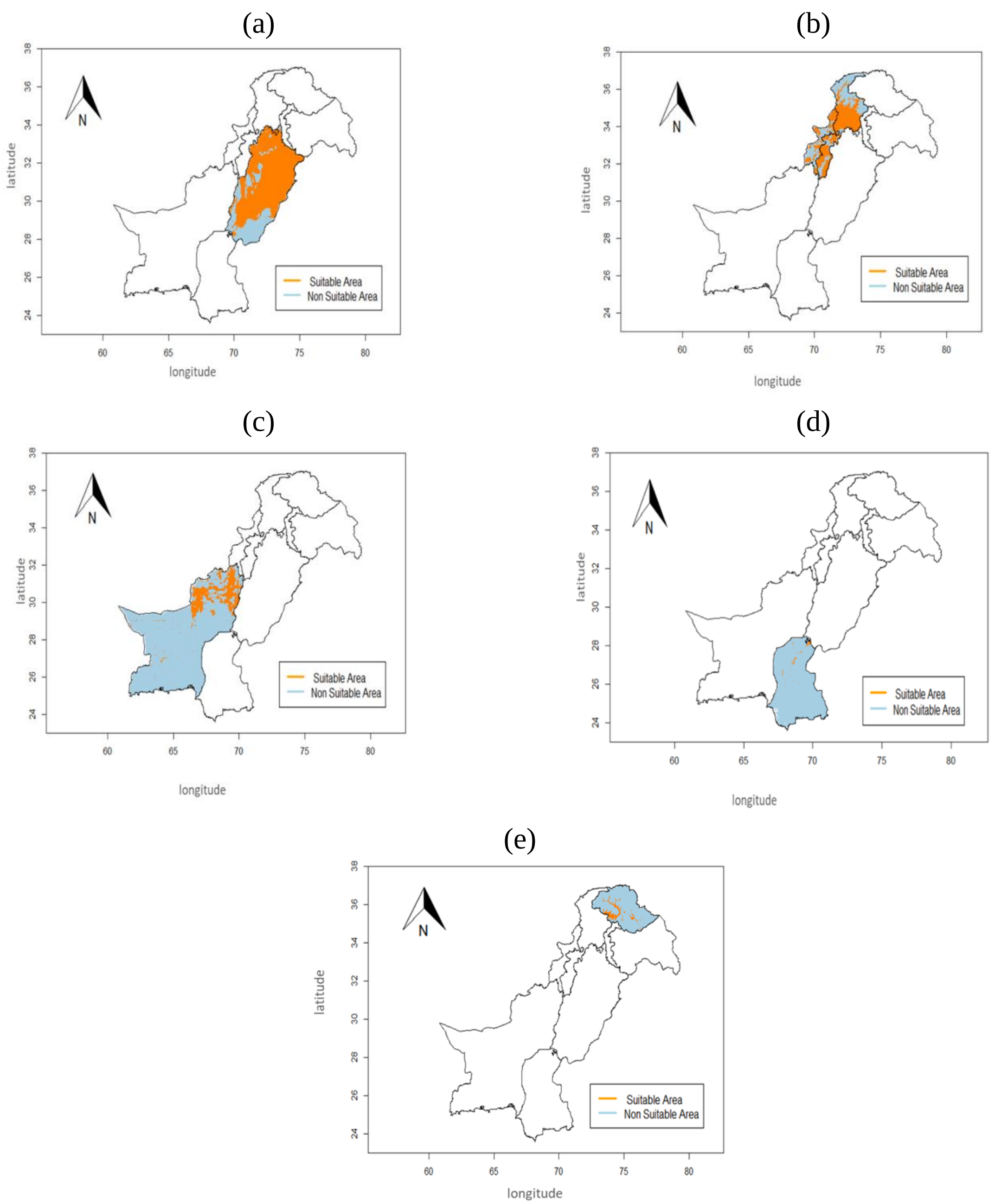


Figure 4.5: Provincial area breakdown for potato crop under current climate (a) Suitable areas for potato cultivation in Punjab (b) Suitable areas for potato cultivation in FATA and KPK (c) Suitable areas for potato cultivation in Balochistan (d) Suitable areas for potato cultivation in Sindh and (e) Suitable areas for potato cultivation in Gilgit Baltistan

4.1.4 Maxent Model Evaluation

In the case of current distribution of potato using Maxent, the results of AUC show that the model performed well and simulated results were good enough.

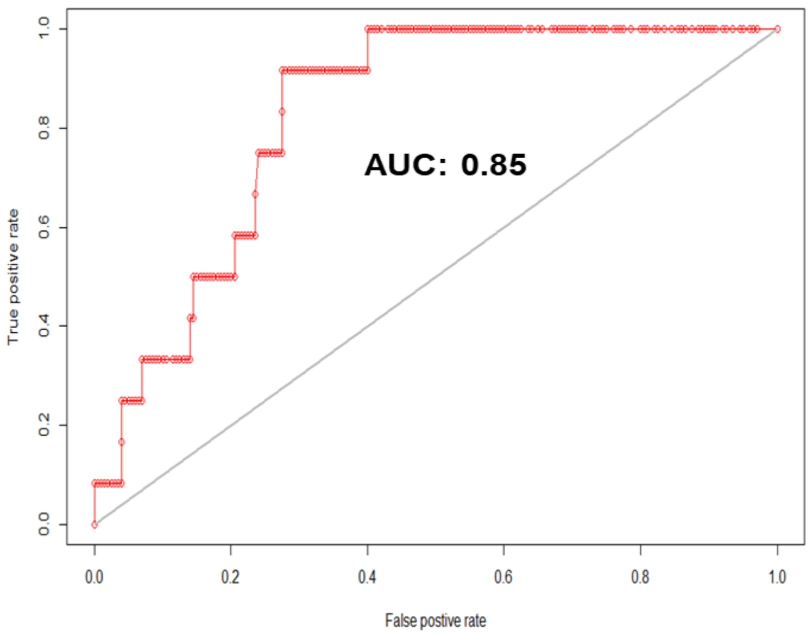


Figure 4.6: Area Under Curve for Current Distribution

4.1.5 Variable Importance for Current Climate Distribution

To have a look on the importance of different variables in the study Jacknife test has been used. Number of models are generated with the help of this test. The basic principle of this test is to exclude one variable and creation of model with all the other variables. Then a model is created using each variable in isolation. In addition, a model is created using all variables, as before (Song et al., 2012). The importance of the variables is denoted with pink and blue bars. The jackknife test is shown in figure 4.7. After running Jackknife for variable importance, the most important variables were found to be isothermality, precipitation of wettest quarter and annual precipitation.

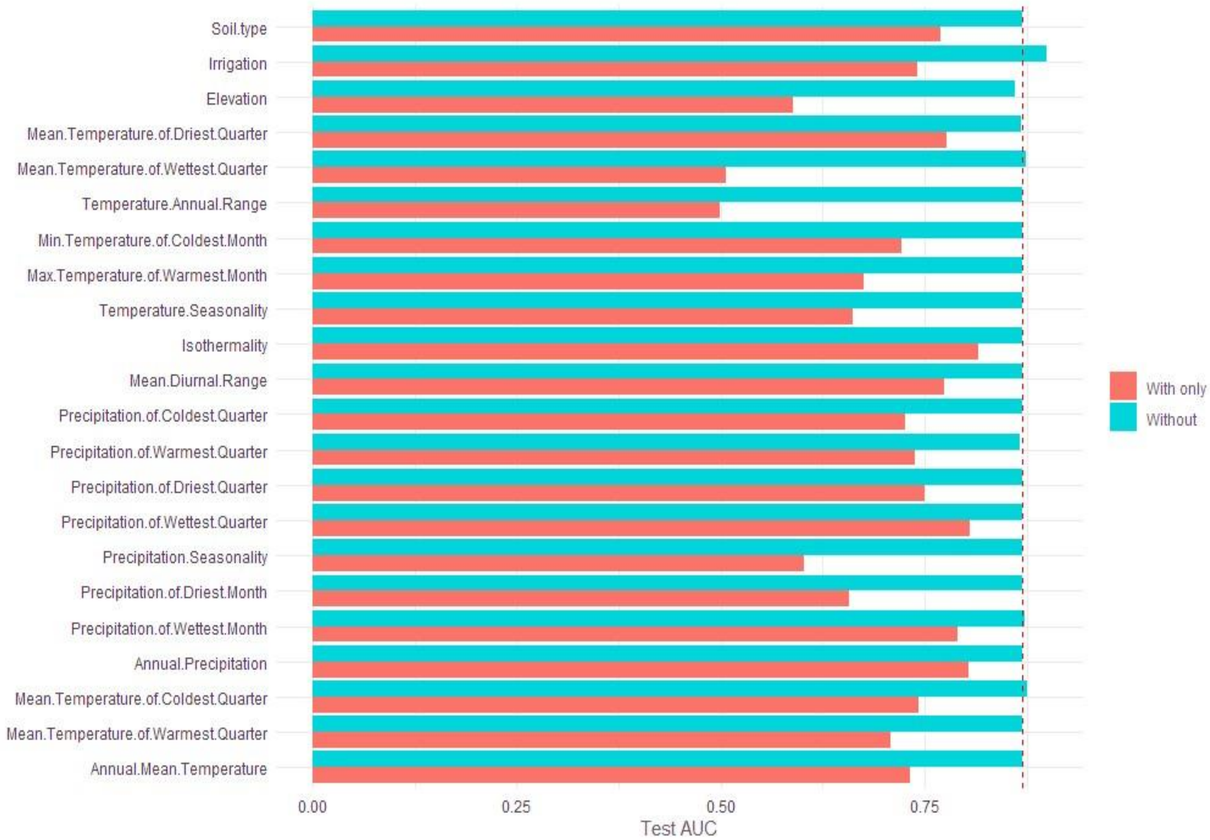


Figure 4.7: Jackknife test of variable importance for current distribution of Potato.The most important variables were found to be isothermality, precipitation of wettest quarter and annual precipitation

4.1.6 Pearson Correlation (Current Potato Distribution)

The multicollinearity test was conducted by using Pearson Correlation Coefficient (r) to examine the correlation of the most important variables with other variables (fig 4.9). According to the positive and negative values of correlation, isothermality has highest positive correlation value with mean temperature of driest quarter and is highly negatively correlated with temperature seasonality. Precipitation of wettest quarter is in positive correlation with precipitation of warmest quarter and in negative correlation with mean temperature of driest quarter. Another important variable annual precipitation is found to positively correlate with precipitation of wettest quarter and negative correlation was seen with mean diurnal range.

