UNIVERSITY FOR DEVELOPMENT STUDIES

MODELLING THE EFFECT OF BIOCLIMATIC FACTORS ON THE PRESENT AND FUTURE DISTRIBUTION OF AFRICA LOCUST BEAN (PARKIA BIGLOBOSA) IN AFRICA

ASARE SAMUEL ABABIO

2021

UNIVERSITY FOR DEVELOPMENT STUDIES

FACULTY OF AGRICULTURE, FOOD AND CONSUMER SCIENCES

DEPARTMENT OF HORTICULTURE

MODELLING THE EFFECT OF BIOCLIMATIC FACTORS ON THE PRESENT AND FUTURE DISTRIBUTION OF AFRICA LOCUST BEAN (*PARKIA BIGLOBOSA)* IN AFRICA

BY

ASARE SAMUEL ABABIO (B.Ed. AGRICULTURAL SCIENCE)

(UDS/MHT/0002/18)

THESIS SUBMITTED TO THE DEPARTMENT OF HORTICULTURE, FACULTY OF AGRICULTURE, FOOD AND CONSUMER SCIENCES, UNIVERSITY FOR DEVELOPMENT STUDIES IN PARTIAL FULFILMENT OF THE REQUIRMENT FOR THE AWARD OF MASTER OF PHILOSOPHY HONOURS DEGREE IN HORTICULTURE

NOVEMBER, 2021

DECLARATION

I, Asare Samuel Ababio, hereby declare that except for references to other people's work which have been duly acknowledged, this thesis submitted to the Department of Horticulture, Faculty of Agriculture, Food and Consumer Sciences, is the results of my own investigation and has not been presented for any degree elsewhere

…………………………………	…………………………………
ASARE SAMUEL ABABIO	DR HYPOLITE BAYOR
(STUDENT)	(SUPERVISOR)
DATE………………………	DATE………………………

…………………………………	…………………………………
PROF ABDULHALIM ABUBAKARI	DR. HYPOLITE BAYOR
(CO-SUPERVISOR)	HEAD OF DEPARTMENT
DATE………………………	DATE………………………

ABSTRACT

With the advent of climate change, it is important to predict the effect it has on the potential distribution of species in order to identify those that are vulnerable for planning conservation strategies. This study was carried out to predict the effect of 2080 climate change on the potential distribution of *Parkia biglobosa*. A factorial combination of three resolutions (2.5, 5 and 10 arc-minutes; retrieved from WorldClim database) and eight sample sizes (5, 10, 25, 50, 75, 150, 200 and 305) were sampled from *Parkia biglobosa* location data retrieved from Global Biodiversity Information Facility (GBIF) and modelled using MaxEnt (version 3.4.4). All 19 bioclimatic variables in addition to four soil layers (retrieved from Harmonized World Soil Databases (version 1.2)) were used for the modelling with 15 replications. A virtual species model was also ran based on the mean annual rainfall (mm) and mean annual temperature (°C) of the GBIF data on the software "Virtualspecies". Models were evaluated using Area Under Curve (AUC), True Skill Statistic (TSS) and Kappa Statistic. Jaccard Similarity Index was also calculated between the predicted ranges and the true range for the virtual species. Results indicate that *Parkia biglobosa* has the potential to expand its present range by about 110% in the future 2080. Small sample sizes (<50) predicted imprecise ranges and also were less accurate in terms of evaluation statistics than larger sample sizes (≥50). The effect of resolution largely depended on sample size. Comparison between evaluation statistics and Jaccard Similarity Index suggests that evaluation statistics may not reflect accuracy of range prediction. It is concluded that, human assisted dispersal may be necessary to aid *Parkia biglobosa* to achieve its future potential.

DEDICATION

This work is dedicated to my parents Mr. Martin Asare and Mrs. Victoria Kwashie, my siblings; Campbell and Innocent, and Rev. Fr. Julius H. Sunu. May God bless them.

ACKNOWLEDGEMENT

My appreciation and gratitude go to my supervisor and his family, Dr Hypolite Bayor and Prof Abdul-Halim Abubakari for their immense support, patience, guidance and love that has made this dissertation a success. May God restore you in time, grace and health. I am also grateful to Dr Shu-iab Jakpa, Dr M.M Dawuda and the other Lecturers of the Department of Horticulture for their encouragement and challenge.

Special thanks to the staff and students of UDS Basic School, the families of Mr Joshua Ziblim and Mr. Isaac Ansah especially Cassandra, Mary and Martha. Stephen Ser-yir, Hudu Abdul-Rashid, Yeeb Abraham, Adam Mohammed-Awal, Thomas and Elisha, and all my course mates especially Mohammed Karimu, may the blessings of the Lord be with you for the great support you gave me in diverse ways throughout my study period.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHPATER ONE

INTRODUCTION

1.1 Background

Climate change and land-use change have become very important issues for especially scientific bodies including the Intergovernmental Panel on Climate Change (IPCC) (Odoemene, 2017). This is not because only humans are vulnerable but also importantly, the dramatic impact it is having on all biological communities and ecosystems. The IPCC's special report 2019 (Odoemene, 2017) on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems indicates that observed warming from the preindustrial era has resulted in increased frequencies, intensity and duration of events surrounding heat weaves in many areas.

Climate is defined as the conditions of the atmosphere for a particular region or location over time. It is usually considered as the long-term summation of atmospheric variables which in a short-term period is regarded as the weather. These variables include temperature, wind, humidity, precipitation, atmospheric pressure and solar radiation (https://www.britannica.com/science/climate-meteorology) (Retrieved on the 8th April, 2020). Climate is also regarded as the overall effect of weather and atmospheric conditions over a number of years for a given area or location. Climate includes both the average values of the climatic elements that existed in the past, likewise their extreme ranges and variability and the number of times of various occurrences (https://www.britannica.com/science/climate-meteorology). (Retrieved on the 8th April, 2020)

Climate, can be referred to as a description which considers averages and variability, of relevant quantities of surface variables for example temperature, precipitation, and wind, over a duration starting from months to many (thousands or millions of) years (Browne, 2011). In practice, 30 years is an acceptable standard normal period defined by World Meteorological Organization to define and describe the climate (Browne, 2011). These 30 years was selected as a duration long enough to get rid of variations that have accumulated from year-to-year for description or application of the climate for specific regions or general (Browne, 2011).

Patterns in climate features have drifted over the past centuries and the effect of climate change is already happening. Africa is among the most vulnerable continents facing significant effect of climate change (Kurukulasuriya et al., 2006). Climate conditions over the years have had significant impacts on growth and distribution of plant species (Kotir, 2011). There is also much and continuous observable evidence showing the biosphere has and is reacting to current rapid warming with changing genetic population structure, species distributions, shifting phenology and vegetation dynamics (Franklin et al., 2016).

A UN report (May 6, 2019) states that more than a million species of plants and animals are in danger of extinction with severe implications on the human planet in the next future as a result of climate change, exacerbated by human activities. The Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services went on to say that, all humans will suffer the consequences (Fears, 2019). Climate change has already and immediately caused species distribution shifts (including plants) across the globe. The negative effect of climate change threatens food systems and security, reduction in agriculture productivity and yields, and extinction of some crops (plants) (Kurukulasuriya et al., 2006); it affects the crust of the world's economy, health and livelihood (Fears, 2019).

The impact of climate change is suggested to be spatially variable (Thornton et al., 2008). With continuous land use for human livelihood especially through agriculture together with growing population, large regional variations of changes have contributed to net increase in greenhouse gas emissions, decline in biodiversity and loss of natural ecosystem (Odoemene, 2017). It is expected that, the greenhouse gas emissions will bring a significant effect to increase in global average temperature, resulting in increases in rainfall in certain regions of the globe, possibly causing floods as reported by Klutse et al., (2016), which are likely to affect the growth, production, and distribution of indigenous fruit trees species in Africa. The variation in rainfall pattern therefore is projected to have significant impact on agriculture and the local economy of people in many regions (Klutse et al., 2016). Rainfall is a major determining factor that influences crop choice and yield, as well, contribute to success of livelihood and other socio-economic activities in West Africa. Kotir, (2011) identified that the variability in rainfall (a climatic factor) onset, the period it stops and duration affect dates of planting (sowing), crop growth, yield and food production.

Bioclimatic variables such as temperature and precipitation have varying effects on the tree crops. For example, Olajuyigbe et al., (2013) report that the growth rate of *Diospyros mespiliformis* decreased with decreasing watering frequency in a study, and concluded that since the plant was able to survive the stressed water regime, it is able to survive in the savannah, and evidence of the impact water (rainfall) has on the development of plants. This did not take into consideration the changes and effects of other bioclimatic factors such as the temperature, the unexpected cessation of rains and the cooling effects of the atmosphere.

Some authors lament the deficit of work done to enable scientists to accurately predict what types of species are most threatened with extinction resulting from climate change (Kelly et al., 2014). To make informed choices and provide more precise strategies towards mitigating the effect (negative) of climate change, many studies have tried to investigate the impact of climate change and suggested ways to enhance mitigation actions in various ways. For example, Klutse et al., (2016) indicated, there were changes in extreme events including heat waves, droughts, floods, and many more combined with existing happenings capable of potential displacement of sensitive populations, crop failure or yield reduction, food insecurity, and water scarcity. They mentioned that "much burden is therefore placed on the lives and livelihoods of a region (Africa) already plagued with latent adaptive potential.

Again, on the potential consequence of 1.5℃ and 2.0℃ global warming on continuous dry and wet days throughout the West African region, results indicated a consistent changed pattern in temperature and rainfall across many regions of West Africa. Diasso et al., (2018) reported that, it is significant not to underestimate an increase of 0.5 ℃ rise in the global average which consequently leads to enhanced warming of up to 1.0℃ in some parts of Africa. The projected increase in the continuous dry days may negatively influence the future yield of crops which may result in increasing the danger of crop production and food security in the region (West Africa). These and many other studies give relevance to conservation planning to mitigate climate change.

Predicting to what extent species' distributions respond to variations in climate often uses one of a suitable methodologies at different times called Species' Distribution Modelling (SDMs), Habitat Modelling, or Ecological Niche Modelling (ENM) (Miller et al., 2007; Shabani et al., 2018).

Species distribution modelling, (SDM) make use of digital maps of the environmental variables, and spatial information on the vegetation factors/characteristics of interest (Franklin et al., 2017; Miller et al., 2007), species, type, abundance, usually from a sample of locations. These distributive models are produced using conventional statistical methods which are based on assumptions that, distribution of the vegetation is random and, hence, every observation is independent (Miller et al., 2007).

The standard definition of climate helps in describing the climate and makes it easy to be used as a base for which present conditions can be compared.

Although a combination of different techniques are important for global forecast, region-specific investigations that consider different factors of global change drivers may bring more insight on various interactions among drivers Kotir, (2011); Sylla et al., (2018), example, land-use change and fire, that are not easily accessible and accounted for at the global scale. Regional synthesis across the globe can yield very relevant insights into the causes of vegetation changes.

Understanding important variables that affect the growth and distribution of *Parkia biglobosa* in Africa is relevant since it will inform conservationists on the appropriate strategies to employ in conserving the species where the need arises. The relevance for studying this species is also because of the economic benefit the tree species gives to the people living in the geographical area it is distributed. For example, the matured beans are consumed as a dessert by many in the north of Ghana. Though it is a non-timber tree crop, dry branches are harvested and used as fuel wood domestically. It is also of great importance in the traditional pharmacology of people in most communities where it is located, it provides protein in diets as well, serve as feed supplement for farm animals like pigs among others.

To effectively predict the species' distribution, representative sample sizes and spatial resolution must be factored into the modelling and prediction scheme in order to get a true representation in the range. An accurate prediction for a model of species distribution in Africa needs much work to be done since limited species location data is available or is believed to be available. Available data on the web, for example, the Global Biodiversity Information Facility have some level of inadequacies in uploading location data point in the range distributed.

Although species distribution modelling is important for many uses including conservation planning and mitigation for biodiversity loss in general, the accuracy of the model is critical for it to be useful (Manzoor et al., 2018). The accuracy of

SDMs are influenced by many factors including sample size and the resolution at which the modelling is done.

The effect of sample size on accuracy of models have been studied for various species with most concentration keen on investigating the effect of sample size on model performance and accuracies Hernandez et al., (2006);. Jiménez-Valverde et al., (2009); McPherson et al., (2004), in the temperate regions but little is known of tropical Africa, for example Segurado & Arau, (2004) investigated on sample size, outside tropical Africa. Importantly, most tree species in Africa which are not been cared, like some cash crops (cocoa and rubber) need particular attention to understand the extent to which climate change has or could negatively affect their distribution. Many species including rare species as reported by Hernandez et al, (2006), need thorough investigation taking into consideration sample size and its function on range size and model accuracy. This is because these orphan or rare species have controlled spatial distribution (Hernandez et al., 2006), that is, there are few numbers of known locations. Very few studies have investigated sample size in relation to predicted range sizes and model accuracy on species indigenous to Africa. Although Hernandez et al., (2006), have shown that increasing sample size appeared to increase model accuracy, they also suggested that sample size might be species-specific and dependent on the geographic range of the species. Since *Parkia biglobosa* distribution has not been modelled, it was necessary to investigate how sample size affected the accuracy of SDMs as well as geographic range predictions.

The resolution at which models are developed are important and should be considered in any study (Franklin, 2010). The interval at which species respond to their environmental conditions may vary, therefore different scales for predictive modelling of the distribution of species should be carefully selected in the analysis of models (McPherson & Jetz, 2007). Results of the combination of the species ecological characteristics and the spatial resolution on model accuracy might differ, given that the selected spatial resolution employed for analysis may range from a close environment to a larger range (McPherson & Jetz, 2007).

Studies on the impact of climate change among other things must be examined to predict species distribution of indigenous fruit species.

Climate change and its effect on biodiversity in the African continent have made it prudent for species distribution modelling to predict how the distribution of species will be affected in the future. Good model prediction requires the right sample size and resolution. This is necessary for conservation planning into the future 2080. This will ensure the survival of the species now for both biological conservation of the ecosystem and the benefit derived from the species into the future 2080. This study will provide clarity on the minimum required sample size and resolution that will produce accurate models in terms of range and accuracy of models. It will also inform policy to maintain or improve on current mitigation strategies to provide favourable climate and soil conditions that affect the distribution of the species. It is also expected that standard accuracy measures used in predicting accuracy of models could be improved to give good overlaps as may be compared with the

Jaccard Similarity Index. Overall, a model of *Parkia biglobosa* in Africa will be known with determining variables that influence its distribution.

1.2 Objectives of the Study

The main objective of the study is to model the effect of climate variation on the potential distribution of *Parkia biglobosa* in Africa.

1.2.1 Specific Objectives

The specific objectives for the study are as follows;

1. To predict the present (1994) and future (2080) range of *Parkia biglobosa*.
2. To determine the effect of sample size and resolution on the accuracy of modelling the distribution of *Parkia biglobosa* in Africa.
3. Investigate the effect of climate change on the potential distribution of *Parkia biglobosa.*
4. To determine the most important environmental variable that affect the distribution of *Parkia biglobosa* in Africa.
5. To compare standard accuracy measures (AUC, TSS and Kappa) with Jaccard Similarity Index between the true range and the predicted range.

1.3 Study Questions

1. What is the present and future ranges of *Parkia biglobosa* in Africa under climate change?
2. How does sample size and resolution affect range predictions and accuracy predictions?

3. To what extent has climate change affected the potential distribution of *Parkia biglobosa* in to the future?
4. Which environmental variables are the most important for the distribution of the species?
5. Does the traditional accuracy measures predict good accuracies when compared to Jaccard Similarity Index?

1.4 Organization of the Study

The thesis is organized into six chapters. Chapter one is the introduction of the study. It focuses on the background of the study objectives of the study and the organization of the study. Chapter two reviews and discusses literature relevant to the topic to establish a theoretical approach for the research. The areas of literature considered are relevant to the study and provides evidence for analytical discussion to support the study. Chapter three focuses on, methodology applied to obtain the needed information for this study. Again, it presents research design and analysis. Chapter four present results of findings of the research within the context of the study objectives. Chapter five, focused on discussion of the study while chapter six focused on the conclusion, and recommendations base on the findings of the research. Appendixes are provided after references used in the study to give a much practical picture of some issues.

CHAPTER TWO

LITERATURE REVIEW

2.1 African Locust Bean (*Parkia biglobosa*)

Parkia biglobosa (Jacq), also known as "African locust bean", is a tropical tree in Africa popular for its uses as a non-timber wood species which provide food and medicine of various kinds. It has its fruits, leaves, bark, seeds, roots and stems in different economical and medicinal uses. *Parkia biglobosa* is an evergreen, deciduous, branchy crowned tree that is commonly found in the savannah and partially more dry locations of the Sahelian zone and the wetter areas south to the Guinean ecological zone of Africa (Lompo et al., 2017). It serves critical importance in the rural economy of most communities in Africa especially in the West African region. Its products are fast becoming export commodities. It provides not only fruits which are eaten as desserts but also the seeds, which are mostly processed into condiments (known as "Dawadawa" in Ghana) which serves as a spice/condiment in most meals (Dawadawa is a strong flavour and tasty soup condiment which is rich in protein). It is a source of timber and fuel wood for domestic purposes but also relevant in the traditional pharmacology.

It is a dicotyledonous plant which is categorized under the family Fabaceae - Mimosoideae. Though it is deciduous. It is a perennial plant that grows up to about 1 metre in a year: young seedling and between 7 and 20 metres high, and in some cases up to 30 meters. It is reported to start flowering at 5 to 7 years (Houndonougbo et al., 2020). The species is able to withstand fire (Abdulhamid et al., 2017), and has a thick dark grey-brown bark. The pods of the tree, which are referred to as locust beans, are green in the beginning and become dark brown when they are fully mature. They are usually 30 to 40 cm long on average, with some extreme lengths of about 45 cm in length. A pod is capable of containing up to 30 seeds (Lompo et al., 2017; Oyerinde et al., 2018).

P. biglobosa have different uses, including fodder, food, medicine, green manure, fuel wood, timber and many other economic purposes.

Fruits of *Parkia biglobosa* are commonly gathered from the wild and locally used as food, different from one community to another in the species location area. Some parts especially the bark of the tree is used for medicines as well as provide a wide range of commodities including tannins significant in the leather industry and to the livelihood to the communities.

2.2 Natural Range of the Species

Geographically, the species is distributed in some parts of Africa with a description considered in a range map, compiled by Hall et al., (1998) to show climatic zones and geographic boundaries. Naturally, the range covers about 20 countries mainly in the African savannah, north of the equator especially in the Suddanian vegetation zone, parts of which are partially in the drier, north of the Sahelian and wetter regions, south of the Guinean vegetation zone (Lompo et al., 2017). The species' land area covers diverse habitats but mainly on deep loamy and sandy soils (Lompo et al., 2017), characterized with annual rainfall ranging between 700 mm to 2,600 mm in the North and South respectively, and in exceptional situations, around 4,500 mm in countries such as Sierra-Leone and Guinea (Hall et al., 1998). Almost all the West African countries are endowed with *Parkia biglobosa*. The population density of the tree differs from one geographic region to another with a reported 40 trees found in a hectare (Lompo et al., 2017). Land use activities is believed to have affected the species' distribution over the past two decades since the assemblage of its occurrence data by Hall et al., (1998).