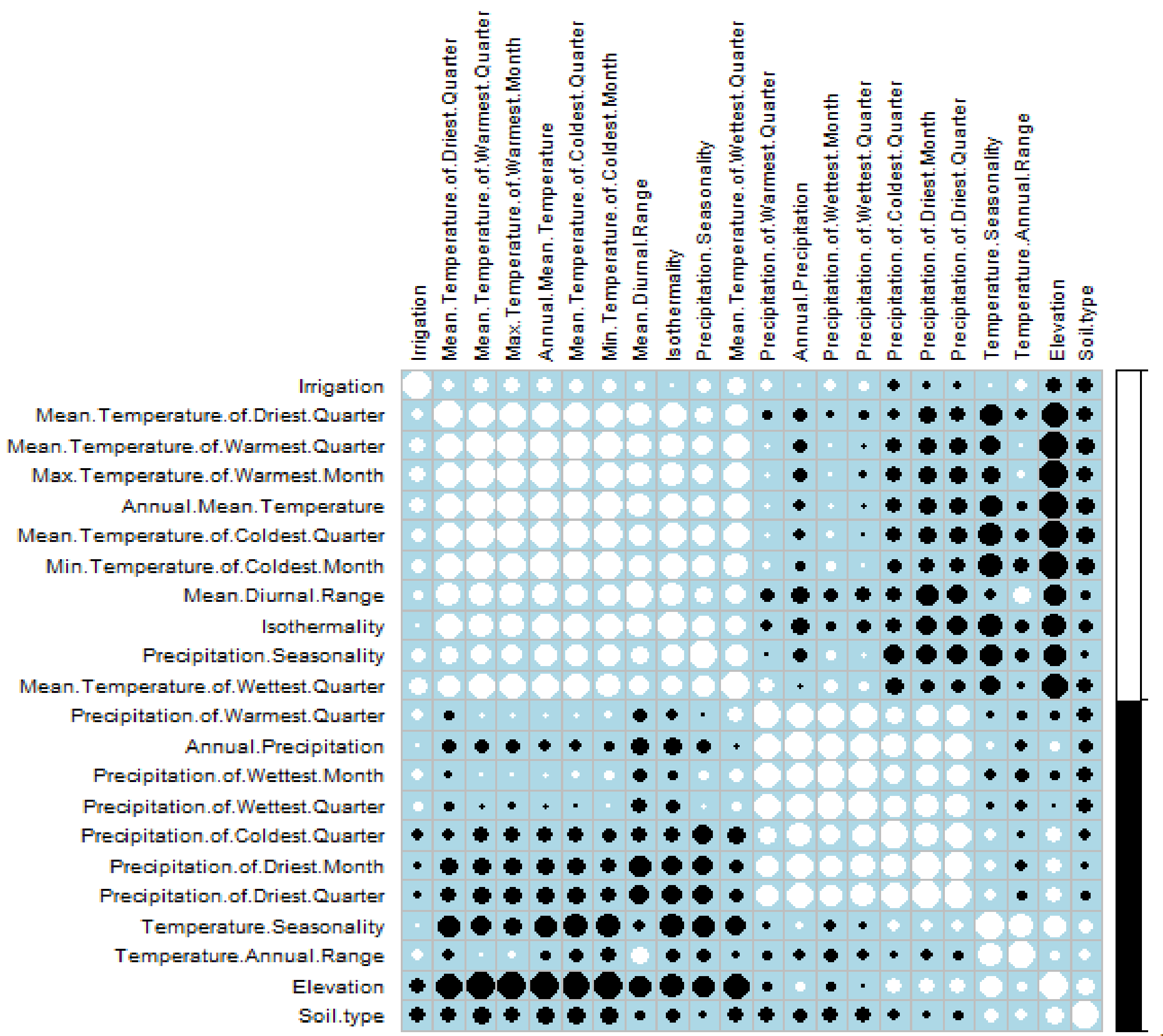


Figure 4.8: Pearson Correlation among Bioclim and other variables for Current Climate distribution

4.2 Potato Predicted Distribution under RCP 4.5 (Model 1)

A CMIP5 model MPI-ESM-LR was used as Model 1 for this study and future distribution of potato was predicted for 2070 under RCP 4.5. RCP 4.5 is a moderate scenario where there is going to be a stabilization of temperature by 2100 which means the radiative forcing level stabilizes at 4.5 W/m^2. The results for RCP 4.5 show a slight decrease in the suitable area for potato in comparison with the area shown in current distribution.

4.2.1 Potato Predicted Distribution using Maxent (RCP 4.5)

In this scenario, reduction in predicted suitable area has been noted for potato cultivation. Map for the raw values of Maxent future distribution for potato under RCP 4.5 has been shown in the figure 4.9 (a) .

(a) (b)

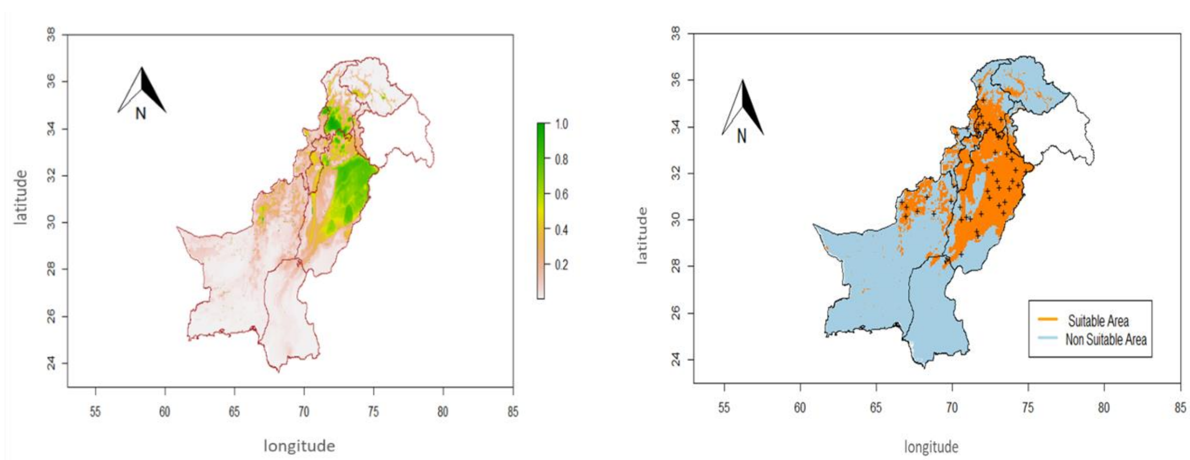


Figure 4.9: (a) Raw Map for Future Climatic Distribution (RCP 4.5) and (b) Threshold map of Potato in Pakistan

Threshold map (figure 4.9 (b)) was generated to simplify the output and show the suitable climate for potato in more simpler and binary form. The suitable area calculated was about 252449 Km^2 which mostly comprises Punjab and KPK. This area is slightly less than the area calculated for present distribution of potato crop.

4.2.2 Provincial Breakdown of Suitable Climate for Potato under RCP 4.5

Area calculation was done based on the provinces of Pakistan and Punjab was found to be a province with maximum share of suitable area to grow potato in future, though the areas for provinces are predicted to be reduced compared to current suitability.

Area suitable in Punjab figure is found to be 145200 Km^2 (figure 4.10 a) followed by KPK (figure 4.10 b) which is 54390Km^2. Suitable climate in Balochistan (figure 4.10 c)is predicted to be 36602 Km^2 and Area under Sindh (figure 4.10 d) is going to be about 2723Km^2. The share of Gilgit Baltistan figure (4.11 e) is about 6193 Km^2.

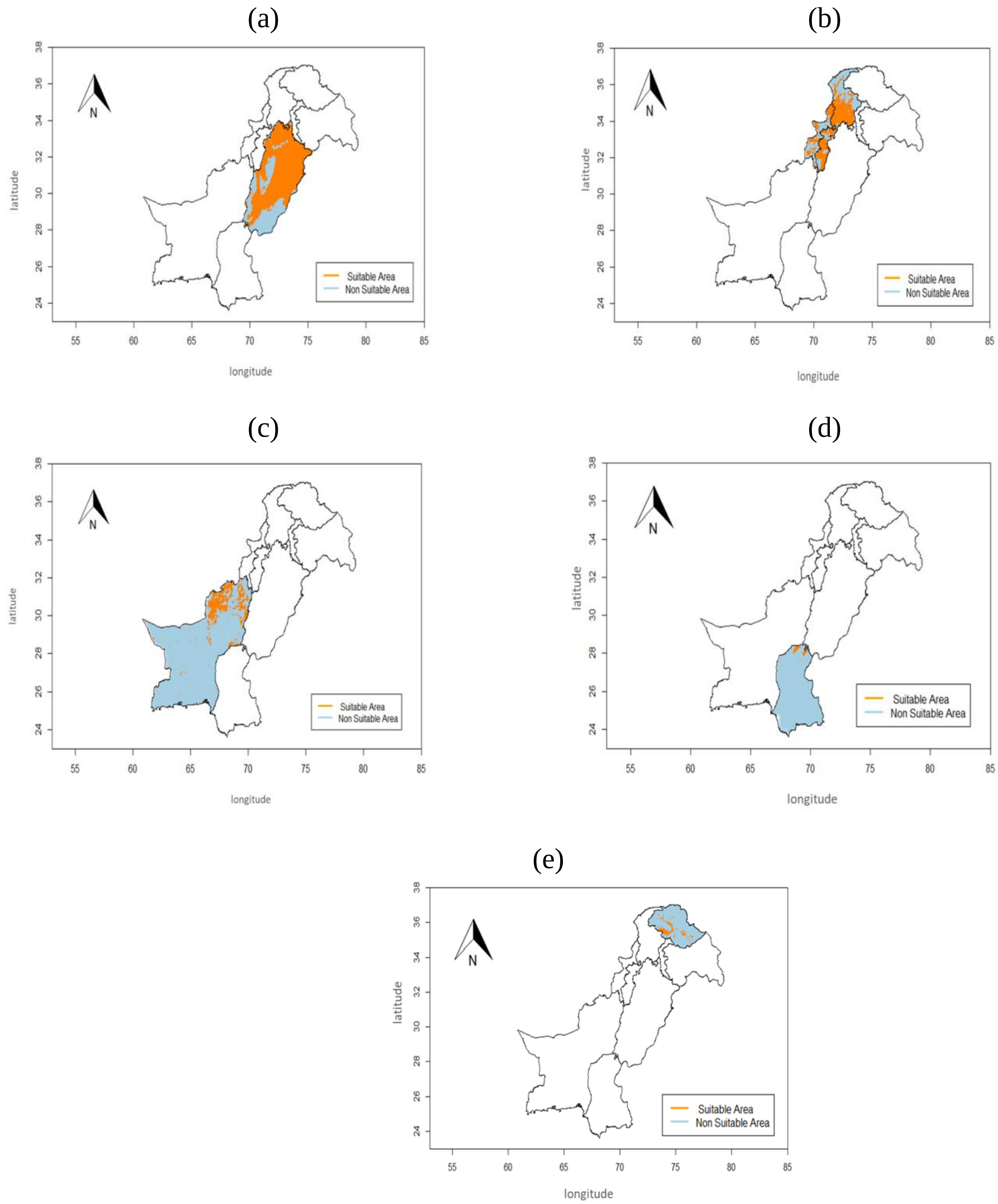


Figure 4.10: Provincial area breakdown for potato crop under RCP 4.5 (a) Suitable areas for potato cultivation in Punjab (b) Suitable areas for potato cultivation in KPK and FATA (c) Suitable areas for potato cultivation in Balochistan (d) Suitable areas for potato cultivation in Sindh and (e) Suitable areas for potato cultivation in Gilgit Baltistan

4.2.3 Maxent Model Evaluation

AUC found for RCP 4.5 showed good performance of our model with the value of 0.82.

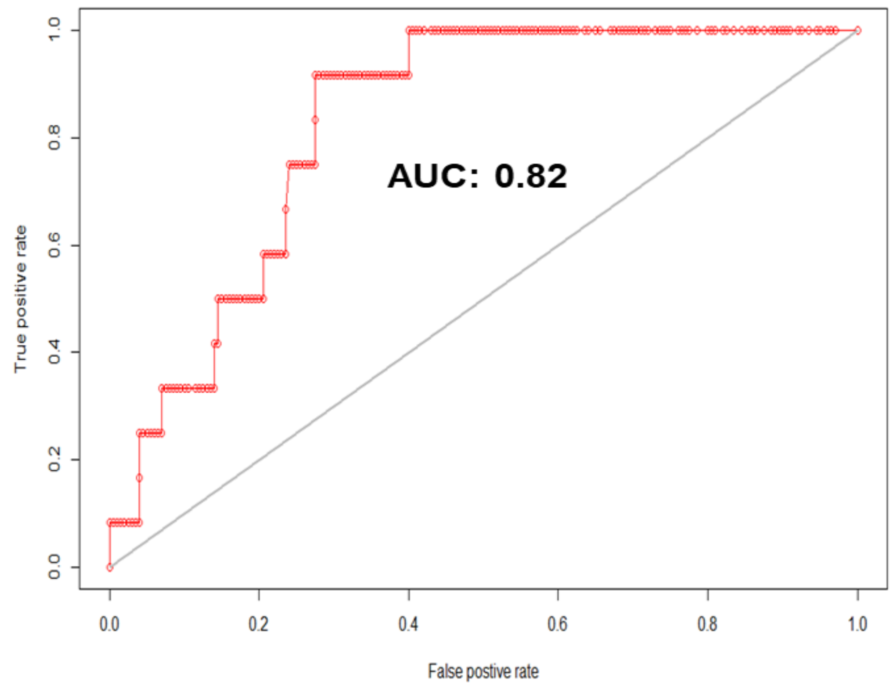


Figure 4.11: Area Under Curve for RCP 4.5

4.2.4 Variable Importance for RCP 4.5 Distribution

After running Jackknife for variable importance, the most important variables were found to be precipitation of coldest quarter and annual precipitation.

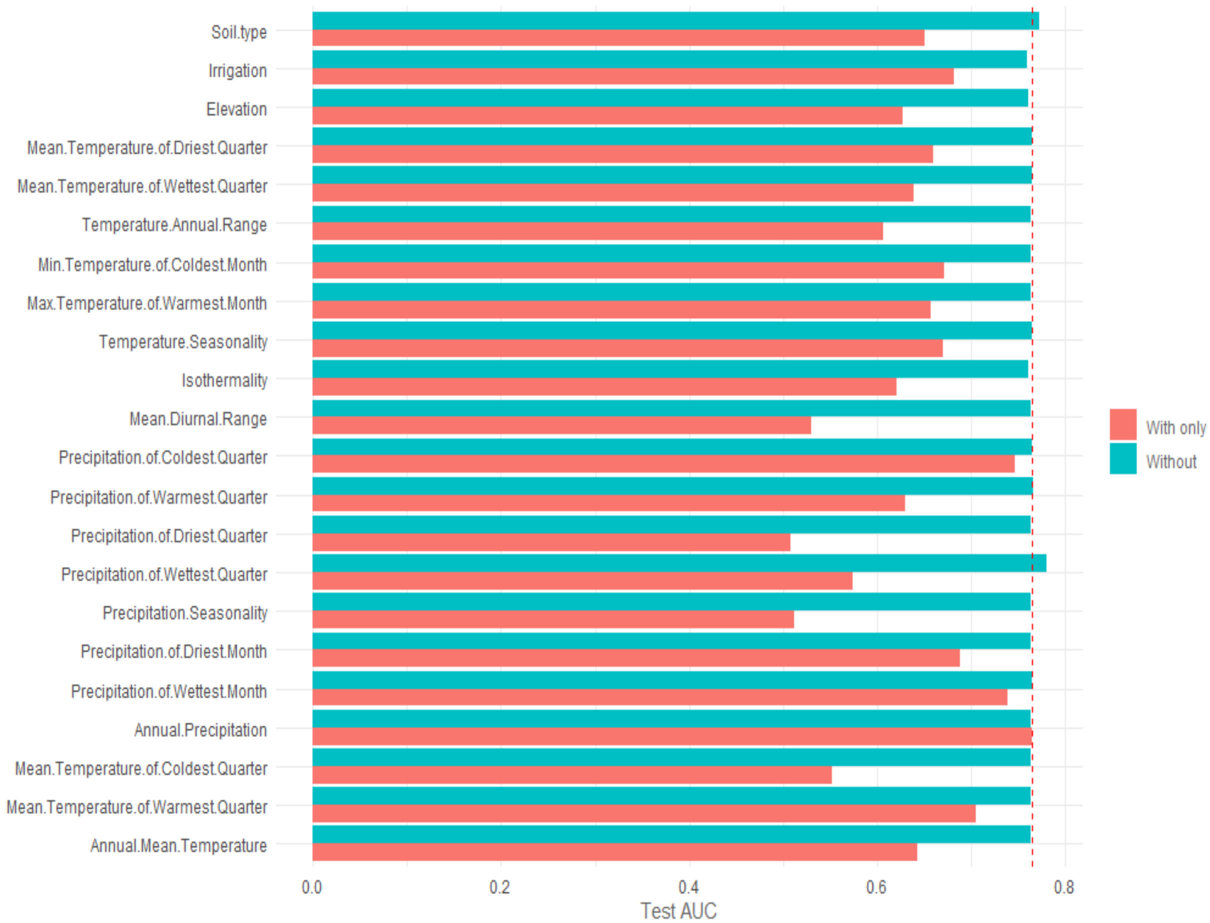


Figure 4.12: Jackknife test of variable importance under RCP 4.5, the most important variables were found to be precipitation of coldest quarter and annual precipitation

4.2.5 Pearson Correlation (RCP 4.5)

The multicollinearity test was conducted by using Pearson Correlation Coefficient (r) to examine the correlation of the most important variables with other variables. According to the positive and negative values of correlation, Precipitation of coldest quarter is in positive correlation with isothermality and in negative correlation with mean temperature of driest quarter. Another important variable annual precipitation is found to positively correlate with precipitation of wettest month and negative correlation was seen with mean temperature of driest quarter.

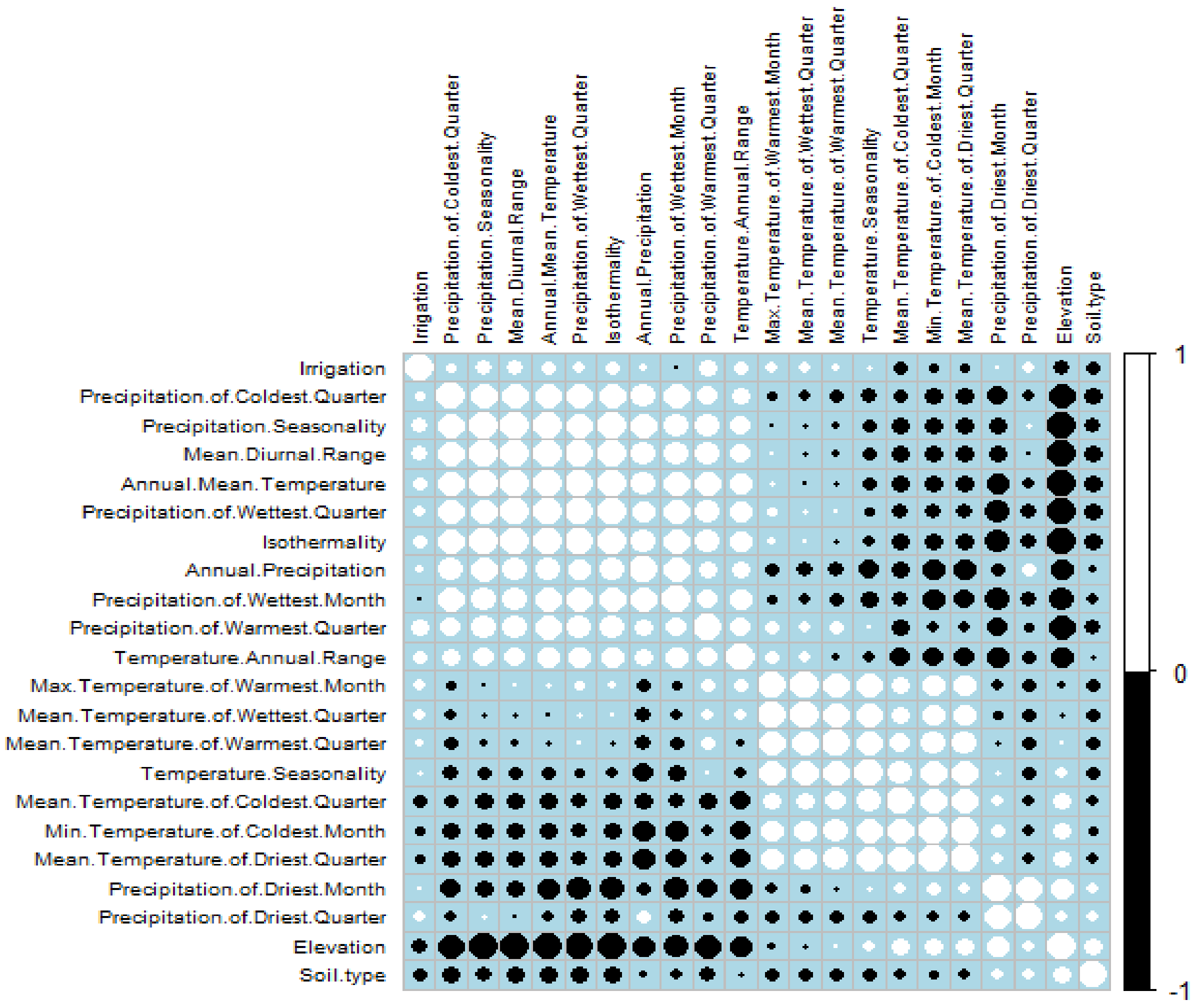


Figure 4.13: Pearson Correlation among Bioclim and other variables for RCP 4.5 distribution

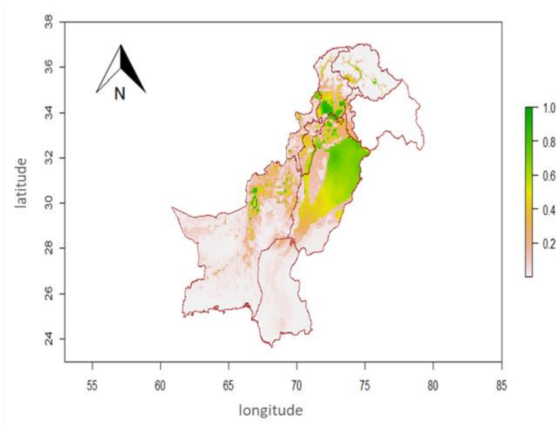
4.3 Potato Predicted Distribution under RCP 8.5 (Model 1)

RCP 8.5 is an extreme scenario where temperature will keep on rising even after 2100 with emission level leading to a forcing level of around 6-7 W/m^2. The results for RCP 8.5 show a drastic decrease in the suitable area for potato in comparison with the area shown in current distribution and in RCP 4.5

4.3.1 Potato Predicted Distribution using Maxent (RCP 8.5)

In this scenario, drastic reduction in predicted suitable area has been noted for potato cultivation. Map for the raw values of Maxent future distribution for potato under RCP 8.5 has been shown in the figure 4.16 (a)

(a)



(b)

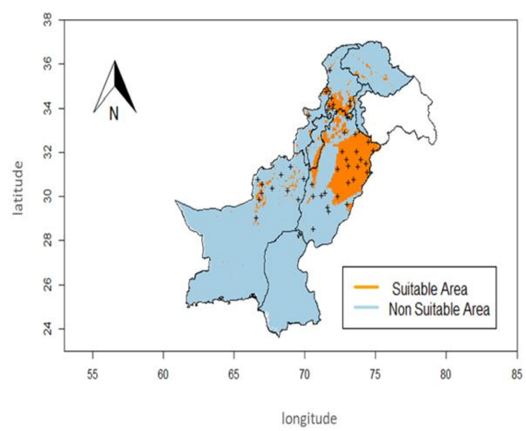


Figure 4.14: (a) Raw Map for Future Climatic Distribution (RCP 8.5) of Potato in Pakistan (b) Threshold map showing predicted (RCP 8.5) suitable climate area for potato cultivation

Threshold map was generated (figure 4.14 b) to simplify the output and show the suitable climate for potato in more simpler and binary form. The suitable area calculated was about 117752 Km^2 which mostly comprises Punjab. This area is less than the area calculated for current distribution and for RCP 8.5.