2.3 Cultivation of *Parkia biglobosa*

Parkia biglobosa' seeds are considered as orthodox seeds (Hong et al., 1996) which can be stored for 10 to 20 years, after which they lose significantly in their germination rate; a decline of approximately 15% from its initial germination percentage (Millogo et al., 2019). Lompo et al (2017), reported that *Parkia biglobosa* seeds tolerate desiccation at a low moisture content while keeping high viability even in freezing conditions. Fresh seeds are reported to reach up to a 95% germination rate when treated with sulphuric acid, to break dormancy (Lompo et al., 2017). Propagation of *Parkia biglobosa* through mature vegetative tissues has been identified to be difficult and mostly unsuccessful (Ræbild et al., 2011). Most propagation practices are through direct seeding and appear to be the most practical means used for improving and rejuvenating existing ranges (Lompo et al., 2017). *Parkia biglobosa* is a fast-growing tree species than some commonly used in agroforestry such as *Faidherbia albida* and *Ziziphus mauritiana* (Lompo, 1999). This characteristic gives the opportunity to increase distribution while promoting conservation, agroforestry and sustainable use of the species (Lompo et al., 2017).

2.4 Medicinal and Nutritive Value of *Parkia biglobosa*

Indigenous healers in various parts of Africa use various parts of the locust bean tree to derive health benefits. It was reported that *P. biglobosa* was one of the leading cited plants species used for treating hypertension, through a survey conducted on healers in Togo, (Makanjuola et al., 2016). Similarly, in a survey conducted in Guinea regarding the use of *P. biglobosa* as anti-malarial plants, *P. biglobosa* was cited among those most often successfully used (Makanjuola et al., 2016). The yellow colour of the pulp is an indication of the presence of nutrient possible to be vitamin A while its sour taste is considered to be due to the presence of ascorbic acid (Oyerinde et al., 2018). It has been reported that, the condiment made from the fermented locust bean seeds have between 12 – 16% carbohydrate, 39 – 47% protein, and 31 – 40% fat of lipids (Ikpeme, et al., 2002). In some areas, extract from the bark of the tree (appendices 3) is applied for the treatment of malaria, pneumonia, wound, and bronchitis (Ojewumi et al., 2016). Studies on biological compounds present in *Parkia biglobosa* leaves has reported the leaves to have antioxidant properties and significantly contribute to the immunity of cells and tissues against the effects of reactive oxygen species as well as other free radicals (Makanjuola et al., 2016).

Some reports emphasise that chemical elements that serve as protective agents from plants known with anti-peroxidative and antioxidant properties play an important role in giving protection to the liver against toxins (Fifamè Grâce Nadège et al., 2016). Aqueous extract of *Parkia* leaves induces an increase in total lymphocytes and TCD4+ in blood and therefore, it may assist in strengthening the immune system (especially of persons with weaker immune systems). Leaf extract of *Parkia* revealed tannins, steroid, flavonoid, terpenoids, cardiac glycoside, and saponin as some bioactive compounds present in the extract for which some are antioxidant agents against factors causing inflammation, hypertension, diarrhoea, diabetes, cardiac failure, bacterial infection, cancer cells, scurvy and membrane lipid peroxidation (Makanjuola et al., 2016).

2.5 Climate Change

Climate ChangClimate change perhaps has become a household name for most people in Africa, mostly linked to increasing temperature and/or flood. Many growers attribute failure in crop yield to decrease or outright cessation of rainfall prematurely to climate change. Climate change according to Hulme (2005), is significantly the most severe problem the world is facing today; considerably more serious than any form of threat from terrorism. Again, Hulme (2005), mentioned that climate change is interchangeably used with global warming and the greenhouse effect. Two variables, rainfall and temperature are considered the most affected climatic variables that are driving the devastating wheels of climate change while wind speed plays rather an important role that many do not consider by increasing the rate of evapotranspiration. Human activities have been identified to contribute a good percentage to global warming with industrialization occupying an important position in the causes of global warming and climate change (IPCC, 2008). Fossil pollutants causing greenhouse effect (Mertz et al., 2009), are mostly attributed to having contributed immensely to increased temperature, causing warming since many irradiated energies remains on the earth (IPCC, 2008).

Major physical events of climate change which are concerns in current times are increasing flood along many coastal regions throughout the world which are attributed to very high intensities of rainfall in short durations, melting of the ice cap and rising sea levels (Mertz et al., 2009), due to higher temperature.

Low ambient temperature has been found to hinder phytoplankton distribution, tree growth and phenology, plant biomass among others by various authors including (Hunter & Lechowicz, 2008; Liang et al., 2002), while high ambient temperature is said to have the greatest relevance in the extinction of some species such as terrestrial mollusc (Baur, 1992). An important phenomenon brought about by climate change is drought on farmland which has affected the abundance of a number of soil animal species, which is attributed to drier climate leading to the extinction of these animals (Hulme, 2005). Studies have focused on the effect of higher temperature and rainfall on the socio-economic livelihood of people often considering temperature and rainfall collectively. Example; Klutse et al., (2016), studied the onset and cessation of rainfall in West Africa. The implication it has on the livelihood of people is huge, especially on rural farmers who are dependent on the weather to grow their crops and make a living. Many years before now, Hulme, (2005), noted for example; Picton 1984; Pollard 1988 to have suggested that temperature and rainfall should not be held in isolation since the combined effect of these variables could be severe. That is, findings from the study of only one of the variables may result in misinformation on policy in terms of conservation planning and sustenance of agroforestry systems since these variables complement each other.

2.6 Effect of Climate Change in Africa

Across the world, the climate has been changing for several thousand years. The Inter-Governmental Panel on Climate Change (IPCC), IPCC, (2008), says climate change is evident and presents environmental, economic and social threats. Again they stated that recent warming in the climate system is unequivocal, as observation indicates an increase in global average air and ocean temperatures, widespread melting of snow and ice, and rising global average sea level (IPCC, 2008). Rising average global temperature and variations in precipitation are evidently and non-debatably impacting ecosystems, biodiversity and human civilization all through the world. A negative consequence of climate change considered the most significant is the risk to agriculture. The implication is rather serious since majority of people especially those in the developing countries depend on agriculture for livelihood (Kurukulasuriya et al., 2006).

Agriculture is vulnerable and delicate to weather and climate factors such as light, temperature and precipitation as well as weather extremes, like floods, storms and droughts (Kotir, 2011; Polak et al., 2016). It is expected that the impact of climate change will not equally influence systems among the population of the world Kurukulasuriya et al., (2006), with strong affirmation that the distribution of impacts will vary in the ability to respond to impacts and resource with which to (overcome) mitigate across nations. Vulnerable among nations are the developing countries, bring about inequalities in income distribution between and within countries (van Vuuren et al., 2011). It is considered that sub-Sahara Africa is most vulnerable and may be the most hard hit from the effects of climate change since it heavily depends on agriculture (Sylla et al., 2018).

Kotir (2011), reported that the vulnerability is because the existing climate is already severe while current knowledge is poorest combined with technological change that has been slow. Climate change is predicted to affect rain, elevate the intensity of droughts and increase mean temperatures, and endangers the abundance of fresh water for agriculture production.

The climate of Africa is distinguished by several climate zones, tropical rain forest, tropical wet and dry, tropical dry, mountain, Mediterranean, middle latitude dry, and humid sub-tropical (Kotir, 2011). Hulme (2005), described the climate of Africa as variable - It is varied because they stretch from humid to equatorial regimes, through seasonal and tropical regimes to sub-tropical Mediterranean-type climates; varying because all these climates show different degrees of temporal variability, particularly with regards to rainfall.

Studies have shown various degrees of changes in mean temperature, rainfall and extreme weather event (Jensen et al., 2008; Mensah et al., 2016). They also show that a change in climatic factors like precipitation, temperature and weather extremes such as floods and drought have been dramatic and expected to remain in coming decades across Africa (Mensah et al., 2016).

Evidence of studied temperatures have shown an increased warming trend since the 1960s (IPCC, 2008). Globally, records of warmth throughout the twentieth century is estimated to be about 0.5℃ per century with relatively higher warming in the June- August ,and September-November seasons then in December-February and March – May (Hulme, 2005). Again, it was recorded, that the late 1980s and 1990s had the warmest years with 1987 and 1998 recording the highest warmth. Temperature changes in the continent are not often uniform and they vary considerably within, and between regions and countries. Hulme (2005), reported mean annual diurnal temperature range increased by 0.7 to 0.9 and increased between 0.5 and 1 since the 1950s in Sudan and Ethiopian regions, and similar rise in temperature also experienced in the Zimbabwe. In South Africa, dianural temperature range declined around the 1950s and 1960s but has been relatively stable since then.

The third annual report of IPCC reports according Bjurström & Polk (2011), shows the earths mean surface temperature increased to 0.2℃ in the twentieth century and it is expected to remain, with a rise of 1.4℃ to 5.8℃ by 2100 (Stocker et al., 2001). The whole of Africa is expected to warm across all seasons throughout this century (Kotir, 2011).

Recent models suggest that the West African region has experienced a 1.5ºC warming since the year 2004 and projections point to a continuum in an increase by 2049 (Mensah et al., 2016), while increasing to 2℃ warming from around 2012 to 2066 (Vizy & Cook, 2012). Temperature increase occurred up to 3℃ global climate levels in the northernmost part of the region while temperature increase in the south is below 1℃. However, the difference in the projected changes between the North's and the South's global warming levels is around 0.5ºC and up to 1℃ in the south and the north respectively (Mensah et al., 2016). Again, they reported an increase in average seasonal rainfall in most part of the region except the northwest. Suggestions claim that increase drying, may result in drought which will negatively affect agriculture. Some other conclusion suggested that the difference in the impacts between 1.5ºC and 2℃ warming is significant and implies that, meeting the threshold of 1.5℃ or 2℃ will result in similar seasonal impacts. That is, rise in extreme rainfall with increasing temperature resulting in flooding in some cities along the coast in West Africa (Klutse et al., 2018).

2.7 Adaptation to Climate Change

An important factor in mitigating climate change and the effect it has or had or would have been to find key adaptation measures, sustainable enough to effect positive lasting changes on the ecological, social and economic systems. Here, adaptation is adjustment in the community that brings changes in climatic stimuli to a greater extent removing the impact on the ecology and social background. These could be done through capacity building; which includes the broadening of the knowledge and ability of the individual, groups and organization to develop adaptation measures by making clearly informed decisions, adapting and implement these adaptation strategies (Adger et al., 2005). These capacity measures they also referred to as "transforming that capacity into action". They insist that adapting to climate change is first an individual task, civil society groups stretching through to governments at local, national and international levels. Communicating climate change information through different forums, giving the potential impacts, venturing into new fields for new opportunities, maintenance of the land and maintaining economic growth are some of the postulated remedies to adapting successfully to adapting to climate change.

For Hulme (2005), the focus should be on making improvements in the flexibility in management of vulnerable communities that form the ecosystem, improvement in the inherent adaptability of plant and animal species as well as develop processes in the system while reducing environmental challenges and social pressure that raises vulnerability to climate variability. Referring to assumptions that, measures for adapting to present climate risk correlates with adapting to future change events (Houghton et al., 2001).

In Africa, technology has moved much slowly behind the rest of the world in adapting to climate change impact especially in adopting higher yielding crop varieties, exploring the benefit of irrigation amidst reduced precipitation and increasing capital investments (Kurukulasuriya et al., 2006). Suggestions to producers of crops and animals (as the most affected though most may not be aware) in Africa welcome investments in infrastructure development throughout Africa at local, regional and international levels in the area of irrigation which will increase the productivity and improve the livelihood of many on the continent (Kurukulasuriya et al., 2006).

Suggestions by some studies advocate for incorporating climate change in long-term conservation and even water use and resources planning, management and governance which is fundamentally important for decision makers (Sylla et al., 2018). For future climate, Sylla et al (2018), quote for example Karambiri et al., (2011); Roudier et al., (2014); Yira et al., (2017) reporting, the direction and magnitude of the projected changes are highly unknown; a statement which point to the need for studies to be refocused especially on future changes resulting from anthropogenic events since some level of significant contribution to mitigating the effect of climate change can be realized when it is embraced with urgency and will.

The need for vigorous work to mitigate the effect of climate change and intense climate variability in many areas also steam from many disasters across the world with respect to weather inconsistencies such as floods and cyclones over the past few years to decades. (Appendix 4) shows some records throughout the world of which Africa is no exception with regards to some disasters that has hit countries in the year 2014, in relation to climate change and its associated phenomenon.

2.8 Species Distribution Modelling

Species distribution models (SDMs) are a combination of verifiable data of the species occurrences or information of a species with data on climatic factors or environmental data. These models are essential in predicting species distributions through landscapes and unravel new discoveries into ecological and evolutionary studies and development, often dependent on extrapolation in time and space. They are widely applied in terrestrial, freshwater and marine studies (Shabani et al., 2016).

Many SDM methods exist which share a common approach in building and projecting models. That is, models are built in relation with known occurrence species dataset and environmental variables. Depending on the objectives of the study, factors such as soil texture, bulk density and land cover among others, relevant for the existence of the species population are considered. Usually, digital maps are used to aid the projection of environmental domains within which the species lives for both current and future projections or estimation of a model, allowing modellers to assess suitable models for the study and provide information on variables that are most critical (Franklin , 2010). According to Yates et al., (2010), most models employ the inexhaustible or no dispersal for estimating percentile gain or loss of climate conditions suitable for the species.

Models can be categorized into classes as; *(a)* ***statistical:*** with some examples as Generalized Linear Models (GLM), Generalized Additive Models (GAM) and Multivariate Adaptive Regression Splines (MARS); *(b)* ***climatic envelopes:*** such as BIOCLIM; and *(c)* ***machine learning techniques*** with examples as: Maximum Entropy (MaxEnt), Classification Regression Trees (CART) and GARP (Elith et al., 2006; Yates et al., 2010). A major significant difference between the methods primarily, is the kind of the species occurrence data being used. While some of these models have been designed to build models with presence-only data, others are designed to use presence-absence data. It is important to note the distinction with the models in other to apply appropriately when developing bioclimatic models aimed for conservation planning and climate change (Yates et al., 2010).

Presence-only data used by Elith et al., (2006) to compare 16 modelling techniques for the range of 226 species from different continents was found to be effective and can be used for modelling species distribution even across regions though novel methods were robust than the more established/traditional methods.

SDM's are increasingly used in ecological and conservational discipline to predict the impact of climate change on biodiversity and species distributions, but model evaluation remains challenging, because reliable data against which simulations of future ranges can be validated are seldom available (Yates et al., 2010). Consequently, model validation is often limited to how well they predict present range or distributions. Preferably model evaluations use an independent dataset (for current distributions), but more practically, most modellers often use data splitting, such that a portion of the data are used to train the model and the other portion held to validate it. One measure of classification accuracy is commonly used, the area under the curve (AUC) of receiver operating characteristic (ROC) plot (while the Kappa Statistic, True Skill Statistic, and Boyce Index) are other commonly used measures of classification of model accuracy (Antoine Guisan et al., 2016; Antoine Guisan & Thuiller, 2005).

Key challenges that affect SDMs include bias associated to presence-only data modelling procedures, model selection and evaluation, manipulation of biotic interactions, and the assessment of model uncertainty (Phillips et al., 2009). For example, the method of modelling may reflect a situation which is capable to misinform and unable to relate sequences in the range (Miller et al., 2007). Sample sizes, sampling techniques methods could contribute significantly to model predictions. For example, data inadequacies or errors in the available data in the GBIF may result to giving inaccurate (under prediction or over prediction of ranges) models prediction (Anderson et al., 2016). These causes has been categorized into three stages in data processing; first is the data collectors, where there could be errors in the data uploads leading to false data presentation, secondly is at the data entry stage and lastly the data consumers, how well the modeller handles the data during the modelling process and the interpretation given during analysis (Anderson et al., 2016).

The accuracy and robustness of the models rely on the choice and selection, variables that form parameters, interaction of geographic and environmental factors, modelling algorithm used, the degree of model calibration, and stages of projection. Present relationship between modelling practices and ecological sciences are generally poor, which limits development in the field.

Threshold dependent accuracy such as the sensitivity (SE) and specificity (SP) are used in many fields including species distribution modelling. The sensitivity measures the probability of models such that, the model accurately predicts an observation at a location. And specificity measures the probability that a known observation absent from a location is correctly predicted (C. Liu et al., 2020) Various disciplines refer to the SE and SP differently; example, in the field of imaging, they are known as producer's accuracy X. Liu et al., (2007) in machine learning and information retrieval, they are referred to as precision (Fawcett, 2006). These indices are used complementarily (Adams et al., 2001), and are being used in species distribution modelling.

Another measure of accuracy which is widely used in species distribution modelling is the Cohen's Kappa. Its adoption is aimed at correcting overestimation of the overall accuracy and measures the degree to which an agreement between observed and predicted is higher compared to the expected by chance alone (Liu et al., 2009).

True skill statistic (TSS) is another measure index widely recognized which was used as a test for diagnosis in medicine. It is defined as the average of the net prediction success rate for present locations and that for absent locations and similar to the arithmetic mean of sensitivity and specificity (C. Liu et al., 2020). TSS values greater than 0.2 but less than 0.6 (>0.2, <0.6) are considered fairly accurate while values ≥ 0.6 are good/ accurate models (ALLOUCHE et al., 2006), Area under the curve (AUC) of the (ROC) receiver operating characteristic is a threshold- independent index which is often used to measure the accuracy of various models in so many disciplines (Canran Liu et al., 2011), mentioned that, it is also used in ecological studies despite some criticism it has received as made by (Lobo et al., 2008). It is said to give a misleading picture of mean model performance and over predict ranges that are not practically significant (example; It is the probability that a model would rank a ramdomly selected species prescence location higher than randomly selected abscence location.

A summary of Liu et al., (2009) on the use of these measures of accuracy pronounced key important issues. First, almost all SDMs studies focuses on the discrimination capacity and reliability is not often evaluated. Two, they mentioned the need for precision of accuracies being estimated as important information for model assessment. Sample size for the test data needed to produce accurate estimate of model performance must be considered. They explained that, it has a close relationship to the statistical features of accuracy measures. Small sample sized test dataset lead to unstable accuracy measurement and may mislead conclusions made on model accuracy.

The Jaccard similarity index is another measure of accuracy used in conservation due to the ability to be applied to determine the relationship between species areas to determine for the optimum size for natural protection. Jaccard do similarities between two taxa where operational taxa is not affected by another taxa during analysis like producing independent values different from those of the operating taxonomic unit. Jaccard similarity index considers attributes of co-rated observations as the main function that significantly increase the similarity values (Bag et al., 2019). It is simply the ratio between the size of intersection and size of the union of sample sets.

2.9 Global Biodiversity Information Facility Data

Access to biodiversity data and information is a critical concern at a time where there is habitat loss across the world. The Global Biodiversity Information Facility (GBIF) unarguably is the single largest biodiversity data in the world. GBIF is an Intergovernmental Organization, providing an internet accessible, interoperable network of biodiversity databases and information technology tools Yesson et al., (2007), with a target to make the world's biodiversity data freely and universally available via the internet and has been described as a cornerstone resource. Biodiversity information from museums, herbaria and other organizations around the world constitute the information presently provided on the GBIF portal (Anderson et al., 2016). However, though the database is bulky, it is characterized with some patchy areas, where some locations, taxa etc. are evenly covered while some areas do not or are absent.

2.10 Worldclim Data

Spatially incorporated gridded climate data which is referred to by Hijmans, Cameron, Parra, Jones, & Jarvis, (2005), as climate surfaces provided climate layers referred to as "Worldclim version 1 database" for global land areas excluding Antarctica. The database consisted of precipitation and long-term monthly mean temperature. Over the decade, the original dataset has been improved to include solar radiation, wind speed and vapour pressure at high spatial resolution (less than $1km^2$). Formerly, these climate variables were only available at lower resolution (10 arc-minutes) and for varying base periods (New et al., 2002).