4.3.2 Provincial Breakdown of Suitable Climate for Potato under RCP 8.5

Area suitable in Punjab (figure 4.15 a) is found to be 86790 Km^2 followed by KPK (figure 4.15 b)which is 18515 Km^2. Suitable climate in Balochistan figure 4.15 c) is predicted to be 9123 Km^2 and Area under Sindh (figure 4.15 d) is going to be only 21 Km^2. The share of Gilgit Baltistan (figure 4.15 e) is about 2142 Km^2.

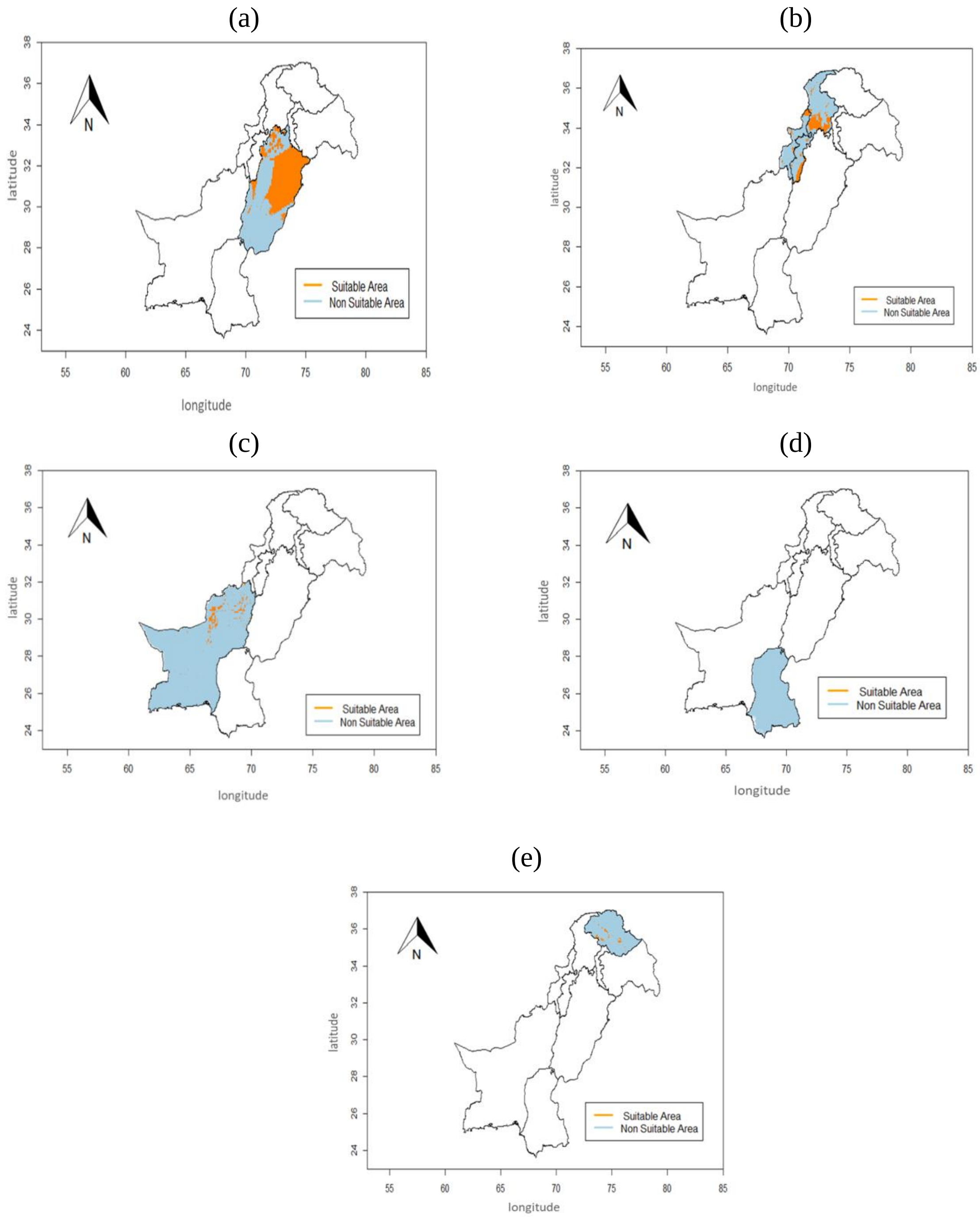


Figure 4.15: Provincial area breakdown for potato crop under RCP 8.5 (a) Suitable areas for potato cultivation in Punjab (b) Suitable areas for potato cultivation in FATA and KPK (c) Suitable areas for potato cultivation in Balochistan (d) Suitable areas for potato cultivation in Sindh and (e) Suitable areas for potato cultivation in Gilgit Baltistan

4.3.3 Maxent Model Evaluation

AUC found for RCP 8.5 showed good performance of our model with the value of 0.8 in fig 4.16.

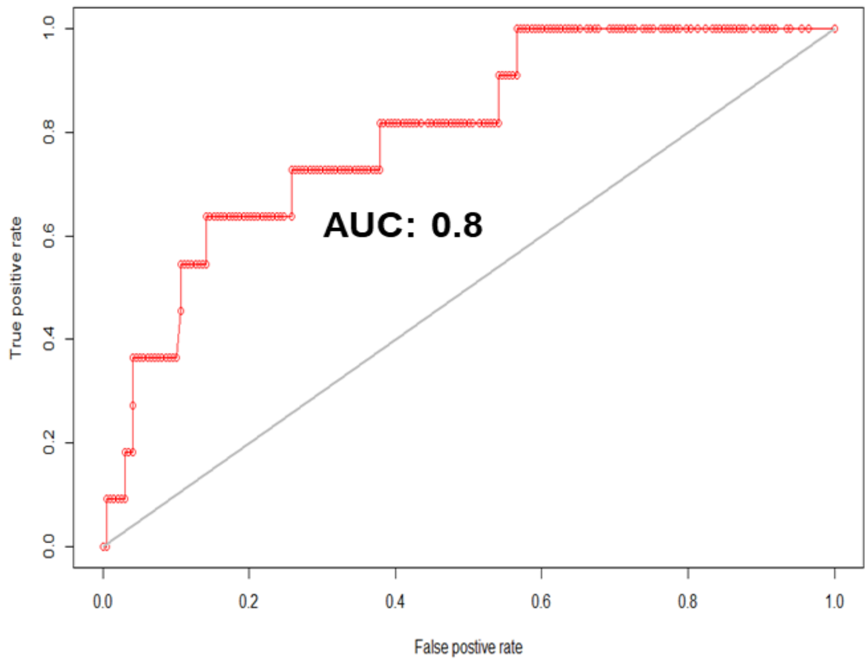


Figure 4.16: Area Under Curve for RCP 8.5

4.3.4 Variable Importance for RCP 8.5 Distribution

After running Jackknife for variable importance, the most important variables were found to be precipitation of wettest month and mean temperature of warmest quarter as shown in fig 4.17.

Figure 4.17: Jackknife test of variable importance under RCP 8.5

4.3.5 Pearson Correlation (RCP 8.5)

The multicollinearity test was conducted by using Pearson Correlation Coefficient (r) to examine the correlation of the most important variables with other variables shown in fig 4.18. According to the positive and negative values of correlation, Precipitation of wettest month is in positive correlation with precipitation of wettest quarter and in negative correlation with precipitation of driest month. Another important variable mean temperature of warmest quarter is found to positively correlate with maximum temperature of warmest month and negative correlation was seen with annual precipitation.

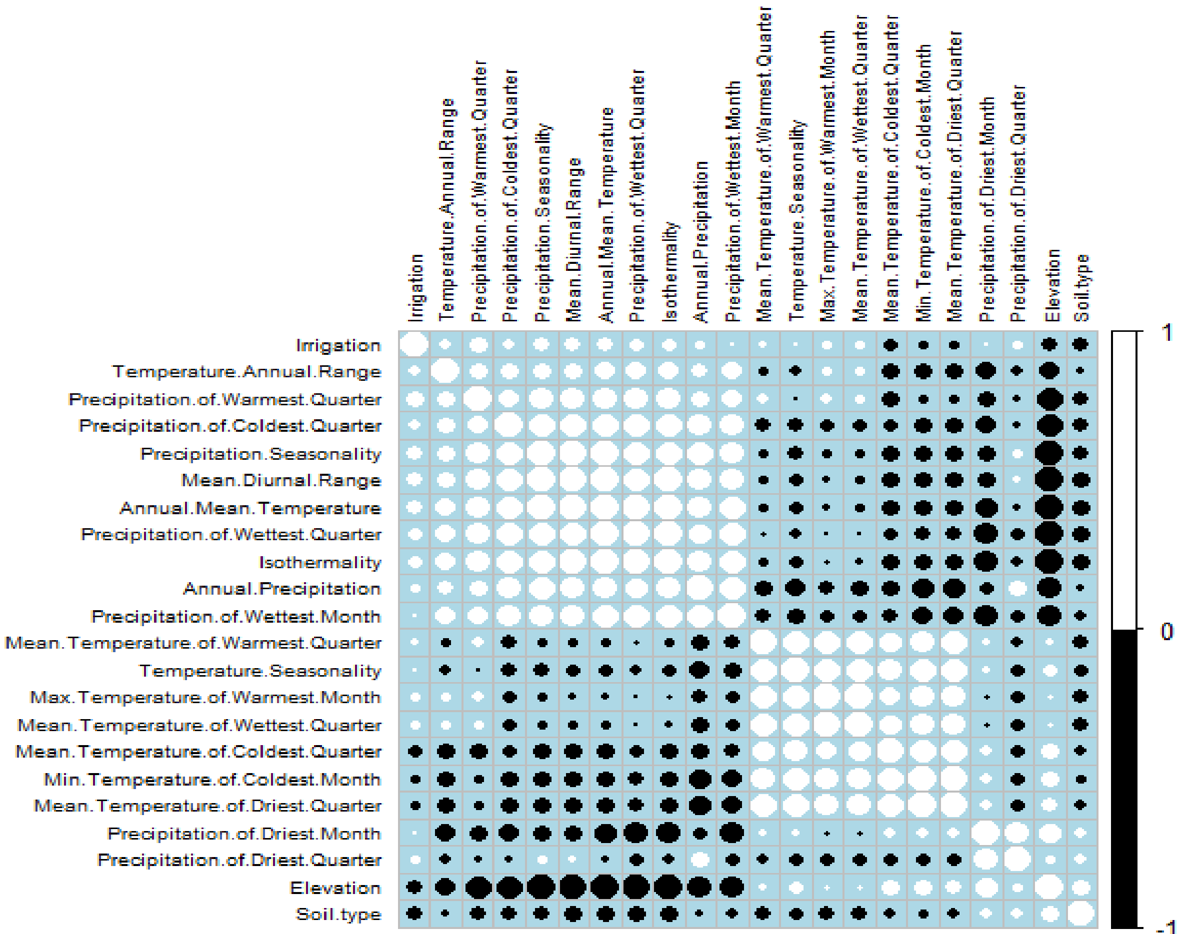


Figure 4.18: Pearson Correlation among Bioclim and other variables for RCP 8.5 distribution

4.4 Potato Predicted Distribution under RCP 4.5 (Model 2)

A CMIP5 model CCSM4 was used as Model 2 for this study and future distribution of potato was predicted for 2070 under RCP 4.5. RCP 4.5 is a moderate scenario where there is going to be a stabilization of temperature by 2100 which means the radiative forcing level stabilizes at 4.5 W/m^2. The results for RCP 4.5 show a slight decrease in the suitable area for potato in comparison with the area shown in current distribution.