Climate data at very high resolution available may be essential for studies in environments of high topography and other areas with varying climate conditions which (Hijmans et al., 2005), referred to as strong climate gradients. Hijmans et al (2005), noted that "the availability of climate data at very high resolution allow for the evaluation of their utility with respect to data at a lower spatial resolution. In addition, the implication of climate data resolution on modelling result is almost unconsidered but current, it can be investigated across larger ranges of spatial resolutions less or equal to one (≥1) km. Though high resolution climate data or layers have brought improvement in investigating modelling results in very small spaces, it cannot be generalized to mean that the quality of the climate data is necessarily high in all places (Hijmans et al., 2005). To these it must be explained that the quality of the variables is spatially varied and are dependent on the variability of the climate in a location or an area, coupled with the quality and density of the observations, and the degree to which a spline can be fitted through it (Hijmans et al., 2005). Challenges in producing climate layers resulting from the low density of climate stations as well as not including relevant drivers, for example; aspect, made the climate layers unable to account for all the variation that may arise at a resolution of 1 km, specifically records of precipitation in mountainous areas.

Worldclim version 2 has only current data and work in progress for the future data, which already have gridded time-series of meteorological variables like the land surface temperature and cloud cover now available from a number of satellite-borne instruments. These have the potential to inform estimates of the variables of interest.

2.11 Sample Size and Model Accuracy

Models for species that have large range and environmental tolerance performs less in accuracy than for the species that have smaller ranges and restricted environmental forbearance (Elith et al., 2006; Wisz et al., 2008; Zurell et al., 2016). Small sample sizes are challenging for statistical analysis and results decreases predictive potential if compared to model occurrences (van Proosdij et al., 2016).

Accuracy should increase with increasing sample size until the highest accuracy potential is achieved. Added to this is that, the highest accuracy potential and the number of records that constitute a sample size for which the true range is predicted is dependent on the scale of species, the extent and spatial resolution of the environmental and species occurrence data available for the modelling and modelling method used (Hernandez et al., 2006). Accuracy complication may arise if only presence only data are used or are available owing to the difficulty to examine false positive prediction errors. Due to this, some modellers prefer to totally ignore commission errors but give concentration to only omission errors. This definitely is not healthy for or otherwise, simply not sufficient for any model because models with no omission errors may possess high commission errors (Hernandez et al., 2006).

Studies have found MaxEnt to be the best predictor even when a small sample size is applied. In the literature, Hernandez et al., (2006), showed that, out of the four models built to manipulate different sample sizes (GARP, Domain, MaxEnt and Bioclim), MaxEnt outperformed the other models with the mean of the least of the sample size (5) having the highest score while Domain had the least mean with the same sample size.

CHAPTER THREE

MATERIALS AND METHODS

3.1 Experiment One

Experiment one focused on using data from the GBIF (Global Biodiversity Information Facility) to model *Parkia biglobosa* while varying the sample size and resolution. This section employs standard techniques so that the present (1994) and future (2080) distribution of *Parkia biglobosa* could be predicted. The effect of both sample size and resolution were assessed on the range prediction and model accuracy.

3.1.1 Occurrence Data

African locust bean was chosen due to its usefulness in the local economy of many communities of Africa (especially Northern Ghana) where it is located and because it is a wild/orphan species. Species data used in this study was retrieved from the Global Biodiversity Information Facility (GBIF) downloaded from https://www.gbif.org. DOI 10.15468/dl.4fxtom (Retrieved 8th January, 2019). GBIF has a pool of several species' occurrence data supplied by ecological, geographic and conservationist organizations, other institutions and private individuals, using remote sensing, satellites and or fossil evidence to record the location and presence of species. It is a facility which is evaluated and updated regularly and occurrence (presence) is presented on digital maps, showing clearly the distribution as well as enjoying reference and access.

To mitigate the tendencies of working with data set with location coordinates which may influence predictions and may produce inaccurate models, species data were verified by filtering to take out occurrences that did not match coordinates or locations using Arc Map in ArcGIS version 10.4.1. The ArcMap was supplied with an electronic map which allows data coordinates to be plotted against the country boundaries. Where data points were marked as errors, they were removed using Microsoft excel (Excel). Observations of plants under cultivation (those which are out of their natural location/cultivated by someone in a geographic location other than their natural range) and observations with apparent errors were also taken away from the data set for modelling. Species locations largely perceived to be misrepresented were removed. (For example; the original (GBIF) dataset when opened in the ArcMap may show species location in the sea. This clearly is a deviation and could be attributed to errors in data entry processes) in conformity with (Magarey et al., 2018).

3.1.2 Samples and Resolution

After thinning the occurrence data and the three resolutions, the number of records remaining were, 888, 644 and 434 for resolution 2.5, 5 and 20 arc minutes respectively. For each resolution, a sample of 107 records were reserved for model validation. Eight different sample sizes (5, 10, 25, 50, 75, 250, 305) were randomly selected from the remaining data without replacement. Each sample was taken independently of the other from the data set. This was done to ensure that sample sizes were maintained the same for all three resolutions. This mode of sampling resulted in 24 data sets (3 resolutions × 8 sample size = 24).

3.1.3 Environmental Variables

Climate data were acquire from the WorldClim website (https://www.worldclim.org) (Hijmans et al., 2005). Raster layers for nineteen (19) bioclimatic variables obtained from the globally interpolated datasets representing annual trends, seasonality and extreme environmental variables which are supposed to be maximum relevant to plant existence (Pearson & Dawson, 2003), were used as environmental variables or data for the modelling. Climate data were cropped to correspond with the map of Africa. All the 19 bioclimatic variables were prepared at resolutions 2.5, 5 and 10 arc-minutes. These set of climatic variables was supplemented with four soil variables retrieved from the Harmonized World Soil Database.

The environmental variables, and their abbreviations are;

BIO1	Annual Mean Temperature (°C× 10)
BIO2	Mean Diurnal Range [Mean of monthly (max temp – min temp)] (°C)
BIO3	Isothermality (BIO2/BIO7) (× 100)
BIO4	Temperature Seasonality (standard deviation × 100)
BIO5	Maximum Temperature of Warmest Month (°C)
BIO6	Minimum temperature of Coldest Month (°C)
BIO7	Temperature Annual Range (BIO5 – BIO6) (°C)
BIO8	Mean Temperature of Wettest Quarter (°C)
BIO9	Mean Temperature of Driest Quarter (°C)
BIO10	Mean Temperature of Warmest Quarter (°C)

BIO11	Mean Temperature of Coldest Quarter (°C)
BIO12	Annual Precipitation (mm)
BIO13	Precipitation of Wettest Month (mm)
BIO14	Precipitation of Driest Month (mm)
BIO15	Precipitation Seasonality (Coefficient of Variation)
BIO16	Precipitation of Wettest Quarter (mm)
BIO17	Precipitation of Driest Quarter (mm)
BIO18	Precipitation of Warmest Quarter (mm)
BIO19	Precipitation of coldest Quarter (mm)
Bulk_dn×10	Soil bulk density ×10
Soil_txture	Soil texture
t_oc×100	Top soil organic content ×100
t_pH×10	Top soil pH ×10

3.1.4 Modelling Software

The Maximum Entropy for species distribution modelling, MaxEnt (ver. 3.4.4) was thew software used for the modelling. Model training was set to auto features. Each model was validated with the 107 records set aside. A threshold of 10 percentile training Cloglog presence was selected as a threshold for converting continuous probability to presence- absence predictions and the software was allowed to write the background predictions. Background predictions, sample predictions and maps of the binary predictions were used for computing the predicted ranges and test statistics namely Kappa statistic, True Skills Statistic (TSS) and the Area Under the Curve (AUC) of the receiver operative characteristic curve.

Three hundred and sixty (360) models (8 sample sizes × 3 resolutions × 15 replicates) were built for the species dataset and for each, a forecast was made for the current (1994) and the future (2080) range using the Had85 representative concentration pathway. The pathway represents one of the severe situations that might arise (van Vuuren et al., 2011). Current predicted area and the future predicted area was calculated using ArcMap 10.4.1. Available at https://www.esri.com/arcgis/about- ArcGIS.

The MaxEnt algorithm used to juxtapose presence positions and variable interactions to close interactions of background positions, and established the maximum entropy probability distribution coming close to uniformity, subject to the limitations thrusted by observed spatial distributions and connected environmental factors. The decreasing of relative entropy between known positions and background location data in such a situation optimizes the maximum entropy probability distribution (Elith et al., 2006).

3.1.5 Model Evaluation

MaxEnt was used to project the species distribution in the study area. Models were evaluated using the True Skill Statistic (TSS), Area Under the Curve (AUC) and Kappa Statistic. Here, a confusion matrix was computed to first calculate the "Sensitivity" and "Specificity" of the species occurrence. Confusion matrix in this modelling studies sought to address the discrepancies in the species occurrence dataset such that, it was used to compute for the possible observed location points and predictions, either present or absent or both observed and predicted absent. In this way, the sensitivity and the specificity are easily derived for the calculation of the True Skill Statistics.

Table 1 below is an illustration of a confusion table, indicating the observed occurrence data and the highest possible predictions which could be made. True positives are locations in which the species were observed to occur and for which the model predict the species to be present. False positives are locations in which the species was observed to be absent but which the model predict the species to be present. False negatives are locations in which the species was observed to be present but which the model predicts to be absent. True negatives are locations in which the species was observes to be absent and for which the mode predict the species to be absent.

Table 1: A Confusion Table Used to Derive Parameters Used in Calculating TSS

		Occurrence data	
		Present	**Absent**
Prediction	**Present**	True positives (a)	False positives (b)
	Absent	False negatives (c)	True negatives (d)

$$\text{Sensitivity} = \frac{True\ positive}{Present\ data}$$

$$\text{Specificity} = \frac{True\ Negative}{Total\ Abscence}$$ (ALLOUCHE et al., 2006; Bean et al., 2012) .

The AUC is an indication of the area under the receiver operating characteristics (ROC) curve. An ideal model would have an AUC to be equal to 1 while models which have 0.5 are regarded to be no different from random (Bean et al., 2012).

Jaccard similarity index measures the extent to which the predicted ranges accurately predict the true range and the true accuracy of the models. Ranging from zero (0) to one (1), values predicted below 0.7 are considered poorly overlapped and those equal to and above 0.7 are good models, predicting the measure to which the models perform accurately.

Table 2: Parameters Underlying the Computation of the Accuracy Measures Used for the Modelling

Measure	**Formula**
Overall accuracy	$\mathrm{A} = \frac{a+d}{n}$ ………………………………………….. Equation 1
Sensitivity	$\mathrm{Se} = \frac{a}{a+c}$ ………………………………………….. Equation 2
Specificity	$\mathrm{S} = \frac{d}{b+d}$ ………………………………………….. Equation 3
True Skill Statistic	$\mathrm{T} = Sensitivity + Specificity - 1$ ……….. Equation 4
Kappa statistic	$\mathrm{K} = \frac{\left(\frac{a+d}{N}\right) - \frac{(a+b)\,(a+c)+(c+d)\,(d+b}{N^2}}{1 - \frac{(a+b)\,(a+c)+(c+d)(d+b)}{N^2}}$ …………………. Equation 5
N	= a + b + c + d ……………………………………. Equation 6

3.1.6 Most Important Variables

To measure the most important variables that affect the models on the distribution of the species, MaxEnt was set to create response curves and calculate the percentage contribution of the variables to the model to measure variable importance. Percentage contributions of each variable were compared and the variable with the highest contribution was taken to be the most important for each of the model. The response curves show the extent to which each of the environmental factors affects MaxEnt prediction. The curves show how the forecasted probability of presence changes as each environmental factor are varied, keeping all other environmental variables at their average sample value.

3.1.7 Transformation and Analysis of Data

The ranges (areas) of the species generated from the models and their accuracy measures (AUC, TSS and Kappa) were analysed using ANOVA in GenStat version 12 ed. Exploratory plots (Box- plots) suggested the data were suffering from heteroscedasticity and the main cause was the sample sizes. Smaller sample sizes had higher variability than larger sample sizes. In order to use ANOVA, log to base 10 and the Box-Cox group of transformation were explored. In most cases, the best transformation from the Box-Cox transformations improved the data significantly and were used before ANOVA was performed.

3.2 Experiment Two

3.2.1 Virtual Species

This experiment concentrates on the development of a virtual species using the Virtualspecies Programme running on R (Leroy, 2014; Leroy, 2019). The virtual species was used to determine which sample size and resolution gave the best model predictions made in the first experiment in terms of the predicted range and model accuracy. This is possible because the virtual species afford the modeler the insight and foreknowledge of species distribution in the range.

3.2.2 Sample Size and Resolution

For consistency and validation purposes, all the eight (8) sample sizes used in experiment 1 were repeated in this experiment. That is; sample size 5, 10, 25, 50, 75, 150, 200 and 305. These were combined with the three resolutions; 2.5, 5, 10 arc minutes and replicated 100.

The virtual species was generated using the routine Virtual species Leroy et al., (2018) running in R (R core team, 2019). Two environmental variables were used, namely Annual mean temperature (BIO1) and Annual precipitation (BIO12). The specific values were determined using the location data obtained from GBIF site. The mean and standard deviation of the two environmental variables (BIO1 and BIO12) were determined (annual temperature: mean = 27℃; standard deviation = 11.1℃; annual precipitation: mean = 181mm; standard deviation = 261mm), and used for the virtual species. This ensured that the virtual species mimicked the real species in at least two environmental variables. By generating the virtual species, the exact range of virtual species was known and this was necessary to determine how close model predicts were to the real range. It also enabled us to put model accuracy validation (AUC, TSS and Kappa) in their right context. Validation statistics that were too high or two low could be detected because by comparing the predictions of each model to the known range of the virtual species, the true accuracy of the model could be inferred.

3.2.3 Modelling Software

This experiment like the first one used Maximum Entropy for species distribution modelling, MaxEnt (ver. 3.4.4) to build the models. For consistency, all settings in experiment one was maintained.

3.2.4 Model Evaluation

MaxEnt models were evaluated using the same accuracy measures used in experiment one. That is, the True Skill Statistic (TSS), Area Under the Curve (AUC) and Cohen's Kappa.

Since the range of the virtual species was known, it was possible to compare the predicted range from each model to the actual range. This enable the modeler to determine precisely which of the model was the most accurate without necessarily using model evaluation statistics.

CHPATER FOUR

RESULTS

4.1 Experiment One: Species Distribution Models of *Parkia biglobosa* Using Global Biodiversity Information Facility (GBIF) Dataset

4.1.1 Present Range Prediction

Analysis of variance (ANOVA) was conducted on the predicted range produced by the models. The interaction of resolution and sample size was significant ($p < 0.001$). The interaction of the resolution and the sample size showed that with smaller sizes (≤ 50), range predictions were not significantly different among resolutions (Figure 1). With larger sample sizes however (75-305) lower resolutions predicted larger ranges than higher resolutions (Figure 1). At lower sample sizes (example sample size 5), there was very high variability in range predictions. For example, the maximum predicted range was 19.5 times larger than minimum. This variability decreases as sample sizes increases. At sample size 305, the maximum prediction was only 1.7 times larger than the minimum prediction. This appeared to make smaller sample sizes (<50) very imprecise and probably unreliable (Figure 1).

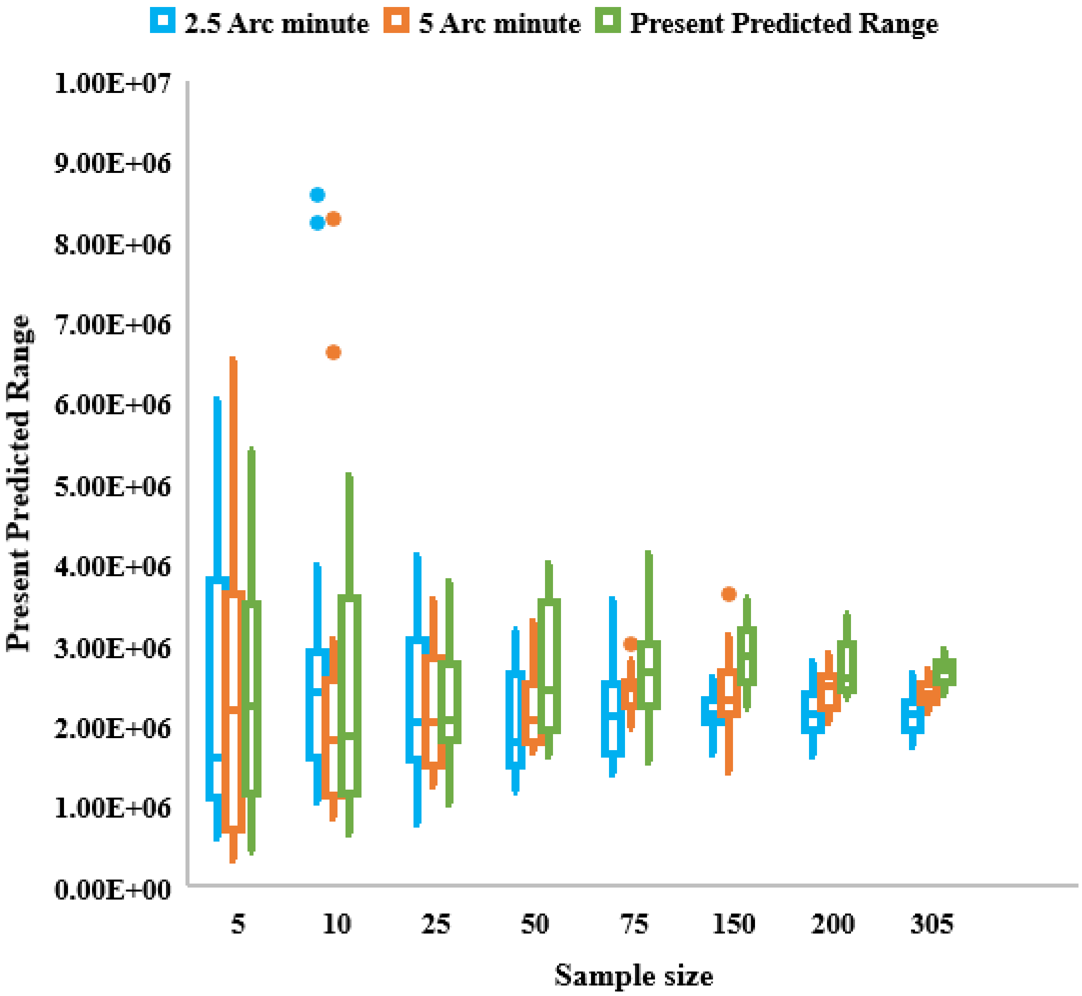


Figure 1: Effect of resolution and sample size on the predicted range

4.1.2 Future (2080) Predicted Range

ANOVA conducted on the future (2080) predicted range showed that the interaction between resolution and sample size was significant ($p<0.005$). At the lower sample sizes (5-50), resolutions were not significantly different within each sample size. However, at sample sizes greater than 50, the coarser resolutions tended to predict larger ranges than the lower sample sizes (Figure 2). The main effects, namely size and resolution were both significant ($p<0.005$). For sample size, predicted ranges decreased as sample size increased (Figure 2). For resolution, predicted range increased with coarseness of the resolution. However, resolution 5 arc minutes was not significantly different from resolution 10 arc minutes.

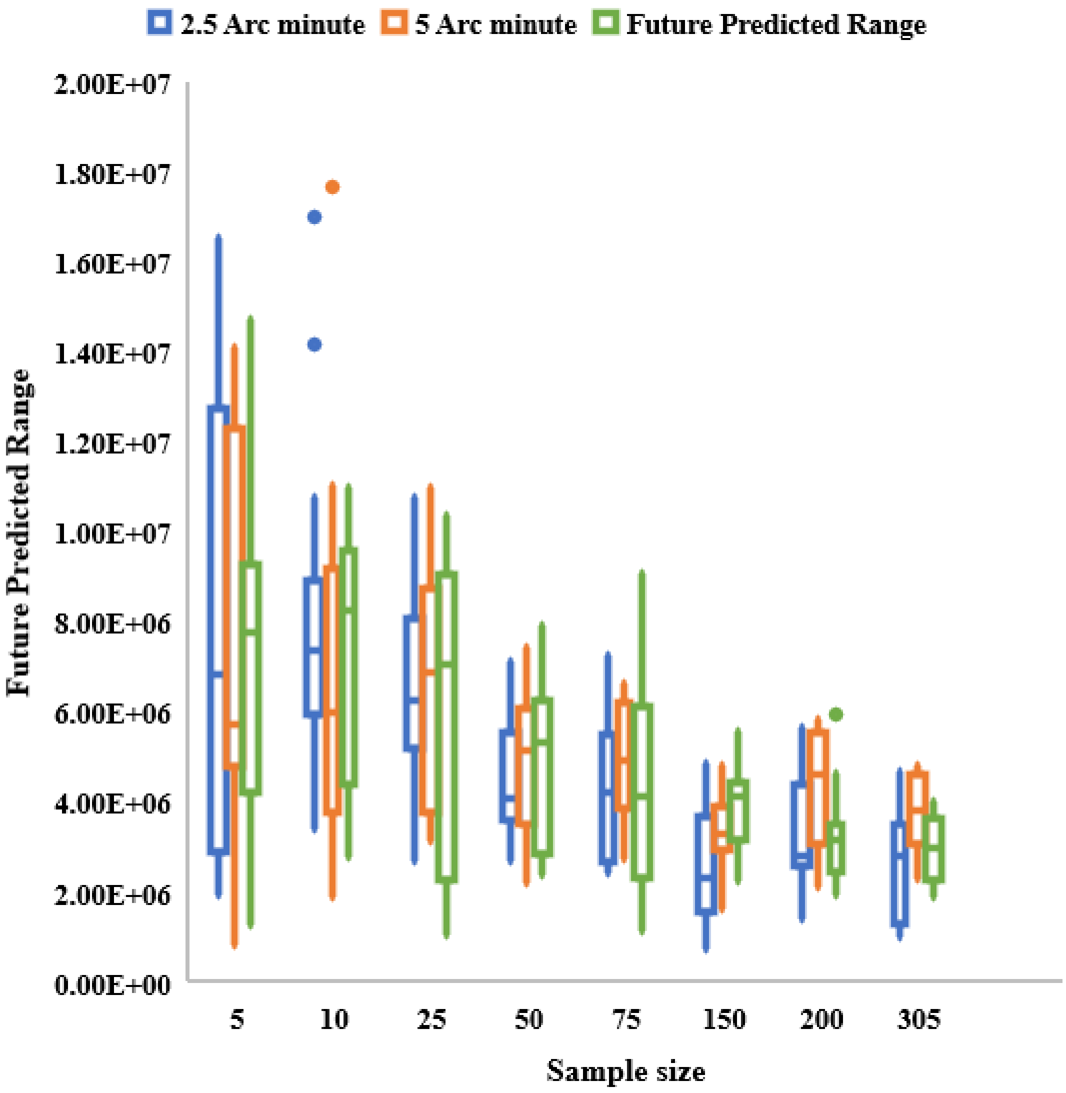


Figure 2: Interaction effect of resolution and the sample size on the future predicted range

4.2 Sample Size and Resolution on The Accuracy of The Modelling Distribution (Model Accuracy Assessment)

4.2.1 True Skill Statistic (TSS)

The ANOVA of TSS showed that interaction effect of the sample size and resolution was not significant ($p > 0.956$). Sample size (Figure 3) showed significant differences ($p < 0.001$). TSS values increased with increasing sample size and each higher value was significantly different from the one immediately below except sample size 200 and 305 which were not different (using lsd). The TSS values ranged from 0.4 for sample size 5 to 0.8 for samples size 305. This shows that all sample sizes gave fairly accurate models for TSS since a TSS value of ≥ 0.2 are often regarded as fairly accurate.

The resolution was also significant ($p<0.005$). TSS values decreased as the coarseness of the resolution increased. Each resolution was significantly different from the other (lsd <0.05; Figure 4).

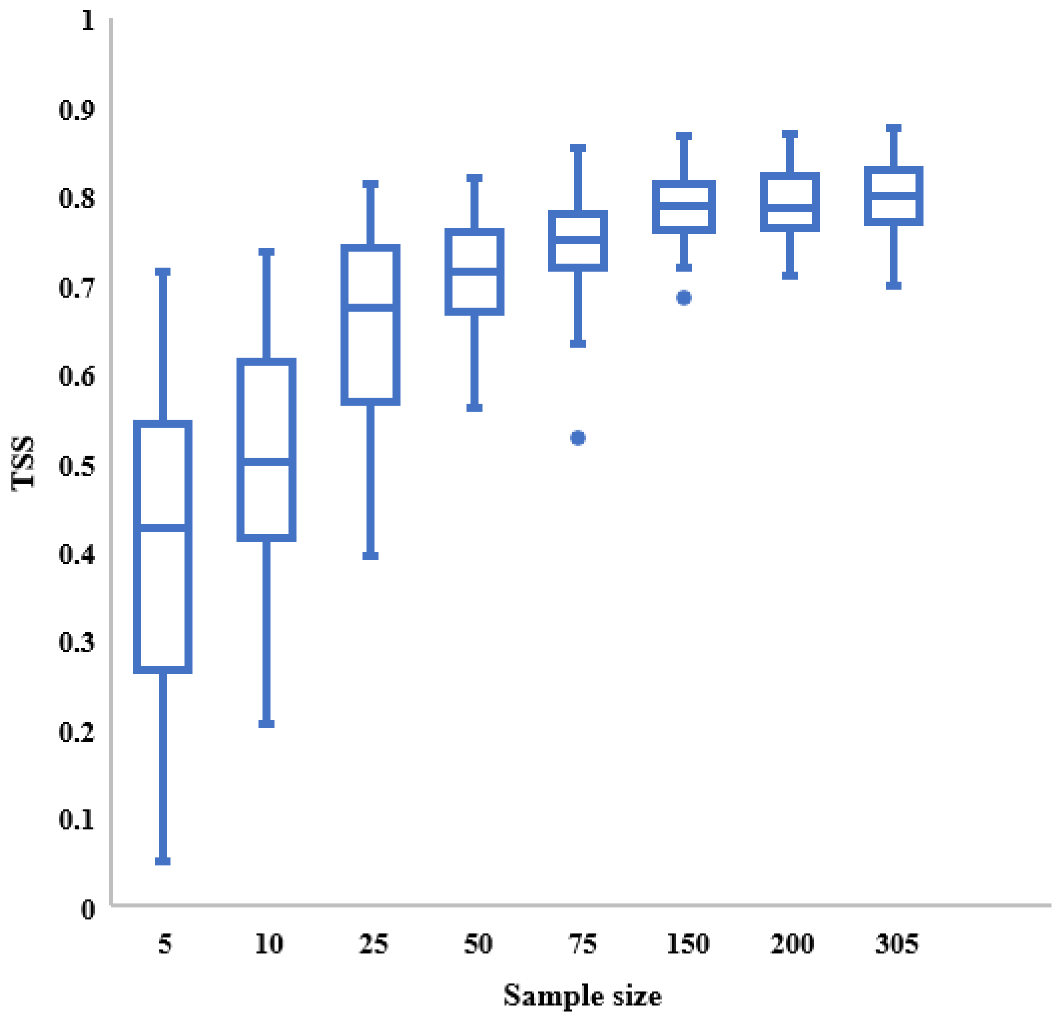


Figure 3: Effect of sample size on TSS

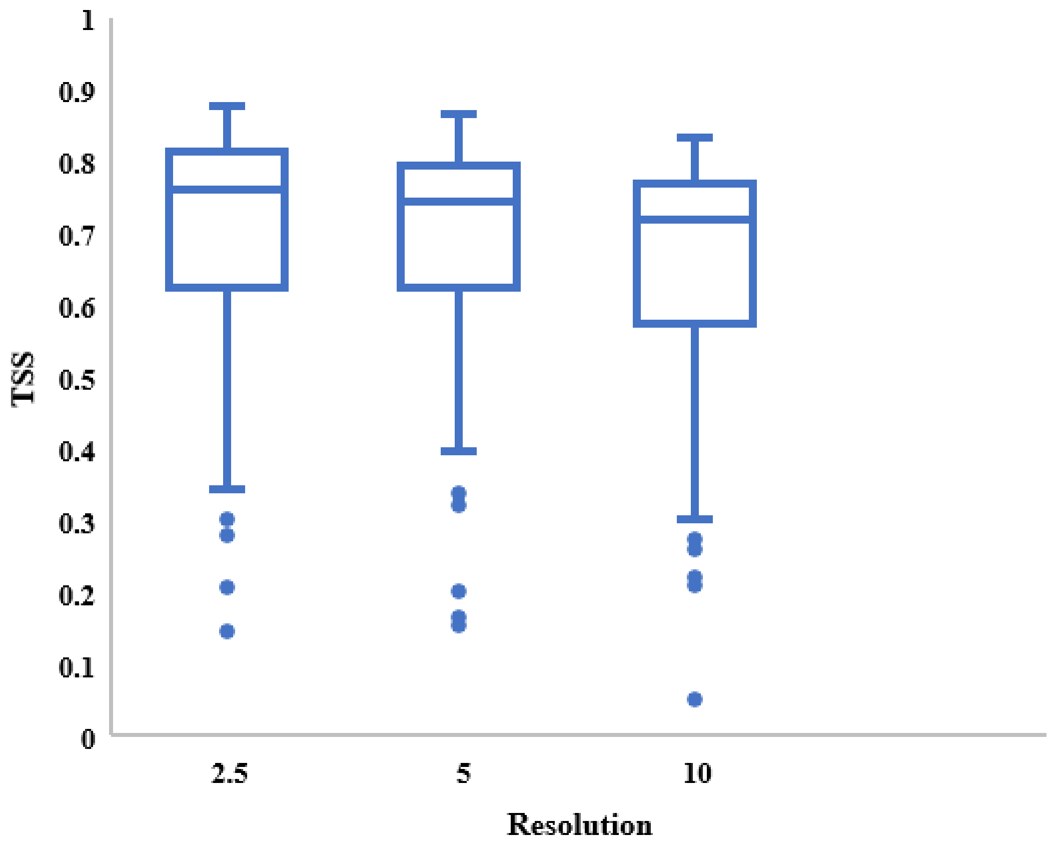


Figure 4: Effect of resolution on TSS

4.2.2 Kappa Statistic

The ANOVA of Kappa showed that the interaction between resolution and sample size was not significant ($p > 0.274$). However, sample size was significant ($p < 0.001$). Kappa values increased with increase in sample size and that the relationship with sample size was non-linear. increase in sample size. Smaller sample sizes (5, 10 and 25) were all significantly different ($p < 0.001$) from each other, however, sample size 25 compared with sample sizes from 50 up to 200 were not significant ($p > 0.05$). There was no significance between sample sizes 50, 75, 150 and 200, sample size 305 was significantly different from the rest of the sample sizes ($p < 0.001$) (Figure 5). The Kappa Statistic values ranged from 0.10 for sample size 5 to 0.30 for sample size 305. The results indicate that all sample sizes from 5 to 200 produced models with poor accuracies (Figure 5). Sample size 305 however produced a model with marginal accuracy.

Resolution was significant ($p < 0.001$) The values of Kappa decreased as coarseness of the resolution increased. Each level of resolution was significantly different from the others (LSD, $p < 0.05$) (Figure 6). Kappa values ranges from 0.15 to 0. 18 for resolution 10 and 2.5 arc minutes. This is an indication that, all resolutions produced models with very poor accuracies.

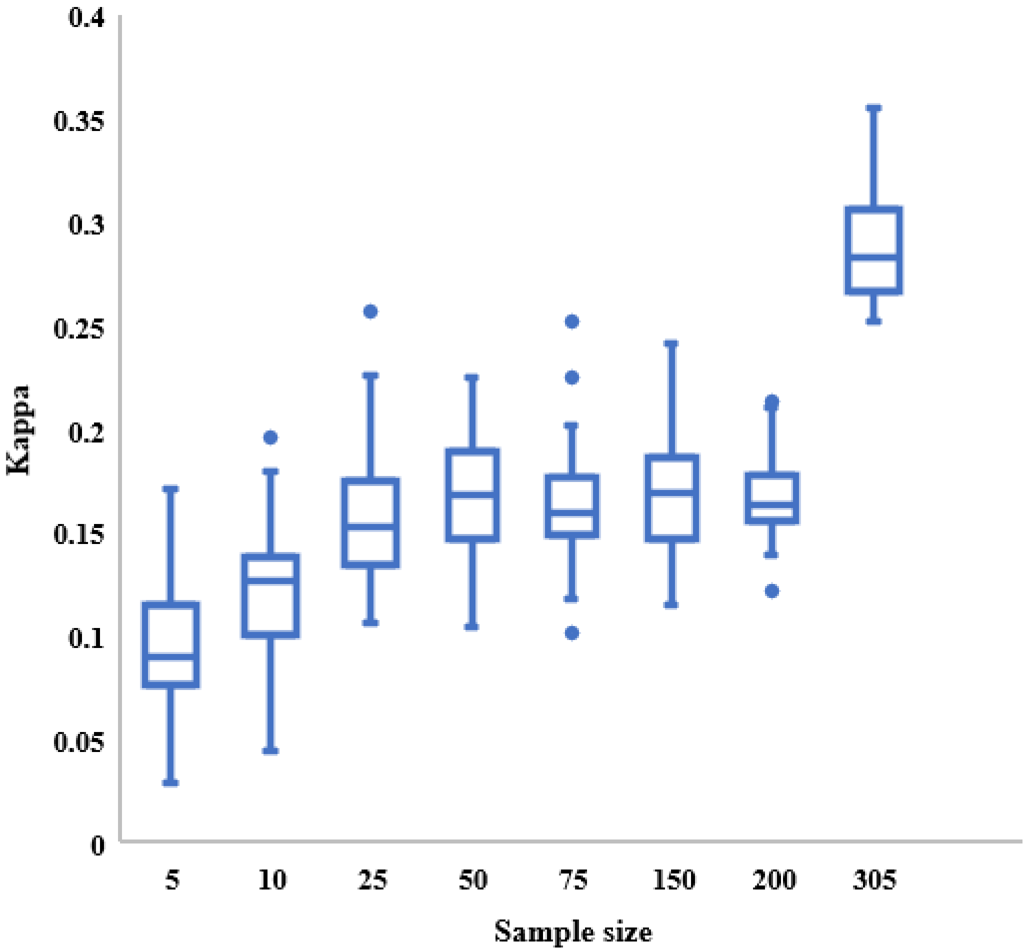


Figure 5: Effect of sample size on Kappa

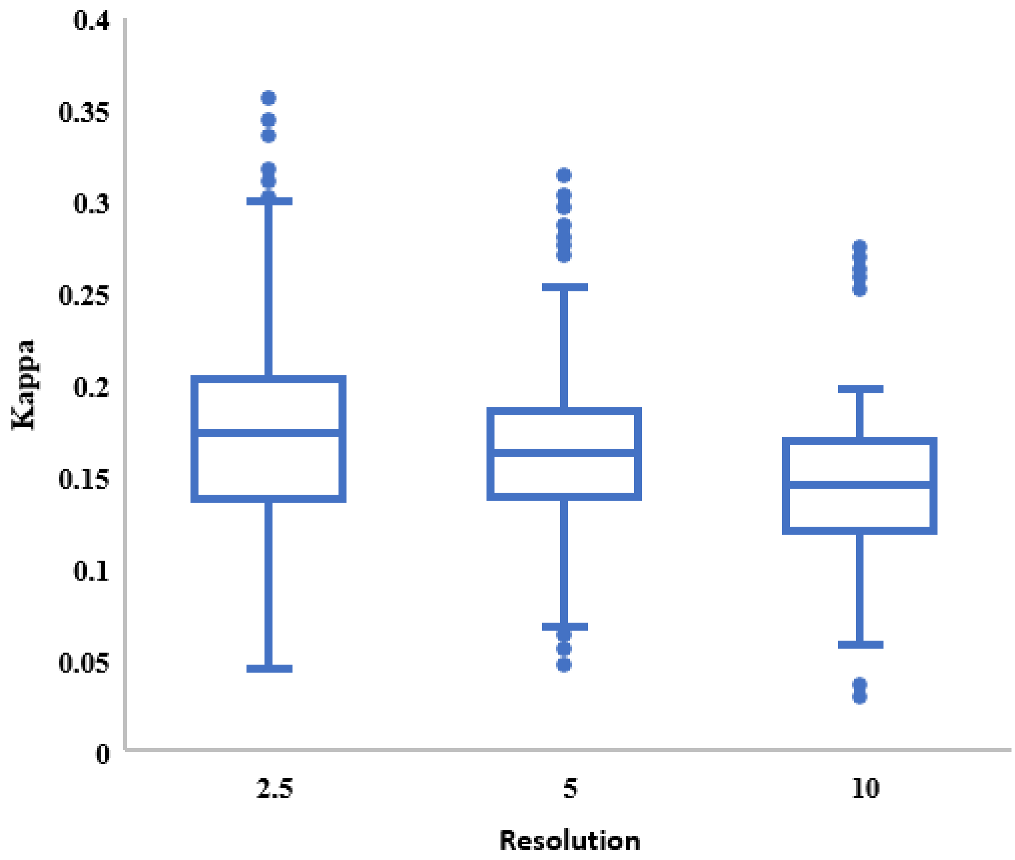


Figure 6: Effect of resolution on Kappa

4.2.3 Area Under the Curve (AUC)

The ANOVA of AUC showed that the interaction between the resolution and the sample size was not significant ($p>0.05$). The main effects of resolution and sample size were both significant ($p<0.001$) (Figure 7 and Figure 8). For sample size, AUC values increased with increase in sample size. All sample sizes were significantly different from each other except for sample size 200 and 305 which were not different from each other but had significantly higher AUC values than all other sample sizes (Figure 7).

For the resolution, AUC values decreased with increase in the coarseness of the resolutions. All three resolutions were significantly different from each other

(Figure 8). The results suggests that model accuracy increases with increase in sample size as well as increase in resolution at which the modelling was done.