4.4.1 Potato Predicted Distribution using Maxent (RCP 4.5)

In this scenario, reduction in predicted suitable area has been noted for potato cultivation. Map for the raw values of Maxent future distribution for potato under RCP 4.5 has been shown in the figure 4.19 (a

(a) (b)

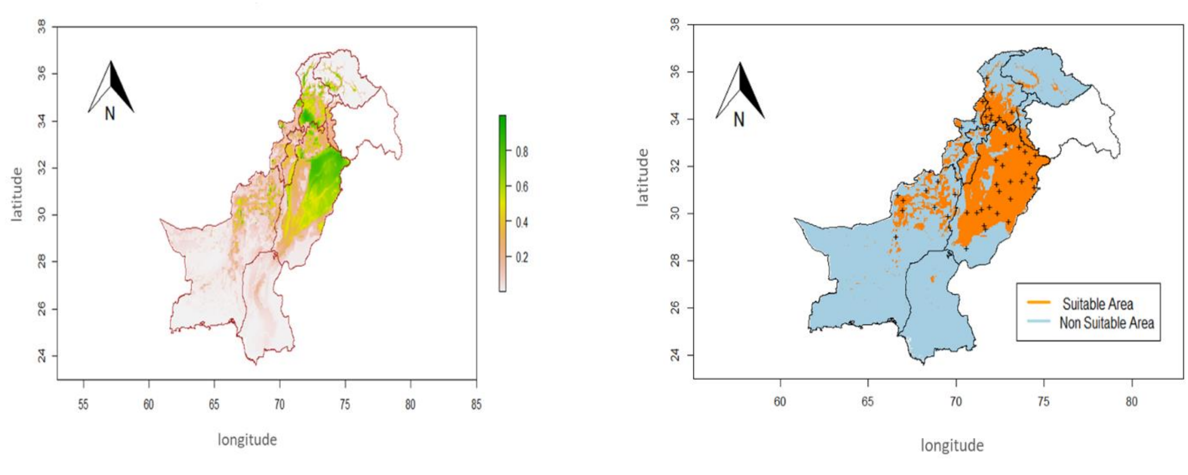


Figure 4.19: (a) Raw Map for Future Climatic Distribution (RCP 4.5) of Potato in Pakistan (b) Threshold map showing predicted (RCP 4.5) suitable climate area for potato cultivation

Threshold map was generated to simplify the output and show the suitable climate for potato in more simpler and binary form. The suitable area calculated was about 237132 Km^2 which mostly comprises Punjab and KPK. This area is slightly less than the area calculated for present distribution of potato crop as shown in figure 4.19 (b).

4.4.2 Provincial Breakdown of Suitable Climate for Potato under RCP 4.5

Area calculation was done based on the provinces of Pakistan and Punjab was found to be a province with maximum share of suitable area to grow potato in future, though the areas for provinces are predicted to be reduced compared to current suitability.

Area suitable in Punjab is found to be 149193 Km2 shown in figure 4.20 (a) followed by KPK in figure 4.20 (b) which is 43053 Km2. Suitable climate in Balochistan depicted in figure 4.20 (c) is predicted to be 35300 Km2 and Area under Sindh (figure 4.20 (d)) is going to be about 518 Km2. The share of Gilgit Baltistan figure is about 3556 Km2 shown in figure 4.20 (e).

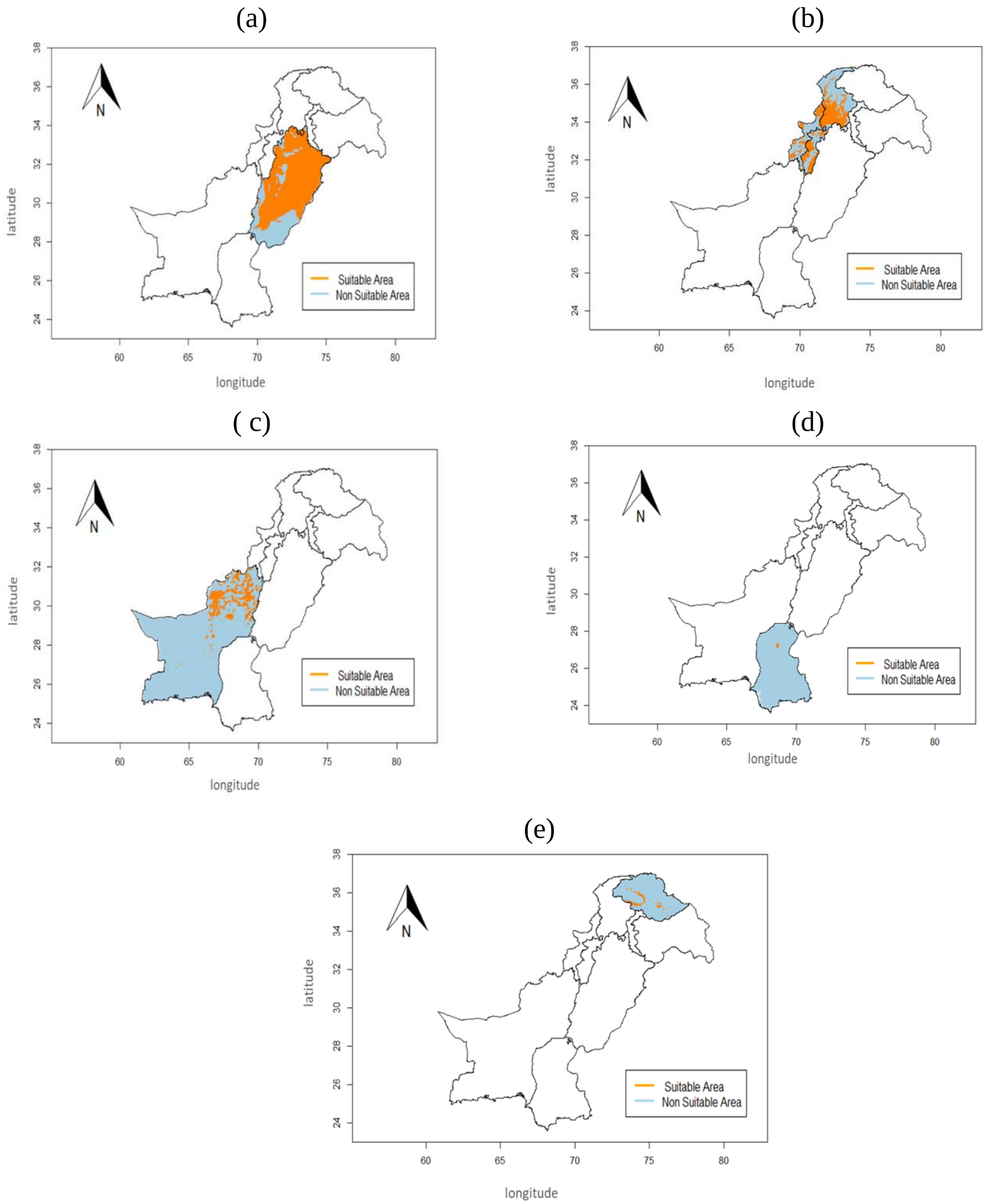


Figure 4.20: Provincial area breakdown for potato crop under RCP 4.5 (a) Suitable areas for potato cultivation in Punjab (b) Suitable areas for potato cultivation in FATA and KPK (c) Suitable areas for potato cultivation in Balochistan (d) Suitable areas for potato cultivation in Sindh and (e) Suitable areas for potato cultivation in Gilgit Baltistan.

4.4.3 Variable Importance for RCP 4.5 Distribution

After running Jackknife for variable importance, the most important variables were found to be precipitation of coldest quarter and annual mean temperature as shown in the figure below.

Figure 4.21: Jackknife test of variable importance under RCP 4.5 where important variables are precipitation of coldest quarter and annual mean temperature

4.4.4 Pearson Correlation (RCP 4.5)

The multicollinearity test was conducted by using Pearson Correlation Coefficient (r) to examine the correlation of the most important variables with other variables. According to the positive and negative values of correlation, Precipitation of coldest quarter is in positive correlation with isothermality and in negative correlation with elevation. Another important variable annual mean temperature is found to positively correlate with isothermality and negative correlation was seen with elevation.

4.5 Potato Predicted Distribution under RCP 8.5 (Model 2)

RCP 8.5 is an extreme scenario where temperature will keep on rising even after 2100 with emission level leading to a forcing level of around 6-7 W/m^2. The results for RCP 8.5 show a drastic decrease in the suitable area for potato in comparison with the area shown in current distribution and in RCP 4.5.

4.5.1 Potato Predicted Distribution using Maxent (RCP 8.5)

In this scenario, drastic reduction in predicted suitable area has been noted for potato cultivation. Map for the raw values of Maxent future distribution for potato under RCP 8.5 has been shown in the figure 4.22 (a).

(a) (b)

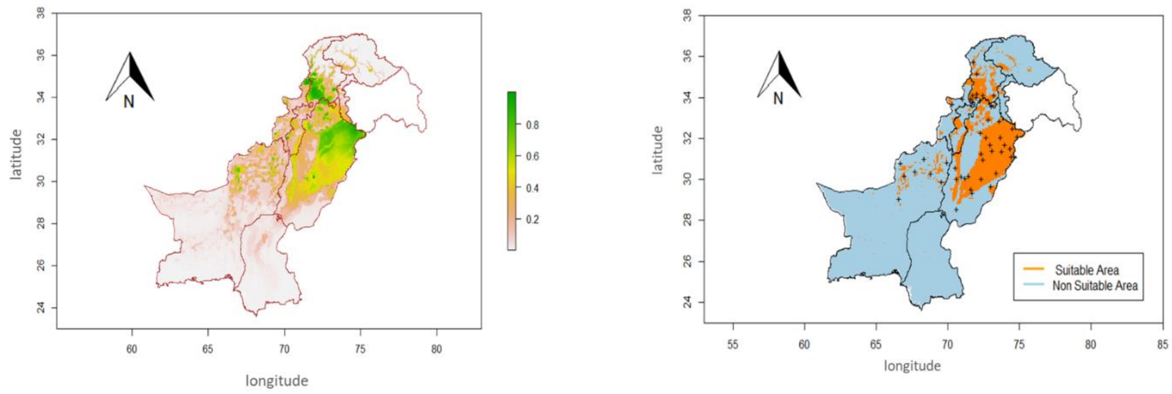


Figure 4.22: (a) Raw Map for Future Climatic Distribution (RCP 8.5) of Potato in Pakistan (b) Threshold map showing predicted (RCP 8.5) suitable climate area for potato cultivation

Threshold map was generated (figure 4.22 (b)) to simplify the output and show the suitable climate for potato in more simpler and binary form. The suitable area calculated was about 154598 Km^2 which mostly comprises Punjab. This area is less than the area calculated for current distribution and for RCP 8.5.

4.5.2 Provincial Breakdown of Suitable Climate for Potato under RCP 8.5

Area suitable in Punjab (figure 4.23a) is found to be 106086 Km^2 followed by KPK (figure 4.23b) which is 31440 Km^2. Suitable climate in Balochistan (figure 4.23c) is predicted to be 11183 Km^2 and Area under Sindh (figure 4.23d) is going to be only 57.7 Km^2. The share of Gilgit Baltistan (figure 4.23e) is about 3460 Km^2.