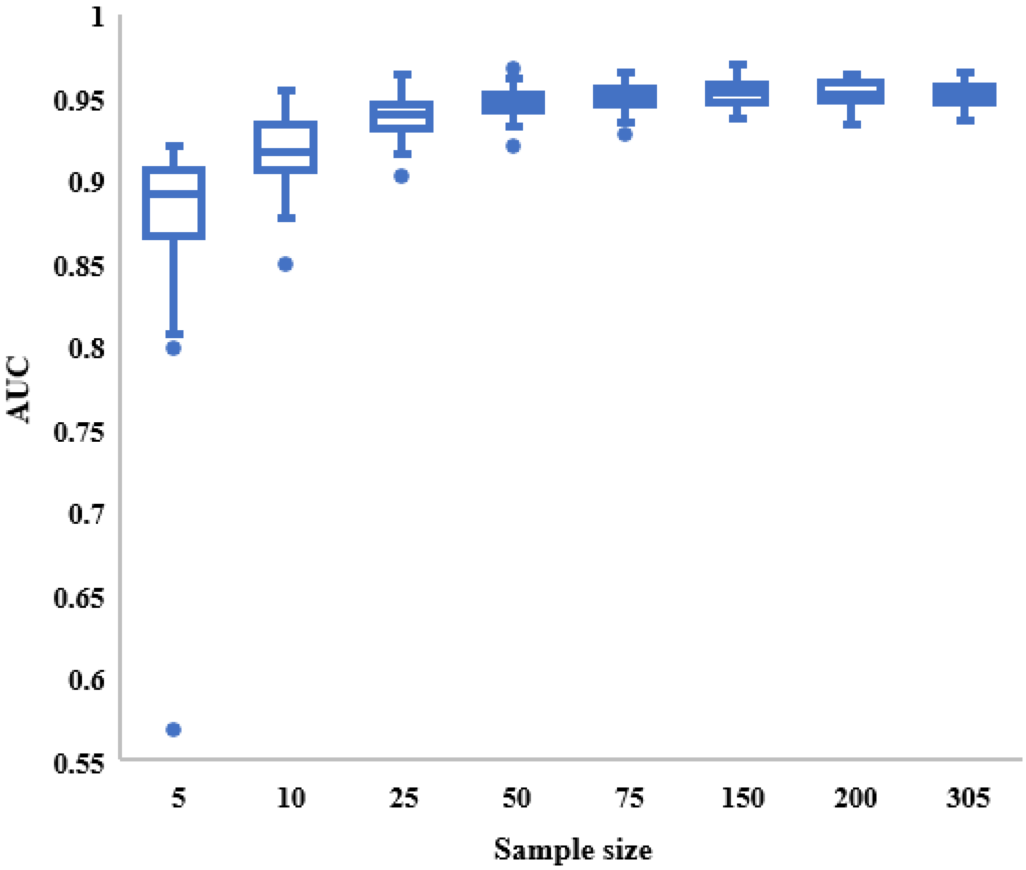


Figure 7: Effect of sample size on AUC

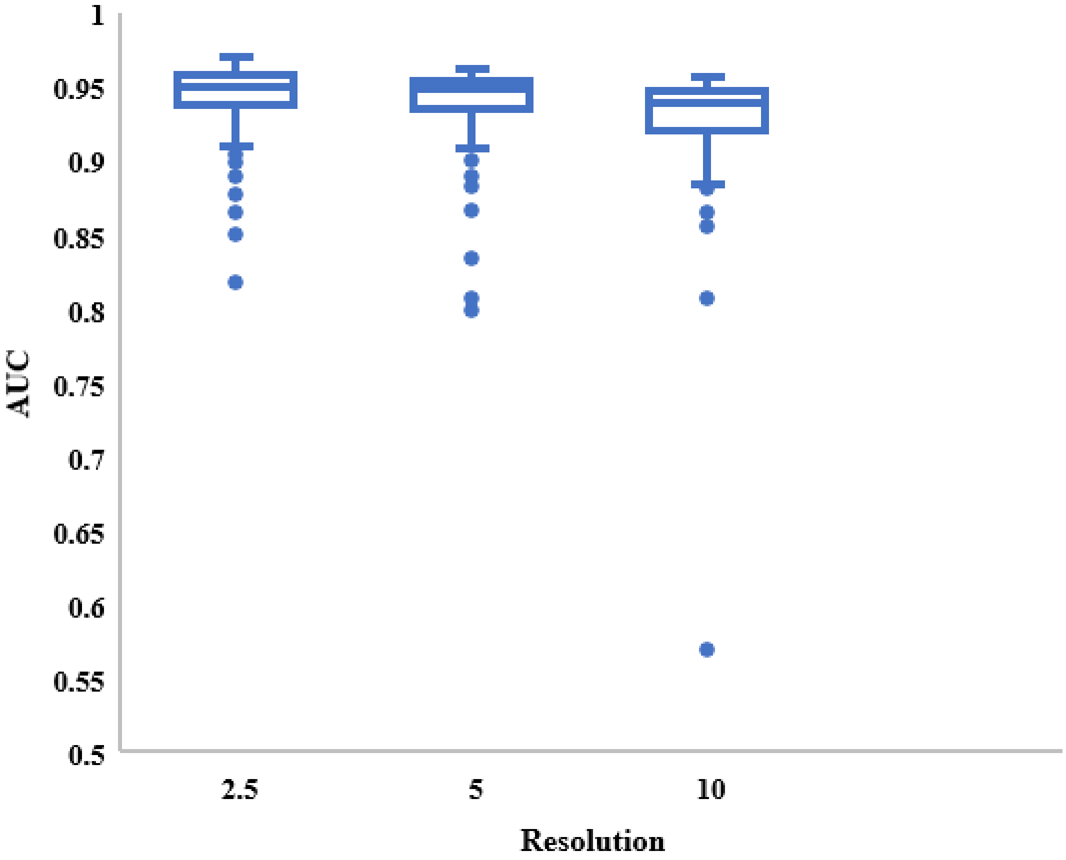


Figure 8: Effect of resolution on AUC

4.3 Pictorial Comparison of the Predictions from Sample Size 5 and 305 with the True Range

The following presentations are maps of the distribution of the species as modelled by MaxEnt. The minimum, average, maximum and the true range predicted areas for the least and the highest sample sizes were selected with resolution 5 arc minutes and are represented in the maps. These maps served to indicate the variability associated with predictions baseed on sample size 5 compared with predictions based on sample size 305. While there is very high variability in sample size 5, in predicting the range of the species, sample size 305 predicted a very stable range and closest to the true range of the species. For the minimum with sample size 5, (Figure 9), map A, there was very high underprediction while the maximum for sample size 5 (map C) over predicted the true range. Only the mean prediction (map B) predicted closest to the true range. However, maps A, B, C representing the minimum, the mean and the maximum predictions for sample size 305 all predicted similar range sizes, closet to the true range (Figure 10). This implies that, estimating the range size of the species with sample sizes with fewer records gives imprecise predictions.

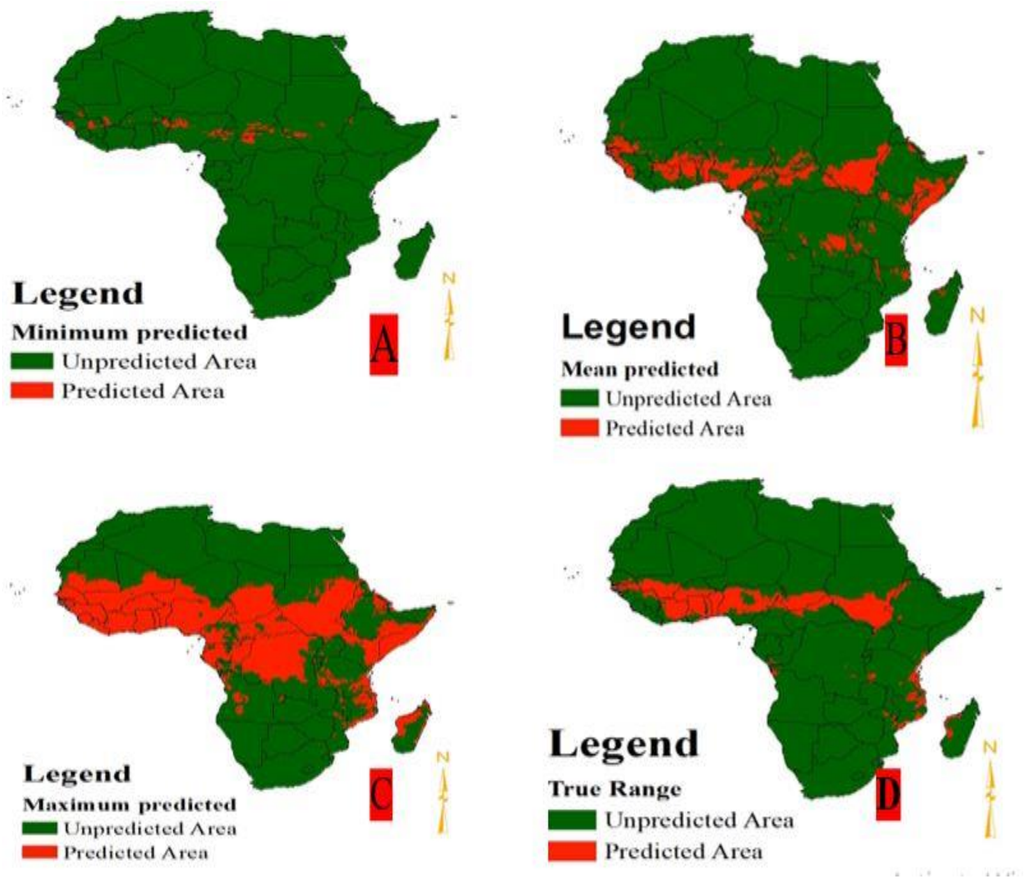


Figure 9: Maps of predicted (A=Min, B=Mean, C=Max, D=True range) distribution of *Parkia biglobosa* with sample size 5

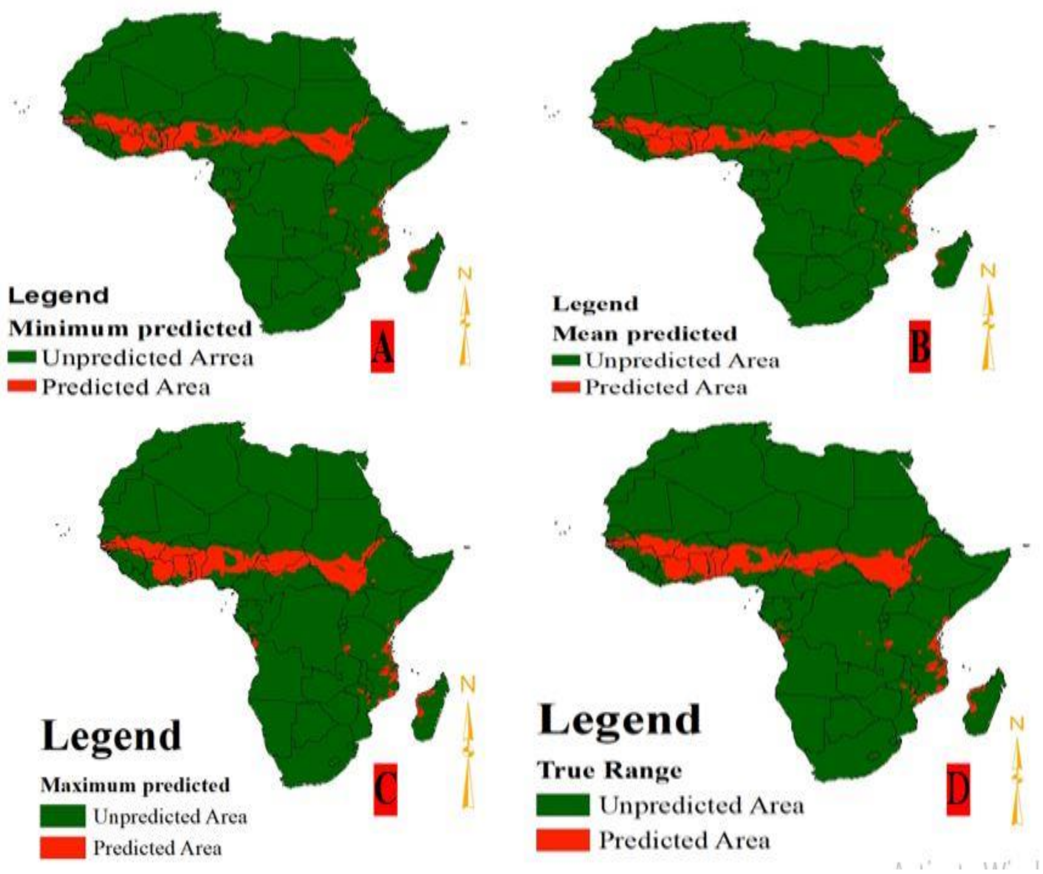


Figure 10: Maps of predicted (A=Min, B=Mean, C=Max, D=True range) distribution of *Parkia biglobosa* with sample size 305

4.4 The Effect of Climate Change on the Potential Distribution of *Parkia biglobosa*

A paired t-test between the present predicted range and the future predicted range showed that the range of *P. biglobosa* will likely expand by about 110% (t-test, $p<0.001$) by 2080.

Although this potential may exist there are obvious challenges that may obstruct its realization namely, the ability of the species to disperse to the new locations and human degradation of the environment through several activities.

4.5 Most Important Environmental Variable

The results from MaxEnt showed that, BIO 16 (Precipitation of the wettest quarter) was the most important variable determining the distribution of *Parkia biglobosa*. This was followed by BIO 11 (Mean annual temperature of the coldest quarter), and BIO 4 (Temperature seasonality) came third as the most important variable for models of the distribution of the species. However, the most important environmental variable varied with sample sizes. Large sample sizes (Table 3), shows that BIO16 was selected for the most important variable followed by BIO11 and BIO4 and this was consistent across all resolutions.

Table 3: Percentage contributions of environmental variables to determining the distribution of *Parkia biglobosa* with sample sizes 305 and resolution 5 and 2.5 arc minutes

SN	Variable	Percent (%) Contribution	Variable	Percent (%) Contribution
1	Bio16	31.8	Bio16	31.1
2	Bio11	16.4	Bio11	18.4
3	Bio4	12.4	Bio4	10.6
4	Bio19	6.8	Bio19	7.7
5	Bio12	5.9	Bio1	5.9
6	Soil texture	5.2	Bio9	5.4
7	Bio1	4.8	Bio12	5.1
8	Bio6	4.3	Bio6	3.5
9	Bio5	3	Soil texture	2.9
10	Bio9	1.9	Bio18	1.8
11	Bio10	1.5	Bio2	1.6
12	Bio2	1.4	Bio10	1.4
13	Bio18	1	Bio15	1.3
14	Bio15	0.9	Bio7	0.8
15	Bio14	0.9	Bio3	0.7
16	Soil organic matter (t_ocx100)	0.5	Soil bulk density (Bulk_dnx10)	0.6
17	Bio8	0.4	Soil organic matter (t_ocx100)	0.6
18	Bio13	0.2	Bio5	0.2
19	Bio17	0.2	Bio14	0.2
20	Bio3	0.2	Bio8	0.2
21	Bio7	0.1	Soil pH (t_phx10)	0.1
22	Soil bulk density (Bulk_dnx10)	0	Bio13	0.1
23	Soil pH(t_phx10)	0	Bio17	0

However, for small sample sizes, the most important environmental variable differed from resolution to resolutions. For example, in resolution 5, the most important environmental variables include BIO11, BIO14 and BIO15 and resolution 2.5 arcminutes includes BIO11, BIO16 and BIO18.

Table 4: Percentage contributions of environmental variables to determining the distribution of *Parkia biglobosa* with sample sizes 305 and resolution 5 and 2.5 arc minutes

SN	Variable	Percent (%) Contribution	Variable	Percent (%) Contribution
1	Bio11	58.6	Bio 11	61.2
2	Bio14	24.1	Bio 16	8.6
3	Bio16	8.5	Bio18	7.8
4	Bio15	2.8	Bio8	7.1
5	Bio8	2.6	Soil texture	3.1
6	Bio10	2.4	Bio13	2.1
7	Bio13	0.5	Bio3	1.9
8	Bio2	0.5	Bio14	1.8
9	Soil organic matter(t_ocx100)	0	Bio2	1.1
10	Bio19	0	Bio12	0.9
11	Bio18	0	Bio1	0.2
12	Bio17	0	Bio19	0
13	Bio12	0	Soil organic matter(t_ocx100)	0
14	Soil pH(t_phx10)	0	Bio17	0
15	Soil texture	0	Bio7	0
16	Soil bulk density (bulk_dnx10)	0	Bio15	0
17	Bio9	0	Bio9	0
18	Bio7	0	Soil bulk density (bulk_dnx10)	0

29	Bio6	0	Bio5	0
20	Bio5	0	Bio6	0
21	Bio4	0	Soil pH(t_phx10)	0
22	Bio3	0	Bio10	0
23	Bio1	0	Bio4	0

4.5.1 How Does the Most Important Variable Affect the Distribution of *Parkia biglobosa* (Response of The Species to The Variables)

Parkia biglobosa respond well to precipitation of the wettest quarter. Below 400mm of precipitation, *Parkia biglobosa* has low probability of occurrence (<0.5, Figure 11). At about 700mm, *Parkia biglobosa* attains the highest probability (0.7) of occurrence. However, excessive rainfall beyond 700mm diminishes its cloglog probability of occurrence only marginally. The small sample size (Figure 12: The effect of precipitation of the coldest quarter (mm) on cloglog probability of occurrence of Parkia biglobosa for sample size 5), shows a similar situation.

Error! Reference source not found., Figure 14 and Figure 16). For the precipitation of the wettest quartile (BIO 16), Figure 11 showed that. *P. biglobosa* is quite tolerant to low rainfall (0.6 probability of presence at 500mm). However, it reaches its maximum probability presence (1.0) at 1500mm. Further increase in rain does not harm the species as its probability of occurrence remains constant at 1.0.

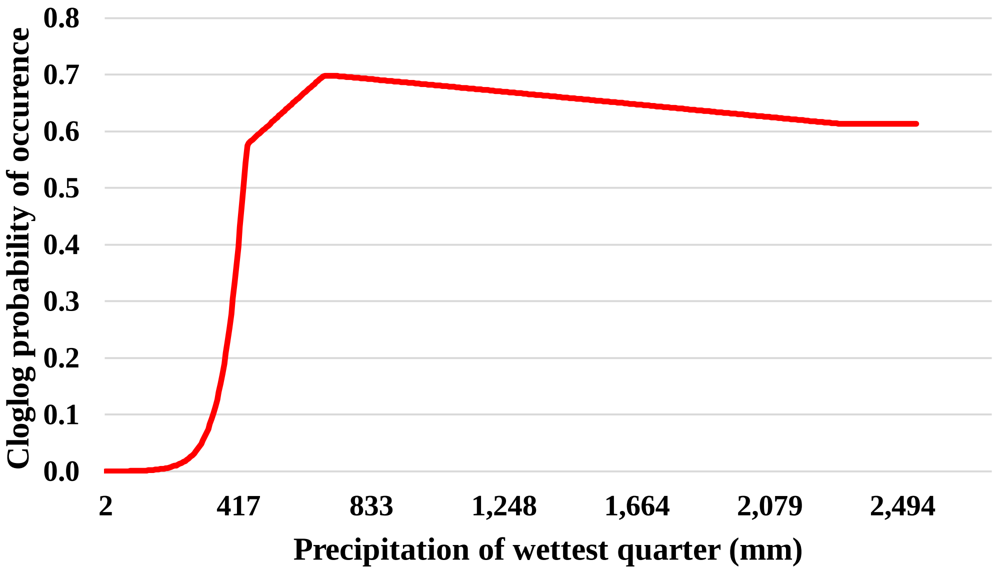


Figure 11: The effect of precipitation of the coldest quarter (mm) on cloglog probability of occurrence of *Parkia biglobosa* for sample size 305

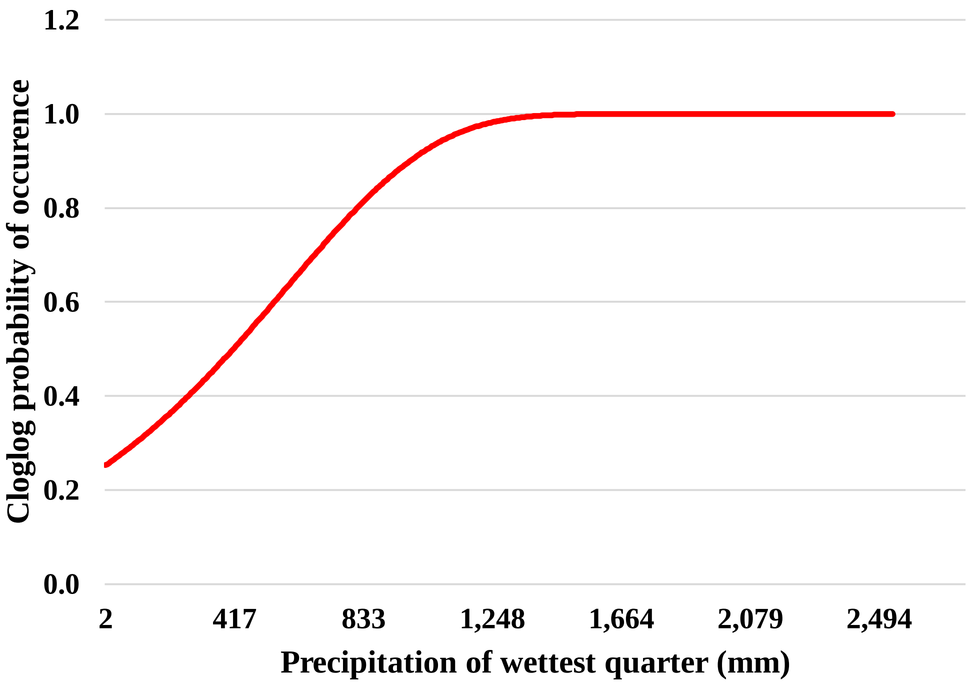


Figure 12: The effect of precipitation of the coldest quarter (mm) on cloglog probability of occurrence of Parkia biglobosa for sample size 5

The second most important variable, (Figure 14) shows how BIO11, the temperature of the coldest quarter affects the distribution of Parkia biglobosa. Parkia biglobosa is very sensitive to temperature of the coldest quarter and only a very narrow range (24-27°C) is optimal for its growth where the probability of occurrence is above 0.5. The probability of occurrence declines rapidly outside of this temperature range. For sample size 5 (Figure 13) Parkia biglobosa appeared to be insensitive to high temperature. However, this might be due to a statistical artefact rather the real situation because of the extremely small sample size.

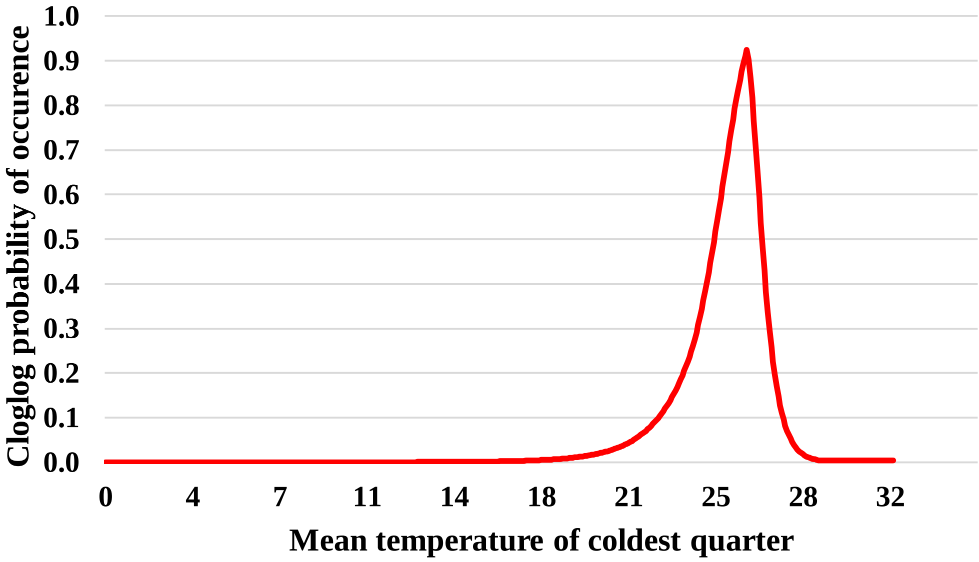


Figure 13: The effect of mean temperature of the coldest quarter (°C) on the cloglog probability of occurrence of *Parkia biglobosa* for sample size 305

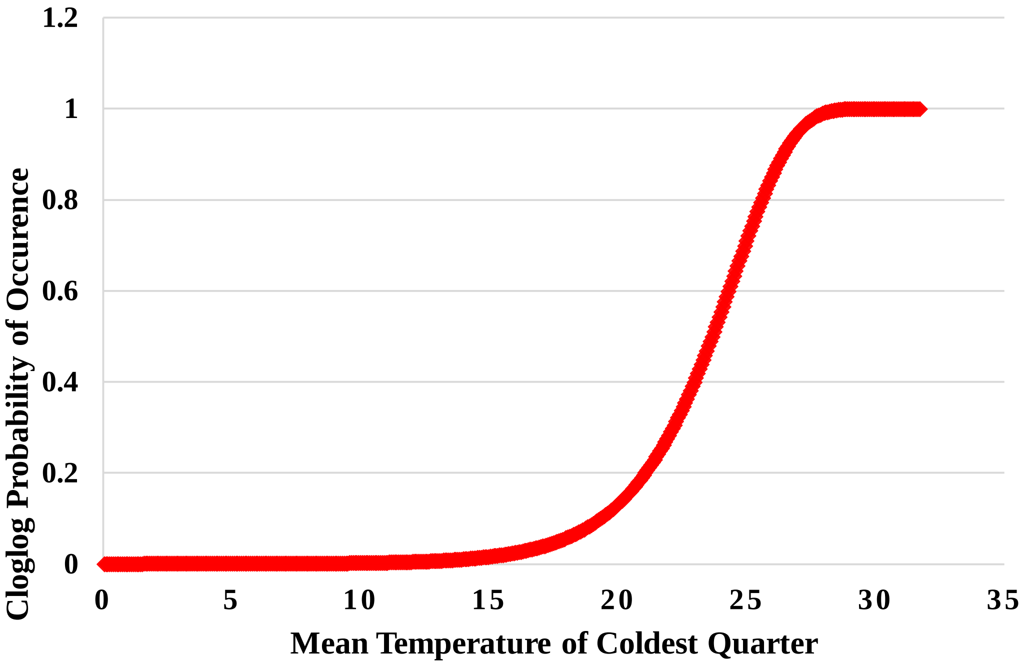


Figure 14: The effect of mean temperature of the coldest quarter (°C) on the cloglog probability of occurrence of Parkia biglobosa for sample size 5

P. biglobosa is also sensitive to temperature seasonality (standard deviation of mean annual temperature) Figure 15. The optimum range suitable for *P. biglobosa* is between 10 -17 standard deviation unit. This means that extreme fluctuations in temperature may not be suitable for the existence of *P. biglobosa* (Figure 16). Figure 16 depicts temperature seasonality for sample size 5. The differences in the graphs may be attributed to the small sample size of Figure 16.

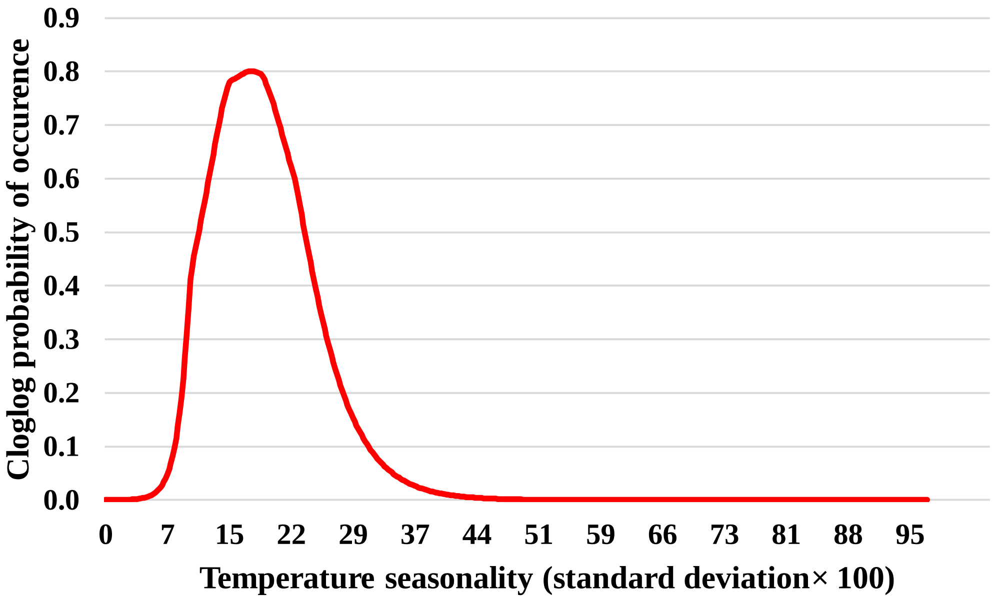


Figure 15: The effect of temperature seasonality (Standard deviation × 100) on the cloglog probability of occurrence of *Parkia biglobosa* for sample size 305

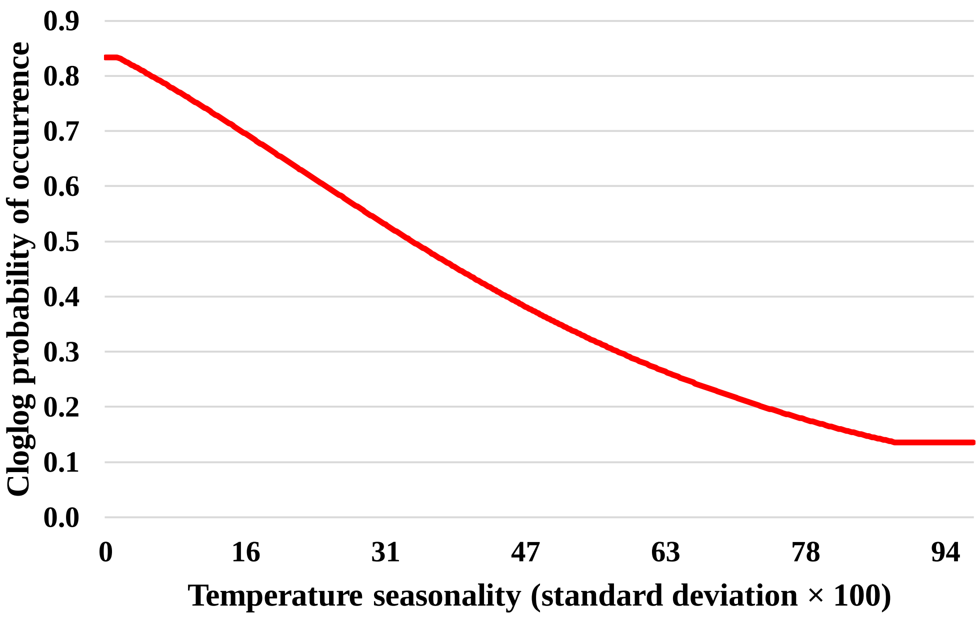


Figure 16: The effect of temperature seasonality (Standard deviation × 100) on the cloglog probability of occurrence of *Parkia biglobosa* for sample size 5

4.6 Experiment Two: Modelling of the Virtual Species

4.6.1 Range Prediction

To determine the extent to which sample size and resolution affects range prediction, virtual species data was analysed and compared with the results obtained from the GBIF dataset. Analysis was based on percentage deviation of predicted ranges from the true range of the virtual species.

The interaction between sample size and resolution was not significant ($p>0.05$). The three resolutions were not significantly different from each other ($p = 0.333$; Figure 17). The three different resolutions predicted an average of 20% below the real range. Sample sizes were significantly different from each other ($p< 0.001$; Figure 18). However, all the sample sizes resulted in under prediction of the true range. The amount of under prediction reduced as the sample size increase. For sample sizes below 50, there was high variability in the ranges predicted, for example, for sample size 5, the range of predictions varied from about 80% under prediction to 200% over prediction (Figure 18). For sample sizes (75-305), the under predictions ranged from 22% to 12% respectively.

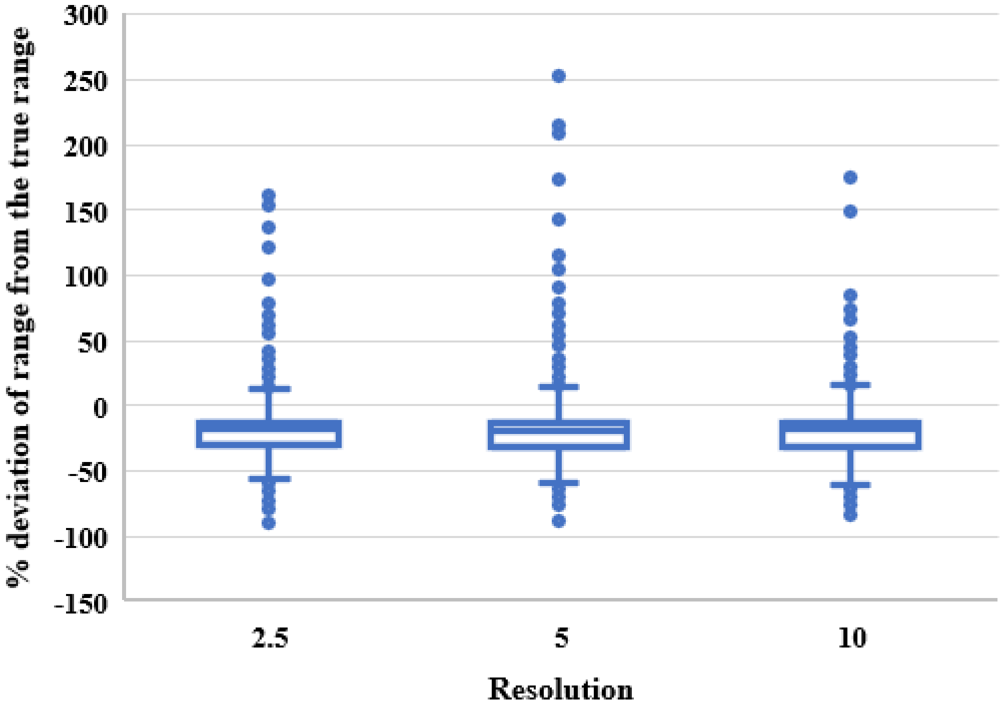


Figure 17: The effect of resolution on the range predicted

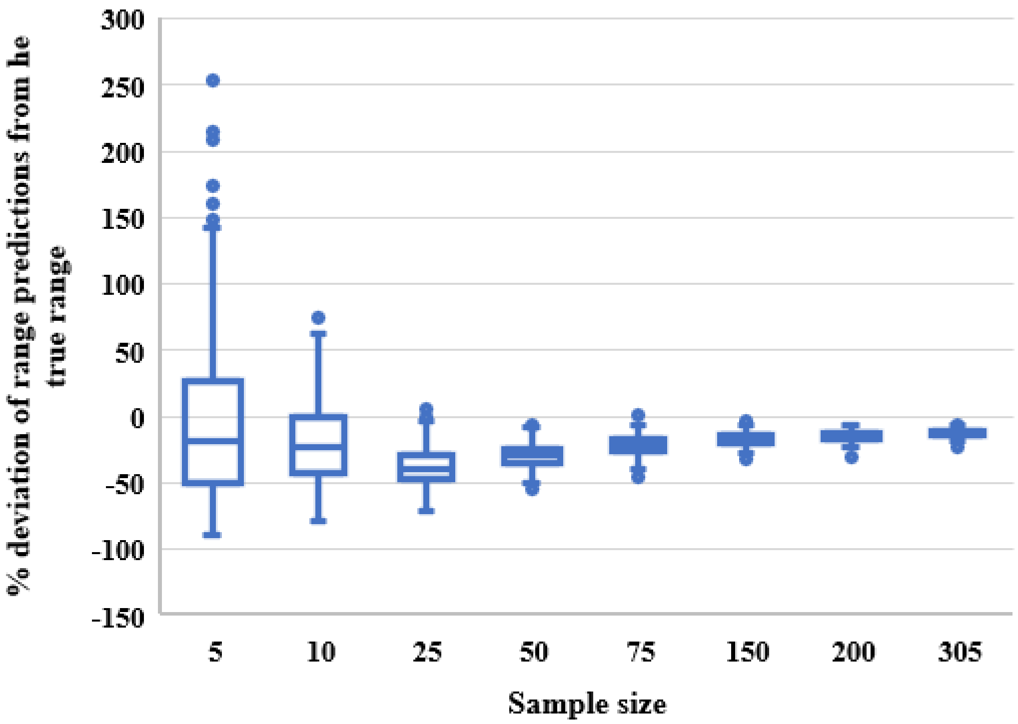


Figure 18: The effect of sample size on the range predicted

4.7 Sample Size and Resolution on The Accuracy of The Modelling Distribution (Model Accuracy Assessment)

4.7.1 The Kappa Statistic for Virtual Species

The ANOVA showed a significant interaction between resolution and sample size with respect to the Kappa Statistic. The interaction indicated that for small sample sizes, the resolutions did not dffer significantly from each other (5-50). For sample sizes above 50, the higher resolutions have higher Kappa values than the lower resolutions (Figure 19).

For sample size it was clear that Kappa values increased as sample size increased (Figure 19) over the entire range of ther sample size. The mean Kappa values range from 0.23 for sample size 5 to 0.44 for sample size 305. This suggest that all sample sizes produced models with fair to good acuracy.

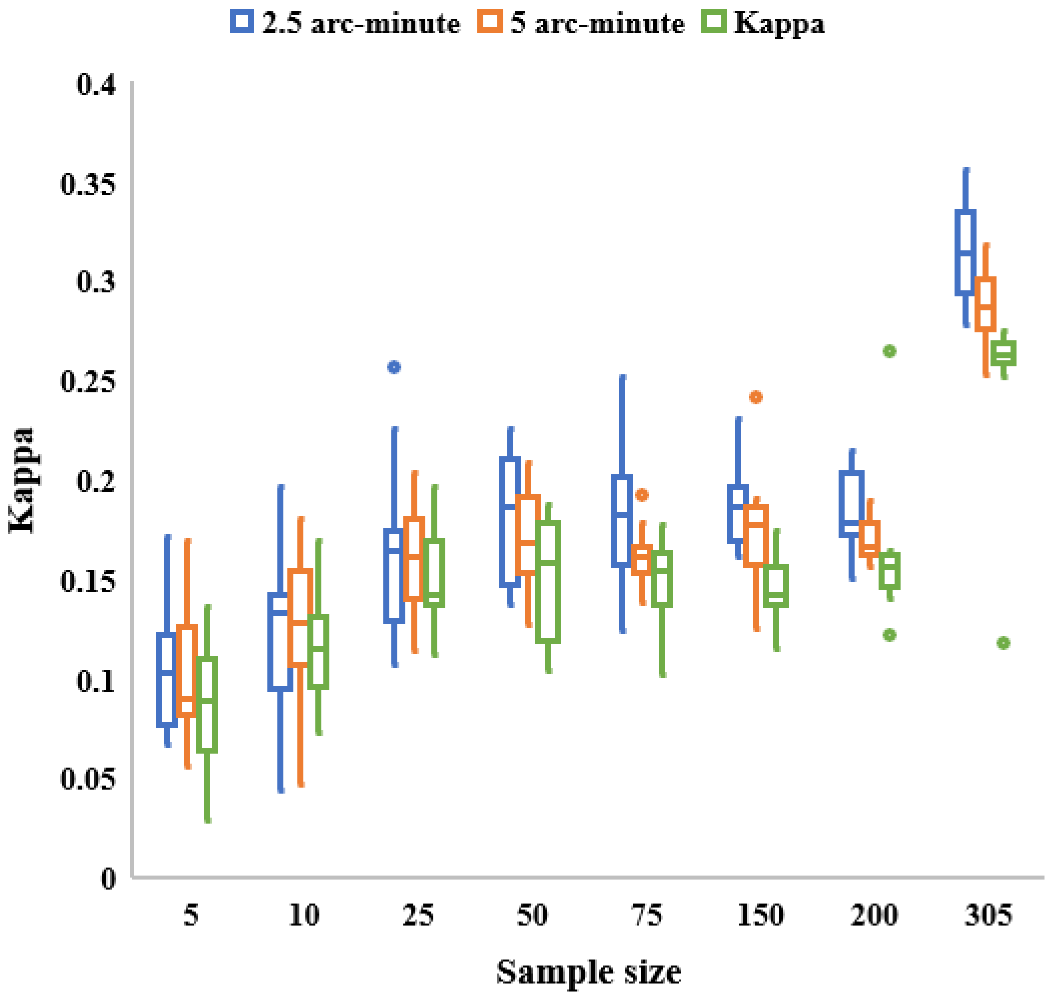


Figure 19: Interaction effect of resolution and sample size on Kappa

4.7.2 True Skill Statistic (TSS) for Virtual Species

The result of TSS shows that, the interaction effect of sample size and resolution was not significant ($p > 0.197$). TSS increased with increase in sample size (Figure 21). All mean differences were significantly different from each other. This suggest that model accuracy increased as sample size increased. The mean TSS values ranged from 0.38 for sample size 5 to 0.76 for sample size 305. This indicates that all sample sizes produced accurate models based on TSS.