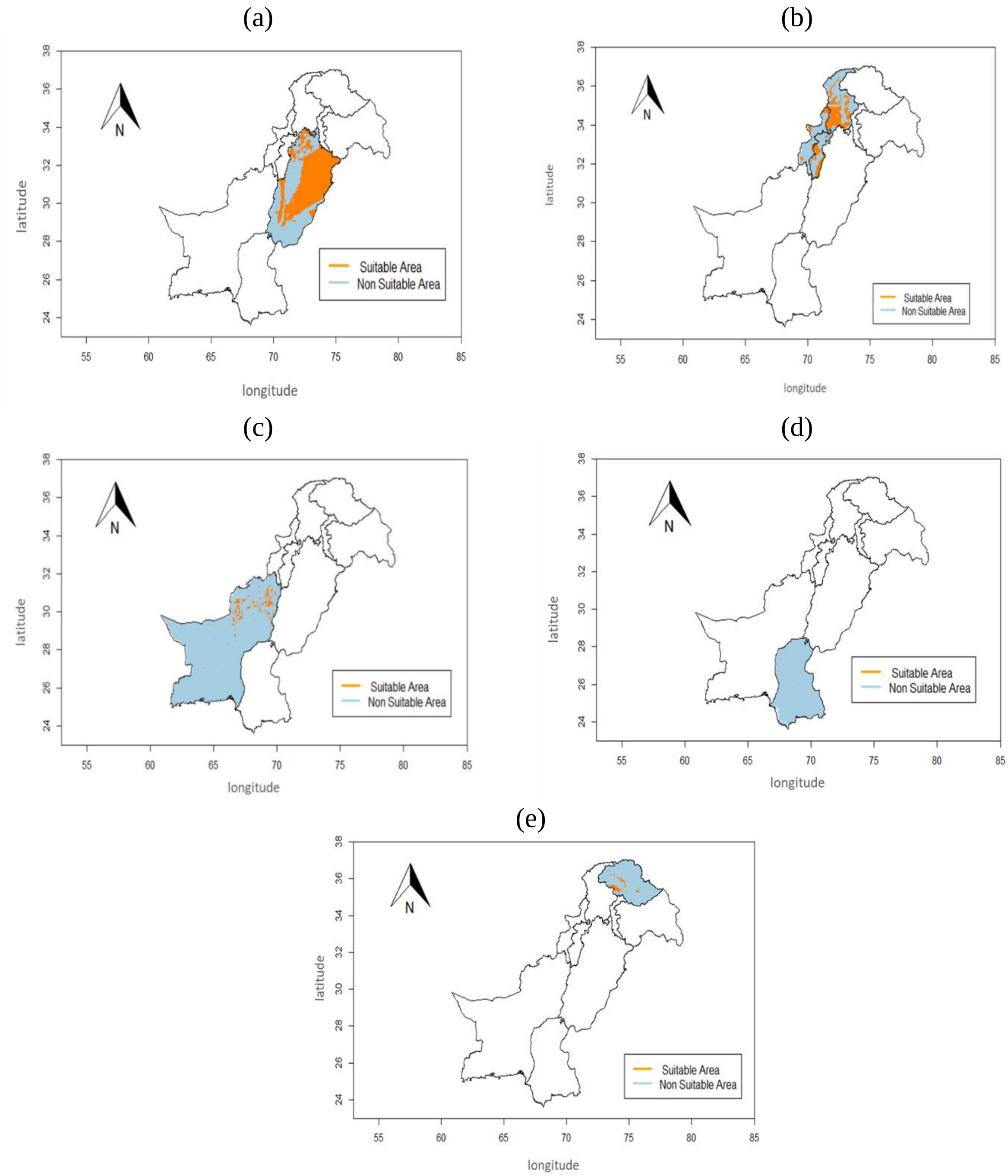


Figure 4.23: Provincial area breakdown for potato crop under RCP 8.5 (a) Suitable areas for potato cultivation in Punjab (b) Suitable areas for potato cultivation in FATA and KPK (c) Suitable areas for potato cultivation in Balochistan (d) Suitable areas for potato cultivation in Sindh and (e) Suitable areas for potato cultivation in Gilgit Baltistan.

4.5.3 Maxent Model Evaluation

AUC found for RCP 8.5 showed excellent performance of our model with the value of 0.9 as shown in figure 4.24.

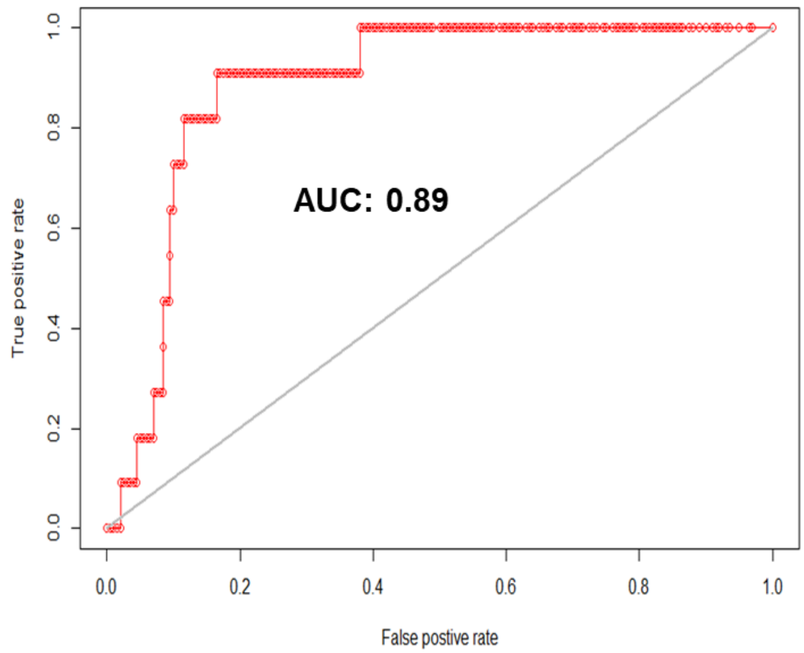


Figure 4.24: Area Under Curve for RCP 8.5

4.5.4 Variable Importance for RCP 8.5 Distribution

After running Jackknife for variable importance, the most important variables were found to be precipitation of coldest quarter and irrigation (figure 4.25)

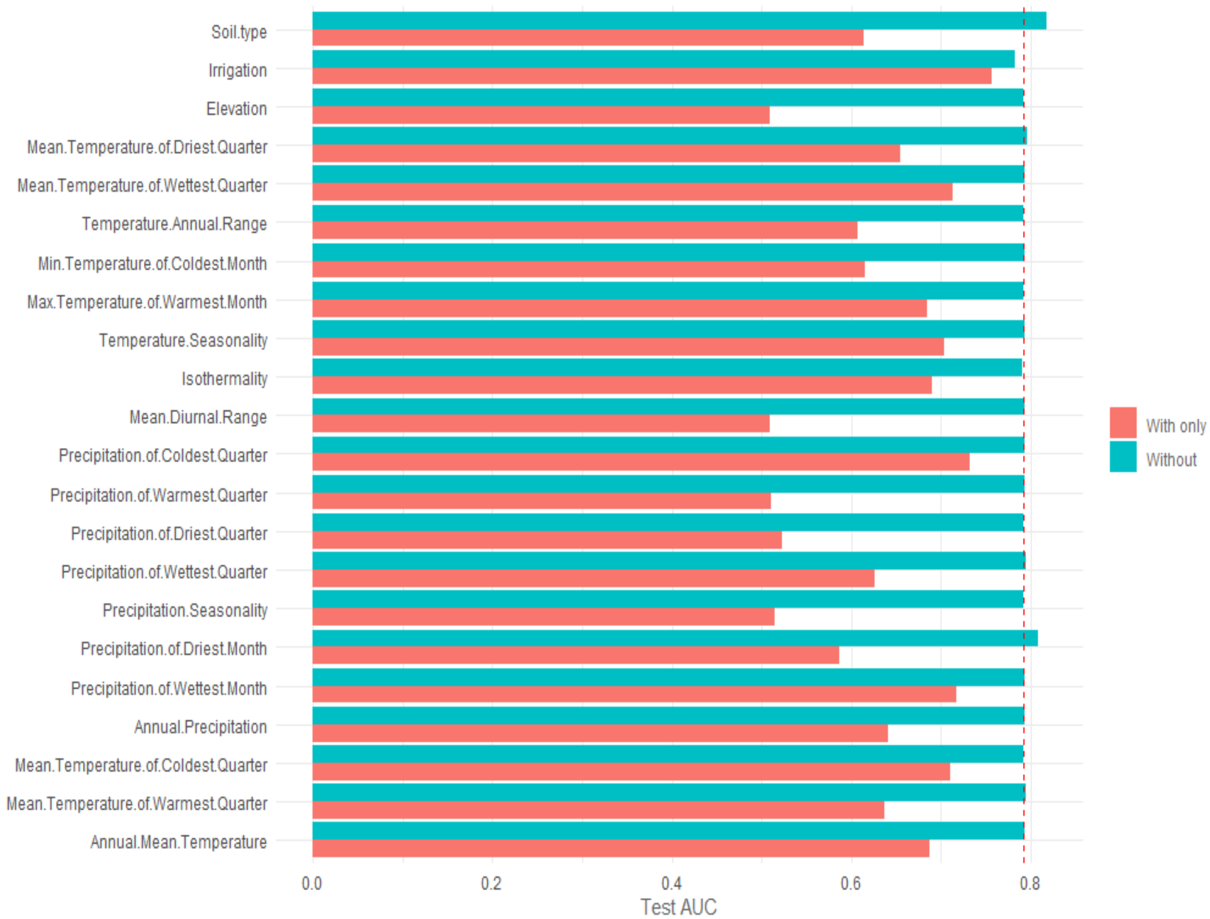


Figure 4.25: Jackknife test of variable importance under RCP 8.5 where important variables are precipitation of coldest quarter and irrigation

4.5.5 Pearson Correlation (RCP 8.5)

The multicollinearity test was conducted by using Pearson Correlation Coefficient (r) to examine the correlation of the most important variables with other variables. According to the positive and negative values of correlation, irrigation is in positive correlation with precipitation of warmest quarter and in negative correlation with elevation. Another important variable precipitation of coldest quarter is found to positively correlate with isothermality and negative correlation was seen with mean temperature of dries quarter as shown in figure 4.26.

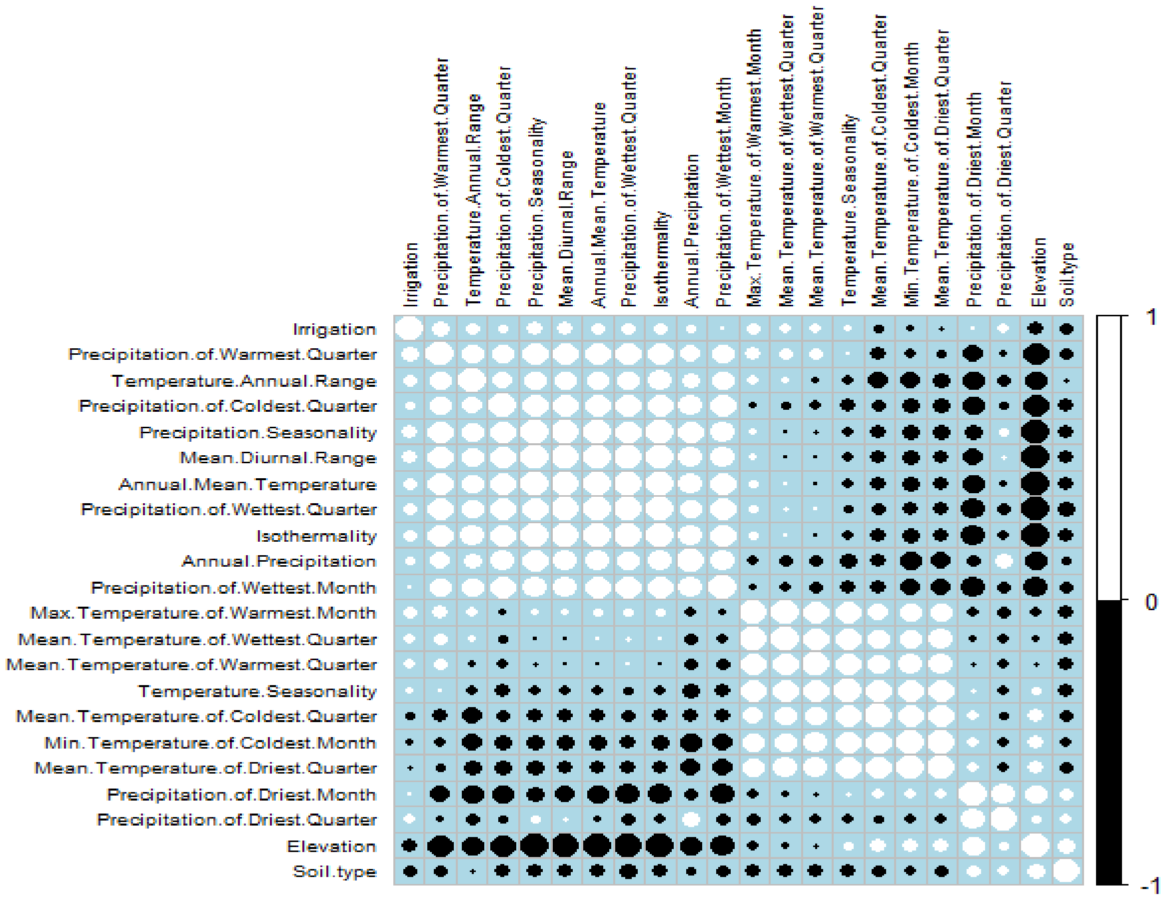


Figure 4.26: Pearson Correlation among Bioclim and other variables for RCP 8.5 distribution

4.6 Reduction in Areas (Model 1 and Model 2)

Area reduction for suitable crop area is shown and both models are show drastic decrease in area under RCP 8.5 for maxent results as shown in figure 4.27

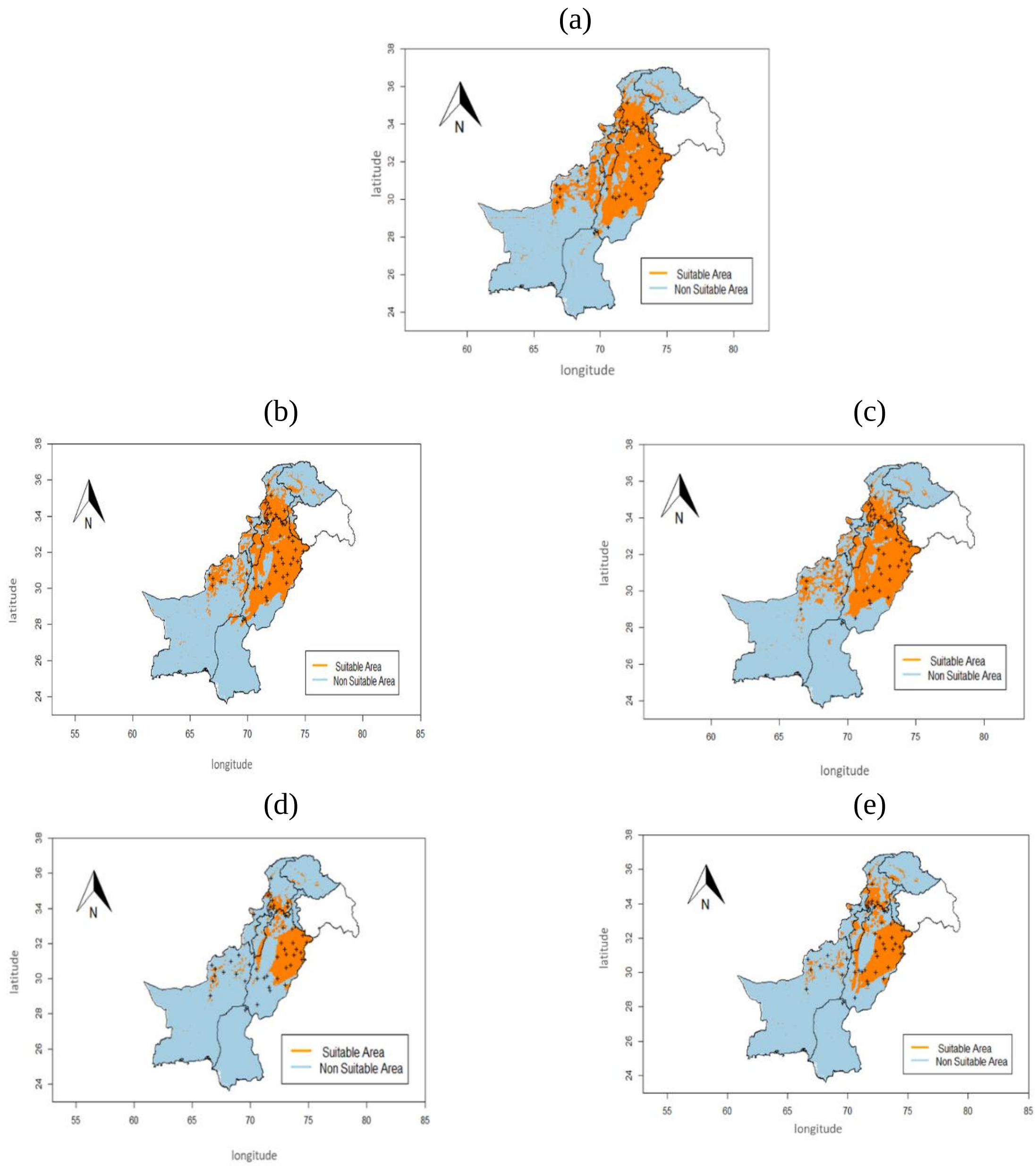


Figure 4.27: Reduction in Areas for RCP 4.5 and RCP 8.5 for model 1 and model 2. (a) suitable current climate area for potato cultivation (b) suitable climate area for potato cultivation for RCP 4.5 (Model 1) (c) suitable climate area for potato cultivation for RCP 4.5 (Model 2) (d) suitable climate area for potato cultivation for RCP 8.5 (Model 1) and (e) suitable climate area for potato cultivation for RCP 8.5 (Model 2)

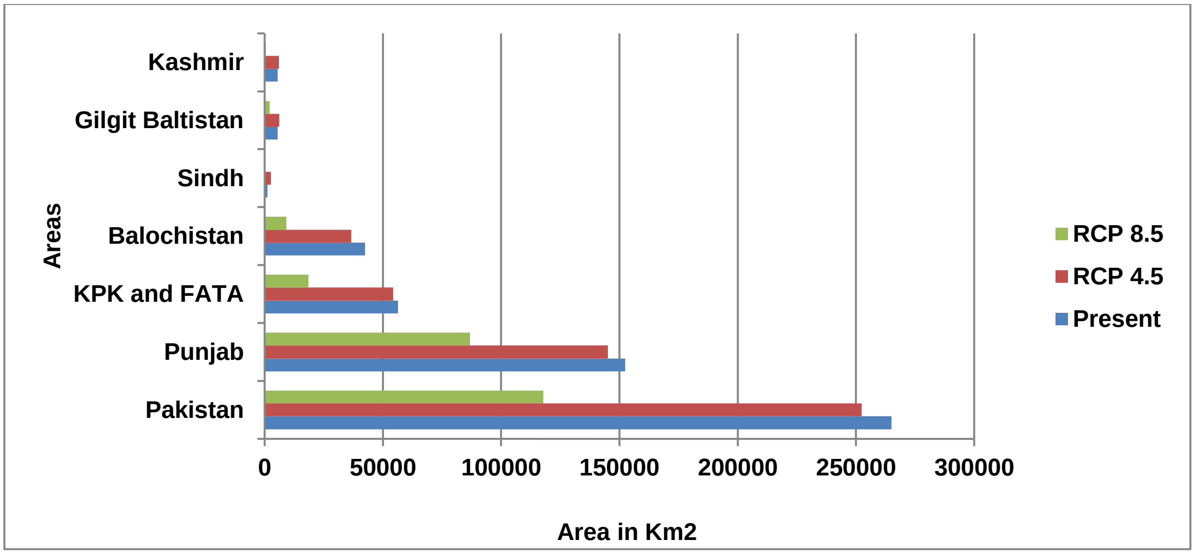


Figure 4.28: Area Reduction in Potato Production (Model 1)

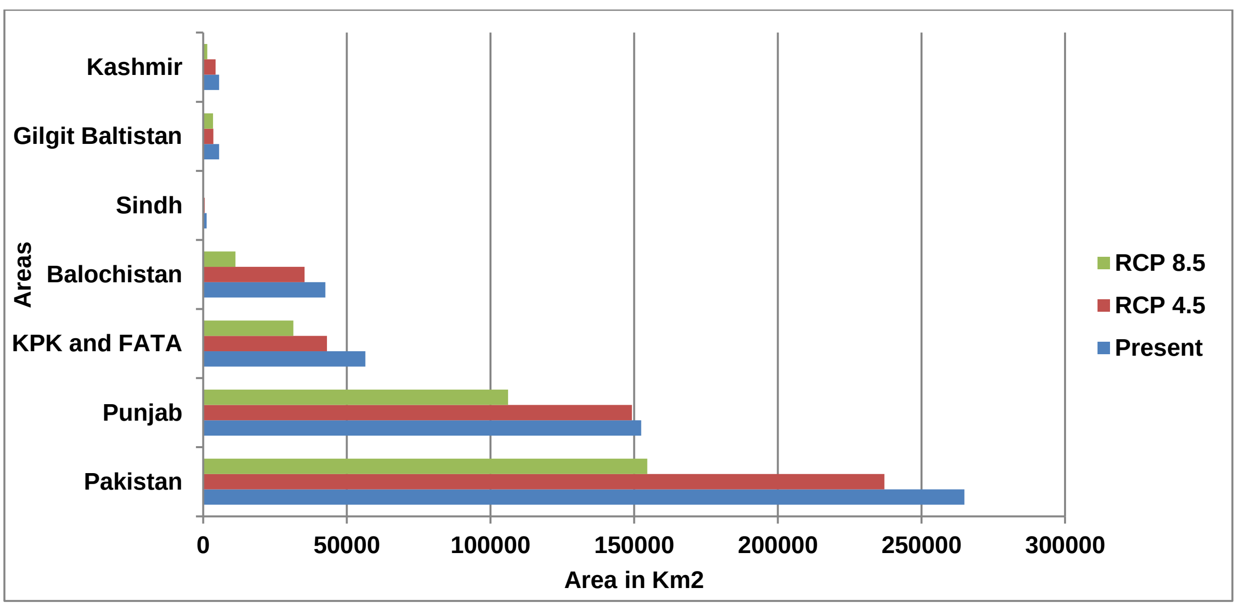


Figure 4.29: Area Reduction in Potato Production (Model 2)

In this research, modeling of the current and future distribution of Potato in Pakistan was performed and all the results of the model had Area Under the Curve value of more than 0.8 indicative of the robustness of the model (Hoveka et al., 2016).