It was also noted that at small sample sizes (5, 10, 25), there was high variability in TSS values which may also suggest that accuracy measurement was imprecise for smaller sample sizes (Figure 20)

Base on the ANOVA, the resolutions were significantly different from each other ($p = 0.038$). Separating the means by LSD, resolution 2.5 and 5 arc minutes were not significantly different from each other (Figure 20; $p>0.05$). However, 10 arc minutes had significantly higher TSS value than the other two.

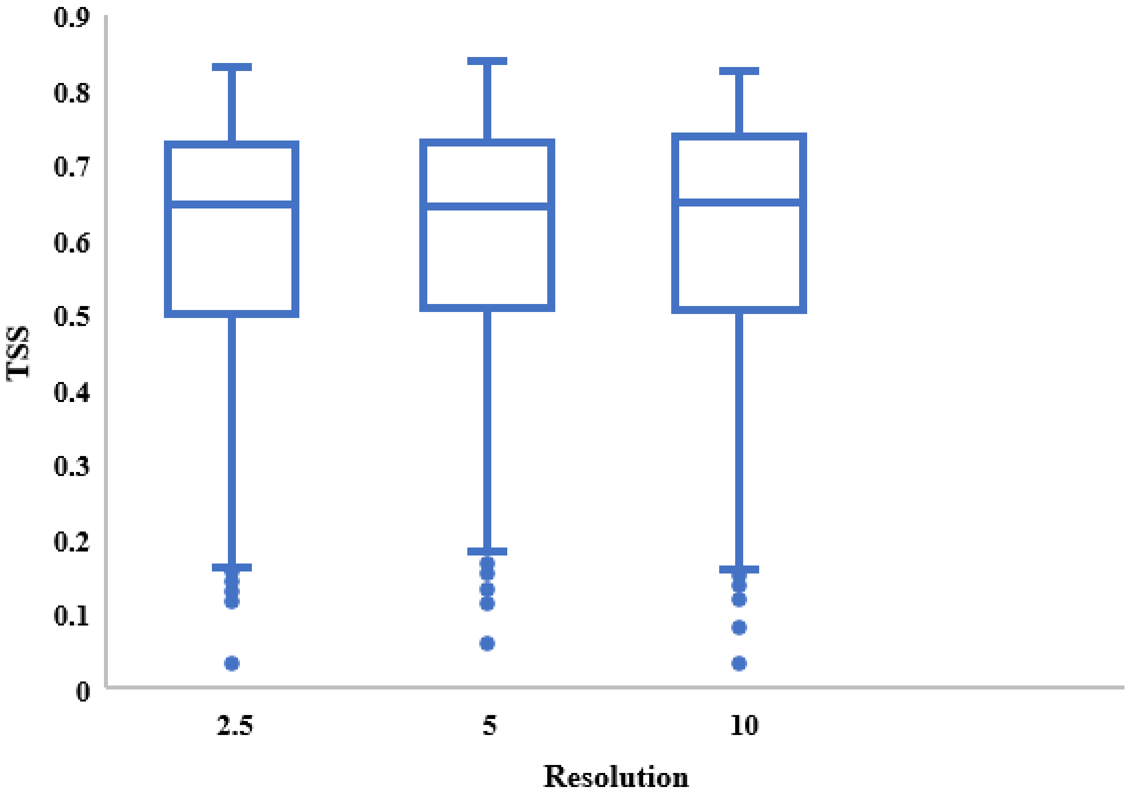


Figure 20: The effect of resolution on TSS

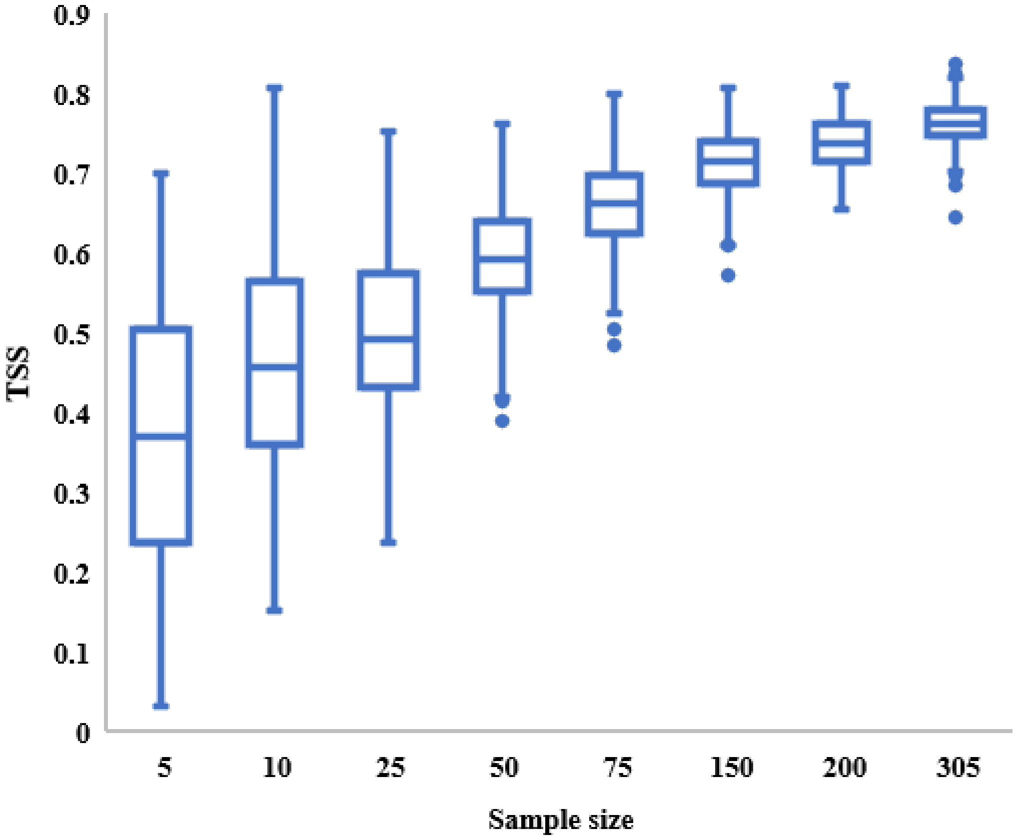


Figure 21: The effect of sample size on TSS

4.8 The Area Under the Curve for Virtual Species

The interaction between the resolution and the sample size was significant ($p < 0.003$; Figure 22). The interaction showed that at sample sizes 5 and 10, the different resolutions were not significantly different from each other but sample sizes 25 to 305; coarser resolutions gave larger values of AUC than smaller sample sizes. The sample size was also significantly ($p<0.001$). With sample sizes, AUC rose from 0.87 at sample size 5 to 0.94 at sample size 75. Then it decreased gradually to 0.93 at sample size 305. Sample size 75 could therefore be regarded as an optimal sample size where the models had maximum accuracy.

The mean AUC values ranged from 0.854 for sample size 5 to 0.936 for sample size 305. Based on these values, it appeared that all the sample sizes were able to produce fairly accurate models since an AUC value of 0.70 is often regarded as a good model.

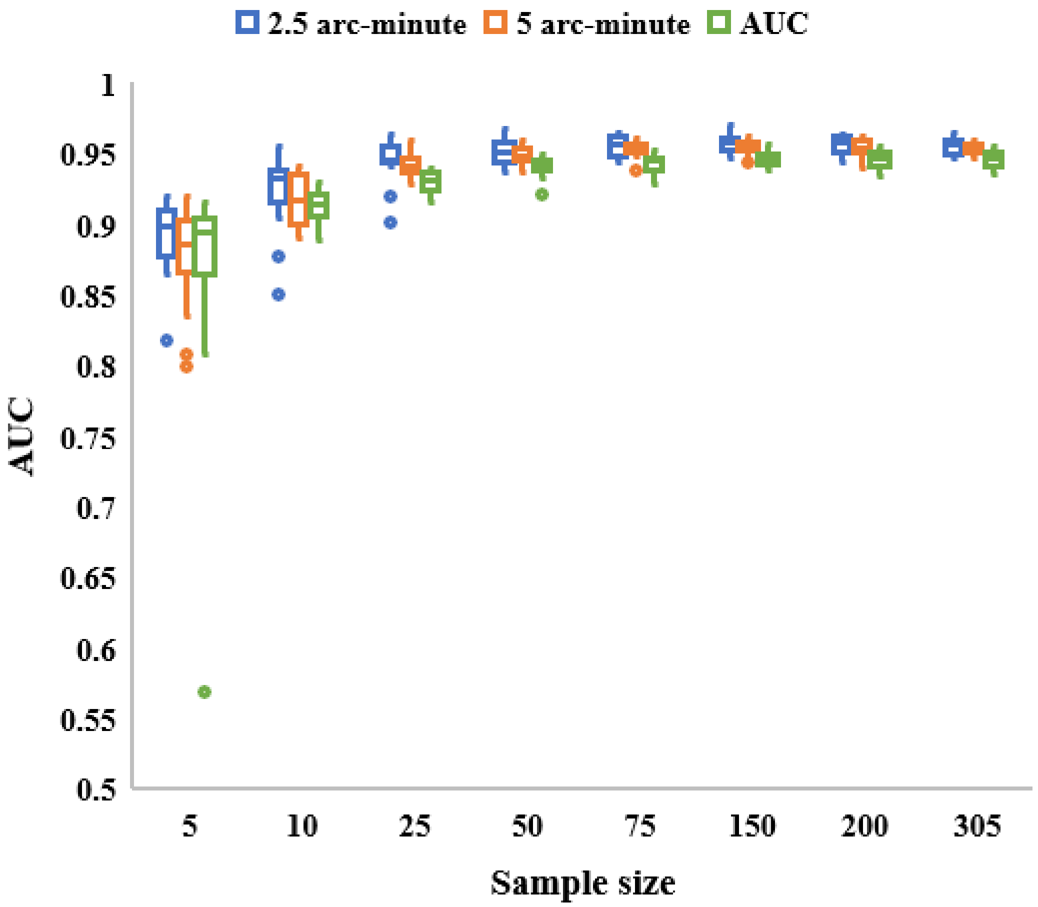


Figure 22: Interaction effect of resolution and sample size on AUC

4.9 Comparison of Standard Accuracy (AUC, TSS and Kappa) with Jaccard's Similarity Index

The ANOVA of Jaccard's Similarity Index showed that the interaction between resolution and sample size was not significant ($p=0.068$). The main effects of resolution and sample sizes were both significant ($p<0.05$). For the sample size, Jaccard's Similarity Index increased as the sample sizes increased (Figure 23). This suggests that larger sample sizes predicted the real range better than the smaller sample sizes. It was also evident that smaller sample sizes produced high variability in Jaccard's Similarity Index implying these predictions were imprecise.

The mean values of JSI varied from 0.31 (s.d. = 0.095) for sample size 5 to 0.83 (s.d. = 0.019) for sample size 305. This clearly indicated that on average models based on sample size 5 predicted ranges which had only about 30% overlap with the true range of the species while models based on sample size 305 predicted ranges which on average overlapped the true range about 80%. Models with sample size 5 could not be regarded as good models since they predicted only 30% of the true range.

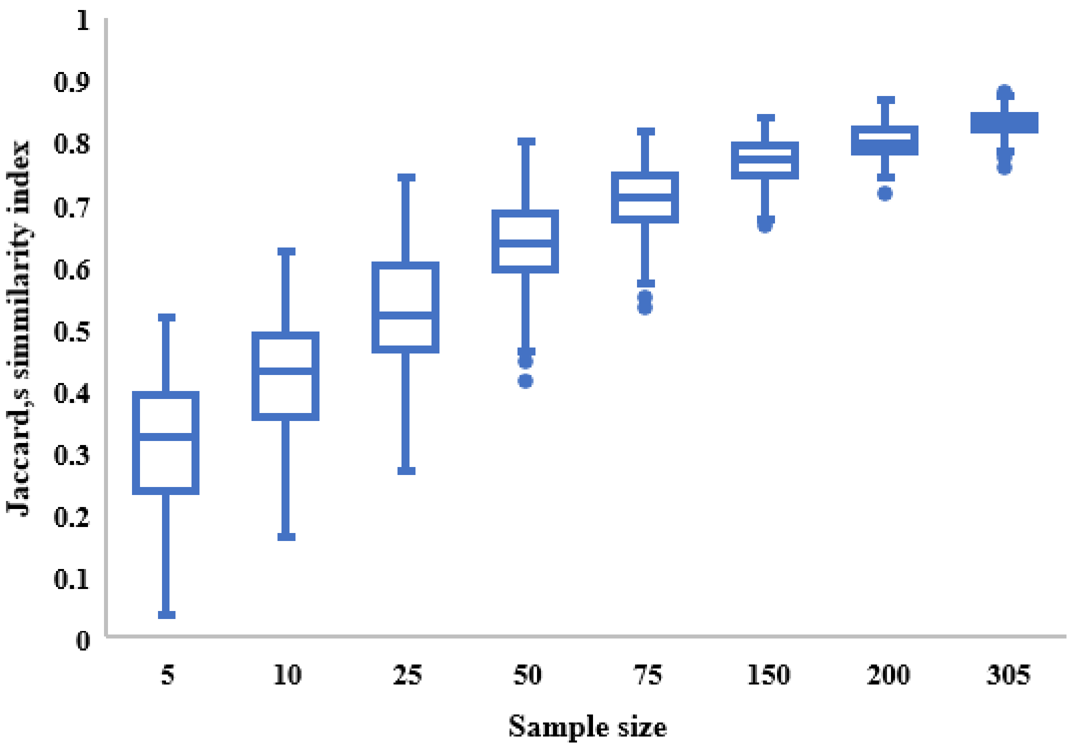


Figure 23: Effect of sample size on Jaccard Similarity Index

For the resolution, Jaccard's Similarity Index decreased as the grids became coarser (Figure 24) since Jaccard's similarity measures the overlap between the predicted ranges and the real range, this result shows that, the finer the resolution, the better the overlap between predictions and real area probably meaning more accurate predictions.

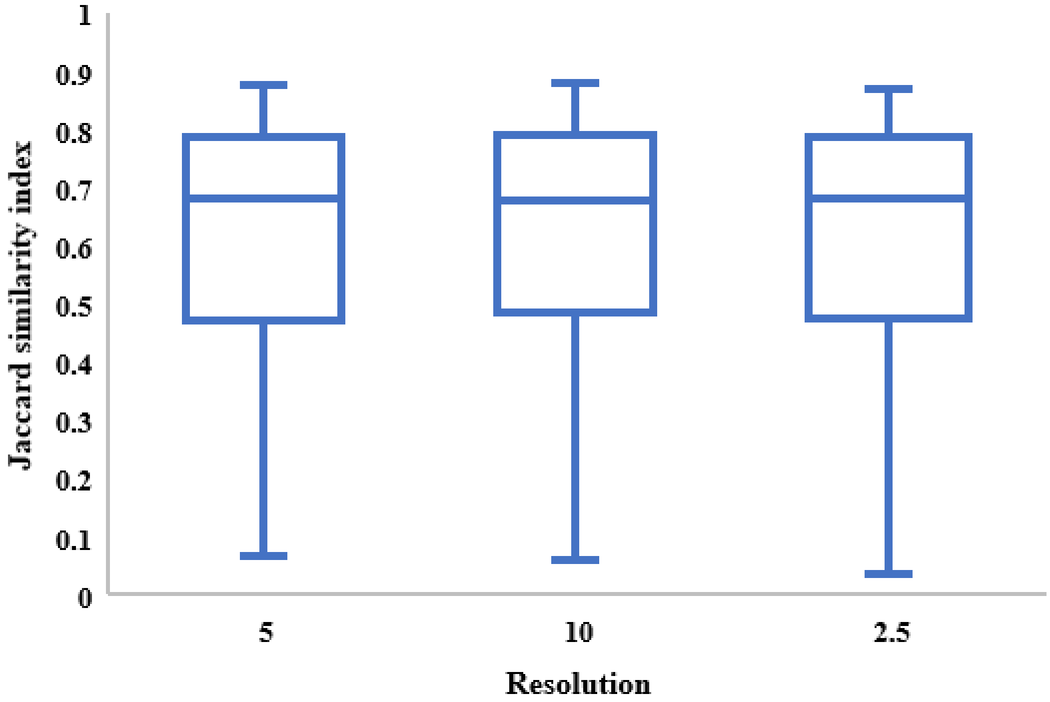


Figure 24: Effect of resolution on Jaccard Similarity Index

CHPATER FIVE

DISCUSSION

5.1 Present Predicted Range

Prediction of range size or areas using the maximum entropy (MaxEnt) algorithm has been reported to produce more representative ranges than α-hull convex polygon method when sample sizes are small (Pena et al., 2014). Results derived from such methods (convex polygon) includes areas that may be unsuitable for the species, or ignore suitable areas; a situation species distribution models may correct for (Pena et al., 2014).

MaxEnt models are often selected for good model performance even with small sample sizes. For instance, van Proosdij, Sosef, Wieringa, & Raes, (2016) found sample size as low as three sufficient for species in narrow ranges and 13 for wide spread species to predict reliable ranges, though small sample sizes have been reported to be unable to predict reliable ranges (Mateo et al., 2010; Wisz et al., 2008). Larger sample sizes produced more reliable range predictions (Feeley & Silman, 2011). The range predictions in this study showed that small sample sizes (< 50) showed very wide variability (Figure 1). For example, for sample size 5, the highest predicted range was 19.8 times larger than the minimum predicted range (compared with 1.6 times for sample size 305). This makes the predictions imprecise irrespective of the mean for small sample sizes.

Some studies have indicated that, high resolutions produce small predicted ranges while lower resolutions give larger predicted ranges (for example; Gaston, 2003; Soe et al, 2009). Using a high resolution for species with narrow ranges capture more details of important environmental variables (Connor et al., 2018). A study by Jiménez-Alfaro, Draper, & Nogués-Bravo (2012) showed that, while high resolution models produced more robust models than the lower resolutions, they nonetheless predict significantly lower ranges, even when the threshold based on the minimal predicted range was used. That is, resolution with larger grid cells (> 1 arc minutes) predicts larger areas especially when large number of occurrence records are used. Results of this study (Figure 1) have shown for small sample sizes, the effect of resolution may be indistinguishable, however, for large sample sizes, high resolutions predicted smaller ranges than lower resolutions as found in (Gaston et al., 2003); Jiménez-Alfaro et al., (2012) and Guisan et al., (2007).