Under the current and future scenarios, classification for the highly suitable, moderately suitable, less suitable and non-suitable area was done and the findings for the current scenario are supported by the other studies (Reddy et al., 2015) for the cultivation of Ceylon Spinach and (Reddy et al., 2015) for the cultivation of Vegetable Roselle in which the results show that Maxent Modeling Approach can be used as a tool to study the climate suitability of different vegetable crop species.

The results for the future distribution of *Solanum Tuberosum* coincide with the findings in the literature (Kumar, 2012; Guo at al., 2017; Zhang et al., 2018; Abdelaal et al., 2019) in which a temperature sensitive species underwent drastic decrease in the suitable habitat under the RCP 8.5 for year 2070.

The key factors shaping the distribution of the species is Isothermality which is the measure of the oscillations of temperature throughout the year. For the growth of the potato optimum temperature is required and if the temperature is more than 25°C or less than 5°C then the crop production might suffer (Majeed et al., 2018). Other variables include annual precipitation (Abdelaal et al., 2019), without the proper means of irrigation including the natural rainfall, the production of potato might suffer in the future. Another factor that is very important in the growth of potato is soil type, without suitable soil type and pH range, the growth of the species is not possible in any scenario just like another species *T. matsutake* highly dependent on the suitable soil type and there would be no yields without favorable soil conditions even if the climate and topography are in the favor of it’s production (Guo at al., 2017).

CHAPTER 5: CONCLUSION AND RECOMMENDATIONS

5.1 Conclusion

Potato is one of the basic food crops in Pakistan. It is very important to see its future distribution in order to cope with food insecurity. Maxent modeling (RCP 4.5 and RCP 8.5) was used to analyze its future distribution with two future models and both models predicted decrease in the suitable climate for Potato cultivation in 2070. The predicted highly suitable cultivation regions are located mainly in Punjab and KPK. Both future models predicted decrease in the suitable climate for Potato cultivation in 2070.

- In model 1 the suitable area reduction for RCP 4.5 is about 10.5% and for RCP 8.5 the suitable area reduced is about 71% with respect to the current suitable area.
- In model 2 the suitable area reduction for RCP 4.5 is about 4.7% and for RCP 8.5 the suitable area reduced is about 55.5% with respect to the current suitable area.
- Important environmental factors shaping suitable habitat of Potato for current climate include isothermality, soil type, temperature of driest quarter, precipitation of wettest quarter and annual precipitation.

5.2 Recommendations

Following are the recommendations to increase potato cultivation regions in future and to cope with food security.

- New varieties of Potato should be introduced that are resilient to changing climatic conditions.

- This research can be used for the future assessment of other non-grain crops and thus the estimates can help to generate more yields by adopting the ways to enhance the growth and cultivation of these crops in the future when food insecurity will be trending. Without better projections of crops in the future, there would be more chances of food shortage and famine so maxent tool is one of the best tools to identify the hotspots that can be utilized for the maximum cultivation of the species.

REFERENCES

Ahmad, Z., Chaudhry, H., Ali, H., Ahmad, S. R., Ashraf, U., Nisar, N., & Tahir, A. (2017). Suitable Habitat Evaluation of Eurasian Otter (Lutra Lutra) In Khyber Pukhtunkhwa, Pakistan. *Transylvanian Review*, *1*(4).

Baig, N., He, C., Khan, S., & Shah, S. A. (2019). CPEC and Food Security: Empirical Evidence From Pakistan. *Journal of Public Administration and Governance*, *9*(1), 191-208.

Bosso, L., Russo, D., Di Febbraro, M., Cristinzio, G., & Zoina, A. (2016). Potential distribution of Xylella fastidiosa in Italy: a maximum entropy model. *Phytopathologia Mediterranea*, *55*(1), 62-72.

Bosso, L., Russo, D., Di Febbraro, M., Cristinzio, G., & Zoina, A. (2016). Potential distribution of Xylella fastidiosa in Italy: a maximum entropy model. *Phytopathologia Mediterranea*, *55*(1), 62-72.

Chahouki, M. A. Z., & Sahragard, H. P. (2016). Maxent modelling for distribution of plant species habitats of rangelands (Iran). *Polish Journal of Ecology*, *64*(4), 453-468.

Chaudhri, T. (2017). The China-Pakistan Economic Corridor (CPEC) (Doctoral dissertation, The Ohio State University).

Combal, B., & Caumont, H. (2016). Supporting Marine Sciences With Cloud Services: Technical Feasibility and Challenges. In *Cloud Computing in Ocean and Atmospheric Sciences* (pp. 59-88). Academic Press.

Elith*, J., H. Graham*, C., P. Anderson, R., Dudík, M., Ferrier, S., Guisan, A., ... & Li, J. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, *29*(2), 129-151.

Esfanjani, J., Ghorbani, A., & Chahouki, M. A. Z. (2018). MaxEnt Modeling for Predicting Impacts of Environmental Factors on the Potential Distribution of Artemisia aucheri and Bromus tomentellus-Festuca ovina in Iran. *Polish Journal of Environmental Studies, 27(3).*

Fatima, S. H., Atif, S., Rasheed, S. B., Zaidi, F., & Hussain, E. (2016). Species Distribution Modelling of Aedes aegypti in two dengue-endemic regions of Pakistan. *Tropical Medicine & International Health*, *21*(3), 427-436.

Field, C.B., Barros, V.R., Mach, K., Mastrandrea, M., 2014. Climate change 2014: impacts, adaptation, and vulnerability. *Working Group II Contribution to the IPCC 5th Assessment Report—Technical Summary, pp. 1–76*

Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation 24 (1),* 38–49.

Food and Agriculture Organization of the United Nations (2017). *The future of food and agriculture—Trends and challenges.* Rome: Food and Agriculture Organization of the United Nations.

Gao, Q., Guo, Y., Xu, H., Ganjurjav, H., Li, Y., Wan, Y., ... & Liu, S. (2016). Climate change and its impacts on vegetation distribution and net primary productivity of the alpine ecosystem in the Qinghai-Tibetan Plateau. *Science of the total Environment*, *554*, 34-41.

Garcia, K., Lasco, R., Ines, A., Lyon, B., & Pulhin, F. (2013). Predicting geographic distribution and habitat suitability due to climate change of selected threatened forest tree species in the Philippines. *Applied geography*, *44*, 12-22.

Ghanbarian, G., Raoufat, M. R., Pourghasemi, H. R., & Safaeian, R. (2019). Habitat Suitability Mapping of Artemisia aucheri Boiss Based on the GLM Model in R. In *Spatial modeling in GIS and R for Earth and Environmental Sciences* (pp. 213-227). Elsevier.

Gregory, P. J., Ingram, J. S., & Brklacich, M. (2005). Climate change and food security. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1463), 2139-2148.

Guo, Y., Li, X., Zhao, Z., Wei, H., Gao, B., & Gu, W. (2017). Prediction of the potential geographic distribution of the ectomycorrhizal mushroom Tricholoma matsutake under multiple climate change scenarios. *Scientific reports*, *7*, 46221.

He, Q., Zhou, G., Lü, X., & Zhou, M. (2019). Climatic suitability and spatial distribution for summer maize cultivation in China at 1.5 and 2.0° C global warming. *Science Bulletin*, *64*(10), 690-697.

Hertel T W, Burke M B and Lobell D B 2010 The poverty implications of climate-induced crop yield changes by 2030 *Glob. Environ. Change* 20 577–85

Intergovernmental Panel on Climate Change, 2014. Climate Change 2014–Impacts, Adaptation and Vulnerability: Regional Aspects. Cambridge University Press.

Jing-Song, S., Guang-Sheng, Z., & Xing-Hua, S. (2012). Climatic suitability of the distribution of the winter wheat cultivation zone in China. *European Journal of Agronomy*, *43*, 77-86.

Jubb, I., Canadell, P., & Dix, M. (2013). Representative concentration pathways (RCPs). *Australian Government, Department of the Environment*.

Kukal, M.S., Irmak, S. (2018). Climate-Driven Crop Yield and Yield Variability and Climate Change Impacts on the U.S. Great Plains Agricultural Production. Sci Rep 8, 3450. https://doi.org/10.1038/s41598-018-21848-2

Majeed, A., & Muhammad, Z. (2018). Potato production in Pakistan: challenges and prospective management strategies–a review. *Pak. J. Bot*, *50*(5), 2077-2084.

Ochieng, Justus., & Kirimi, Lilian & Mathenge, Mary. (2016). Effects of Climate Variability and Change on Agricultural Production: The Case of Small-Scale Farmers in Kenya, Working Papers 229711, Egerton University, Tegemeo Institute of Agricultural Policy and Development.

Peavey, L., 2010. Predicting pelagic habitat with presence-only data using maximum entropy for Olive Ridley Sea Turtles in the eastern tropical Pacific. Master Thesis. Nicholas School of the Environment, Duke University.

Phillips, S. J., & Dudík, M. (2008). Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, *31*(2), 161-175.

Qin, A., Liu, B., Guo, Q., Bussmann, R. W., Ma, F., Jian, Z., ... & Pei, S. (2017). Maxent modeling for predicting impacts of climate change on the potential distribution of Thuja sutchuenensis Franch., an extremely endangered conifer from southwestern China. *Global Ecology and Conservation*, *10*, 139-146.

Raymundo, R., Asseng, S., Cammarano, D., & Quiroz, R. (2014). Potato, sweet potato, and yam models for climate change: A review. Field Crops Research, 166, 173-185.

Reckling, M. O. R. I. T. Z., Döring, T. F., Bergkvist, G. Ö. R. A. N., Chmielewski, F., Stoddard, F., Watson, C., ... & Bachinger, J. (2018). Grain legume yield instability has increased over 60 years in long-term field experiments as measured by a scale-adjusted coefficient of variation. *Asp. Appl. Biol*, *138*, 15-20.

Reddy, M. T., Begum, H., Sunil, N., Pandravada, S. R., Sivaraj, N., & Kumar, S. (2015). Mapping the climate suitability using MaxEnt modeling approach for Ceylon spinach (Basella alba L.) cultivation in India. *Journal of Agricultural Sciences–Sri Lanka*, *10*(2).

Rosenzweig, C., & Parry, M. L. (1994). Potential impact of climate change on world food supply. *Nature*, *367*(6459), 133.

Sahragard, H. P., Ajorlo, M., & Karami, P. (2018). Modeling habitat suitability of range plant species using random forest method in arid mountainous rangelands. *Journal of Mountain Science*, *15*(10), 2159-2171.

Spooner, D. M., McLean, K., Ramsay, G., Waugh, R., & Bryan, G. J. (2005). A single domestication for potato based on multilocus amplified fragment length polymorphism genotyping. *proceedings of the national academy of sciences*, *102*(41), 14694-14699.

Streck, Nereu Augusto. (2005). Climate change and agroecosystems: the effect of elevated atmospheric CO2 and temperature on crop growth, development, and yield. Ciência Rural, 35(3), 730-740. https://doi.org/10.1590/S0103-84782005000300041

Von Grebmer, K., Bernstein, J., Hossain, N., Brown, T., Prasai, N., Yohannes, Y., & Foley, C. (2017). *2017 global hunger index: The inequalities of hunger*. Intl Food Policy Res Inst.

Warren, D. L., & Seifert, S. N. (2011). Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological applications*, *21*(2), 335-342.

Wei, B., Wang, R., Hou, K., Wang, X., & Wu, W. (2018). Predicting the current and future cultivation regions of Carthamus tinctorius L. using MaxEnt model under climate change in China. *Global ecology and conservation*, *16,* e00477.

Yang, X. Q., Kushwaha, S. P. S., Saran, S., Xu, J., & Roy, P. S. (2013). Maxent modeling for predicting the potential distribution of medicinal plant, Justicia adhatoda L. in Lesser Himalayan foothills. *Ecological Engineering*, *51*, 83-87.