From this study, it was observed that studies of sample size and resolution were conducted independently from each other. This study combined the two and found that these factors do not act independently. The range size predicted when studying resolution depends on sample size. For example, Jiménez-Alfaro et al., (2012), observed that lower resolutions predicted larger ranges while higher resolutions predicted smaller ranges. At the same time Guisan et al., (2007) observed no significant effect of resolution on range predictions. Both were observed in this study but depended on the sample size. It is possible that their different conclusions might have been influenced by sample sizes. For smaller sample size (≤50) one might conclude that resolutions are not important in range prediction when studying resolutions (Figure 1). However, for larger samples size (>50), it might be concluded that, coarser resolutions predict significantly larger rangers than finer resolution (Figure 1).

5.1.1 Future Predicted Range

Prediction of range size (future) is based on assumptions made by algorithms chosen in the modelling of the species (Pearson et al., 2006). MaxEnt assumes that the range of suitable environmental variables will be occupied by the species (Phillips et al., 2006).

The result showed that, the interaction between the sample size and the resolution were significant ($p<0.01$). This suggest that predicting far into the future accurately would be a difficult issue since it would depend on both the sample size and the resolution at which the modeling is done. The results also indicated that as the sample size increased, the predicted range decreased (Figure 2). This intend adds to the complexity of predicting future range since sample size alone might lead to either under-prediction or over-prediction of the future true range. Likewise, given a specific sample size, the resolution alone may also lead to over-prediction or over prediction of the future range. Finally, the result showed that with small sample sizes (≤ 25), future range was very variable and imprecise (Figure 2). This may suggest that, for reliable predictions of the future range, sample sizes above 25 may be required.

5.2 Accuracy of Predictions of Models as Affected by Sample size and Resolution

The interaction effect of sample size and resolution on the accuracy of models has not been widely reported by the literature except few like the work of (Soultan & Safi, 2017). This study measured the interacting effect of sample size and spatial resolution on model accuracy using metrics such as the True Skill Statistic (TSS), Area Under Curve (AUC) and Kappa Statistic (Kappa).

5.2.1 True Skill Statistic

TSS is one discriminatory metric that test the accuracy of species distribution models such that, TSS values for a model greater than 0.2 but less than 0.6 are considered fairly moderate accuracy, and values greater than 0.6 are considered good models (ALLOUCHE et al., 2006).

From the results (Figure 3), TSS values increased with increasing sample size. At lower sample sizes, the estimates of accuracy were variable.

Some studies such as that of van Proosdij, et al, (2016), have reported that, small sample size as low as 3 produced an accurate model, other studies were contrary to this findings. For example, McPherson, Jetz, & Rogers, (2004), reported increasing model accuracy with increasing sample size though Jiménez-Alfaro et al., (2012), have reported good model accuracies with small sample size.

For sample sizes, the TSS values ranged from 0.41 for sample size 5 to 0.80 for sample size 305. Accuracies increased with increasing sample sizes, but showed no significant differences between sample size 200 and 305 (Figure 3). However, for sample sizes less than 50 although TSS showed a fairly accurate value, they were very variable (Figure 3), this makes the values imprecise.

For the resolution, the mean TSS values were 0.69, 0.67 and 0.65 for resolution 2.5, 5 and 10 arc minutes respectively. TSS values decreased as the coarseness of the resolution increased. The resolutions were significantly different from the other (Figure 4).

The higher the resolution, the higher the accuracy of the models. Reduction in model accuracy as resolutions decreases has been reported by (Connor et al., 2018; Antoine Guisan et al., 2007; Lee et al., 2004) among others. In most studies, the impact of models developed with lower resolution only degrades the models and make them inaccurate. The accuracy of models in Figure 3 of the results conforms with findings of these authors including (Connor et al., 2018; Antoine Guisan et al., 2007; Lee et al., 2004). However, these authors did not report that small sample sizes (<50) produced imprecise predictions of accuracy.

5.2.2 Kappa Statistic

Kappa statistic values ranges from -1 to +1, with +1 indicating perfect agreement and values ≤ 0 considered as performance no better than random. Kappa values ≤ 0.4 are fairly accurate and values > 0.4 are good models (Landis & Koch, 1977).

In Figure 5, accuracy increased with increasing sample size. Higher resolution predicted higher values of Kappa than lower resolutions. For sample size, Kappa values were 0.1 for sample size 5 and 0.23 for sample size 305. This is an indication that Kappa Statistic produced very poor accuracy for all models except for sample size 305 which had a marginal accuracy.

The accuracy of models based on kappa statistics is greater with species with larger sample sizes (McPherson et al., 2004). These authors called for the use of different methods in assessing the accuracy of models with small sample sizes. Similarly, Bean, Stafford, & Brashares, (2012), disagreed with studies which suggest that accurate models could be produced from small sample sizes, making the claim that, models built with small sample sizes without a fair knowledge of the true range of the species only complicated the confidence of its accuracy. This study confirms the results of Bean et al., (2012), however, the results of this study appear to be conservative since only a sample size above 300 are considered large.

Kappa Statistic values for the resolutions are 0.15, 0.17 and 0.18 for resolutions 10, 5 and 2.5 arc minutes respectively. Lowering the resolution produces less accuracies (Guisan et al., 2007). For resolution, kappa predicted all models to be questionable accuracy (all were below 0.2).

5.2.3 Area Under Curve (AUC)

Models with AUC values of 0.5 and below are models which do not predict any better than random models. (Alberto Jiménez-Valverde, 2012), values of 1 indicate perfect models (Xiaoping Liu et al., 2017) while values above 0.7 are often considered as good models (Drew et al., 2011). Results of this study (Figure 7) showed that, the AUC values ranged from 0.88 for sample size 5 and 0.95 for sample size 305, indicating that, the models predicted very good accuracies.

Higher resolutions have been, reported to show better accuracies while the lower resolutions produce less accuracy (Connor et al., 2018; A. Guisan et al., 2007; Jiménez-Alfaro et al., 2012). Results of this study (Figure 7) corroborate with findings of (Connor et al., 2018; Antoine Guisan et al., 2007; Jiménez-Alfaro et al., 2012), with resolution 2.5 arc minute predicting higher than 5 and 10 arc minutes respectively.

For AUC, all the models in relation to sample size were good. Implications may be, that if smaller sample sizes (<5) were modelled such as van Proosdij et al., (2016) stated 3 sample sizes, they could still have produced good models which makes it more complicated for chosen a threshold sample size for modelling.

5.3 The Effect of Climate Change on The Potential Distribution of *Parkia biglobosa*

Climate change is not only detrimental but beneficial from time to time. An important reason for studies into climate change is especially geared towards mitigation and economic changes. Though one may expect that the effect of the current climate will have negative impact on *Parkia biglobosa*, results from this study showed a significant increase in the range size of the species with a likelihood expansion for the future predicted range (That is, t-test ($p<0.001$) between the present and future predicted range gave about a 110% expansion in the future range by 2080).

This suggest that, though the climate have or may have changed, bioclimatic variables which affects the distribution more are likely to be more beneficial to the species as they might not be affected beyond the current climate. The implication of this result may be severe as there are obvious challenges that may militate its realization. For example, the ability of the species to disperse to the new locations and human degradation of the environment through several activities. This results also, clearly indicate that the species is not under threat to be in the global red list of species that are endangered due to the effect of climate change.

5.4 Most Important Climate Variables

Variable importance changes with increasing resolution (Connor et al., 2018). Modeling at higher resolutions for species specialist probably capture in details difference in environmental variables (Connor et al., 2018). In this study, the sample sizes affected the most important environment variables for the species. Small sample sizes (< 25) produced different most important environmental variables from different models. This finding is similar to Connor et al, (2018). This result may be due to incomplete environmental variable representation in the small sample size (Kadmon et al., 2016).

5.8 How Does the Most Important Variable Affect the Distribution of *Parkia Biglobosa*

The first most important environmental variable affecting the distribution of *Parkia biglobosa* differs with sample size. For large sample size (>50), precipitation of the coldest quarter (BIO16) was selected as the most important variable (Figure 11). The results indicated that, the optimum rainfall for *Parkia biglobosa* is around 700mm during the coldest quarter. In some parts of the range especially in west Africa, the coldest quarter often coincides with the peak of the rainy season. This implies that, the amount of wet season rain is an important determinant factor and from the graph (Figure 11) (Br & Don, 2021) a minimum of 400mm of precipitation is required for the existence for the *Parkia biglobosa.*

The second most important environmental variable was the mean temperature of the coldest quarter (BIO11). From the result (**Error! Reference source not found.**), *Parkia biglobosa* has a narrow range of temperature for the coldest quarter (24 – 27^{0}C). However, the species has a wide geographical range (Dotchamou et al., 2016) which may suggest that it may tolerate temperatures outside this range particularly during warm weather (Br & Don, 2021).

5.8.1 Effect of The Least (5) And Highest (305) Sample Sizes Compared to The True Range of The Species

Many studies have investigated the effect small sample sizes have on the range predicted. However, none have shown (graphically) the extent to which sample sizes affect the prediction of range at a minimum, average and maximum levels. The variability in range prediction made by the small sample size (Figure 9) is a deviation from the predicted true range of the species. Conclusions made on range size and smaller sample sizes however must be carefully studied to avoid under estimation or over estimation.

In Figure 10, range predictions are about the same size of the predicted true range. This corroborate with earlier conclusions and findings that; stable areas are predicted for the species with larger records than lesser ones.

5.9 Experiment Two (Virtual Species)

5.9.1 Range Prediction

Models were validated with virtual species and the ranges were estimated as the deviation of range predictions from true range, where the true range is equal to zero (0). Results of the study showed that the resolution was not significant. The three resolutions underpredicted the true range by an average of 20%. However, the sample sizes were significantly different from each other ($p>0.01$). The sample sizes also underpredicted the true range of the species. It appeared in this experiment, maxent predictions of ranges were smaller than the true range.

Soultan & Safi, (2017), concluded that predicting useful distribution ranges, samples sizes with few records as few as 10 could be representative. Though there have been several suggestions on the minimum required number of record for modelling using different algorithms and MaxEnt algorithm for example used by Wisz et al., (2008), suggested sample size with a minimum records of 30.

In this study, based on the sample size studied, it is clear that maxent underpredicted the true range irrespective of the sample size used. However, it appeared the larger the sample size, the closer the prediction is to the true range.

In this study, smaller sample sizes (< 50) gave very high variability in the ranges predicted, for example, for sample size 5, the range predictions vary from about 80% under prediction and 200% over prediction (Figure 18). These findings corroborate to the results found from the GBIF data. From the result, the amount of under prediction reduced as the sample size increase, implying samples with more records are more likely to predict ranges close to the true range of the species.

5.10 Accuracy of Model Prediction of Predictions as Affected by Sample size and Resolution

5.10.1 Kappa Statistic

Accuracy assessments made using Kappa statistic values ranges from -1 to +1, with +1 indicating perfect agreement and values ≤ 0 considered as performance no better than random. Kappa values ≤ 0.4 are fairly accurate and values > 0.4 are good models (Landis & Koch, 1977).

Studies by Antoine Guisan et al.,(2007) reported that, lowering the resolution for building models makes the models less accurate. From this study, higher resolutions result in higher accuracies for the models than lower resolutions (Figure 19). This result is in line with Guisan et al., (2007).

McPherson et al., (2004) reported that Kappa statistic increases with increasing sample sizes. Bean, Stafford, & Brashares, (2012), also agreed with the assertion of (McPherson et al., 2004). Results from this study agreed with (Bean et al., 2012; McPherson et al., 2004) (Figure 19).

Kappa values in Figure 19 showed that small sample sizes (<50) had Kappa values below 0.40 which cannot be considered as good models (Landis & Koch, 1977).

5.10.2 True Skill Statistic (TSS)

TSS is one discriminatory metric that test the accuracy of species distribution models such that, TSS values for a model greater than 0.2 but less than (< 0.6) are considered fairly moderate accuracy, and values greater than 0.6 (> 0.6) are considered good models (ALLOUCHE et al., 2006).

From the results (Figure 21), TSS values increased with increasing sample size. For lower sample sizes (5-25), TSS values were very variable. The results also indicated that, resolution 10 arc minutes had significantly higher TSS values than resolution 2.5 and 5 arc minutes. For small sample sizes, (5-25), there was high variability in TSS values.

Higher resolutions have been reported to produce good accuracies; that is, the higher the resolution, the higher the accuracy of the models (Connor et al., 2018; Antoine Guisan et al., 2007; Lee et al., 2004). Studies that reported higher accuracies with increase in resolution. Similarly, it has been reported of the loss important of variables which lead to the overestimations at lower resolutions (Connor et al., 2018). It is also known that the impact of models developed with lower resolution only make them inaccurate (Connor et al., 2018). However, accuracy of models in this study (Figure 21) suggested that TSS values increased with reduction in resolution. This deviation may be an inability for some important variables to be accounted for, perhaps similar to the report of Connor et al., (2018).

Alberto Jiménez-Valverde, (2012); van Proosdij et al., (2016), have reported that, small sample size as low as 3 could produce accurate models. However, McPherson, Jetz, & Rogers, (2004), reported the minimum sample size should be at least 30.

In this study, the TSS values varied from 0.38 for sample size 5 to 0.76 for sample size 305. For this study, sample size 50 to 305 produced TSS values above 0.60 which can be considered as good models (ALLOUCHE et al., 2006). Sample sizes below 50 are only random to no agreements (0.38-0.50).

5.10.3 Area Under Curve (AUC)

Models with AUC values of 0.5 and below are models which do not predict any better than random models. (Alberto Jiménez-Valverde, 2012), values of 1 indicate perfect models (Xiaoping Liu et al., 2017) while values above 0.7 are often considered as good models (Drew et al., 2011). Results of this study (Figure 25) showed very high values for all samples sizes (above 0.8). Though MaxEnt AUC values had been shown to decrease with increase in sample size in some studies including (Raes, & Steege, 2007; van Proosdij et al., 2016), accuracy increased slightly with increasing sample size as accuracy have also been reported to increases with a small increase in sample size (Loiselle et al., 2008; Wisz et al., 2008).

Large spread of AUC are indications of poor accuracies while smaller spreads show better accuracy (van Proosdij et al., 2016). Higher resolution has been, in many studies shown better accuracies while lowering the resolution decreases the accuracy (Connor et al., 2018; A. Guisan et al., 2007; Jiménez-Alfaro et al., 2012). Results of this study (Figure 22) corroborate with findings of authors stated above such that, the interaction between resolution and sample size (Figure 22) showed the variability in predicting accuracy with smaller sample sizes and having resolution 10 to predict higher accuracy than higher resolutions.

For each sample size, (except sample sizes 5), higher resolutions outperformed the lower resolutions. Predictions of the effect of resolution models of sample sizes < 25 had similar accuracies especially for sample size 5, which confirms the dependency of spatial resolution on the number of occurrence records.

Implications could be that, stronger accuracy are observed with large sample sizes and at lower resolutions for species with similar geographic characteristics as the one studied.

5.11 Comparison of Standard Accuracy (AUC, TSS and Kappa) with Jaccard's Similarity Index

Jaccard's Similarity Index measures the overlap (0-1) between the true range and the predicted ranges of the models fitted with the data generated by the factorial combination of the resolution and the sample size. Hence, it is the most accurate reproduction of how closely each model predicted the virtual species true range because a value of 0 represents no overlap at all and a value of 1 represents perfect reproduction of the range of the virtual species. Therefore, Jaccard's similarity index can be used to assess the traditional measures of the goodness of fit of models, namely AUC, TSS and Kappa. These measures have been criticized in different ways ranging from whether they are even appropriate to their accuracy measures in species distribution models (Alberto Jiménez-Valverde et al., 2008; Lobo et al., 2008, 2010). In these studies, it was shown that AUC was a misleading measure of accuracy mainly because it was designed to be used for presence-absence data and not for presence-only data (Alberto Jiménez-Valverde, 2012; Lobo et al., 2010; Peterson et al., 2008). Yet AUC is consistently used for models with presence only data. The other measures are vulnerable to the same arguments, i.e., without good quality presence absence data, their accuracy is in question (B. Leroy et al., 2017; Somodi et al., 2017). These concerns required an empirical evaluation of the accuracy of the models for which the extent of overlap (Jaccard's Similarity Index) was used.

The results of the measures of accuracy based on AUC, TSS, and Kappa compared with Jaccard's similarity index reveals important information. Taking AUC for example, all sample sizes produced models with AUC values above 0.70 (0.85 -

0.94) which may be considered as good models (Raes et al., 2007). However, JSI indicated that for models based on a sample size of 5, although the mean AUC value was 0.85, predicted ranges had only about 30% overlap with the true range of the virtual species. In real life situations such models would be obviously misleading and cannot be regarded as good models. It might appear that AUC in this context might be providing inflated accuracy for models with sample size 5. This property of AUC had be reported earlier (Bayor, 2012; Veloz, 2009). Models for sample size 305, however, could be regarded as good models since the ranges they predicted overlapped with the true range on average 80%.

The mean values of kappa ranged from 0.23 - 0.44. Based on Landis & Koch (1977), models with Kappa values above 0.4 are regarded as good models. In this study, this would coincide with sample size 50 and above and an overlap of 64% to 83% using JSI.

For TSS, the values regard as representing good models must be above 0.6 (ALLOUCHE et al., 2006). In this study, the sample size of models with TSS values with 0.6 and above are 50 and above. Therefore, only models with sample sizes above 50 may be regarded as good models as regard to TSS. The JSI showed that this corresponds to 64% to 83% overlap between the ranges predicted by the models and the true range of the species.

CHAPTER SIX

CONCLUSION AND RECOMMENDATION

6.1 Conclusion

This study showed that MaxEnt can be used to predict the present and future ranges of *Parkia biglobosa* successfully. *Parkia biglobosa,* does not appear to be adversely affected in the future (2080) under climate change. The predictions indicate that there is a potential for expansion of its range by about 110%. Sample size affects both the precision of predictions and the accuracy of predictions of MaxEnt models. Small sample sizes (<50) produced less precise predictions with high variability while larger sample sizes produced more precise predictions. Accuracy of models increased with increasing sample size as measured with AUC, TSS and Kappa. The effect of resolution does not appear to affect the precision of prediction but it affected the accuracy of prediction. The effect of resolution on the accuracy of predictions depended on sample size in this study. For smaller sample sizes (<50), the effect of resolution was not clear, probably because of the high variability in the predictions. However, for larger sample sizes (≥50), it was clear that higher resolutions produced more accurate models than lower resolutions. Climate change is likely to improve the potential for range expansion for *Parkia biglobosa*. However, given that the seeds of *Parkia biglobosa* are used for various products for human consumption and therefore actively harvested, coupled with environmental degradation, human assisted dispersal into the suitable ranges may be required to achieve this potential. The three most important environmental variables that affect the distribution of *Parkia biglobosa* into the future are; precipitation of wettest quarter, mean temperature of coldest quarter and temperature seasonality respectively. When the standard accuracy measures of species distribution models (AUC, TSS and Kappa) were compared with Jaccard Similarity Index, the results are interesting. It appears AUC values were overpredicting the accuracy of MaxEnt models while Kappa values were probably underpredicting accuracy of MaxEnt models. This clearly shows that the values of the traditional accuracy measures (AUC, TSS and Kappa) do not necessarily correspond with the accuracy of range prediction.

The results and conclusion of this experiment must be interpreted cautiously since the premise is based on a snapshot of data retrieved from the Global Biodiversity Information Facility at a specific point in time. GBIF data is continuously increasing and may change significantly with time.

6.2 Recommendation

Since the species has the potential to expand, human beings can facilitate the expansion by actively aiding dispersal into the new ranges.

Species distribution modellers using small sample sizes (<50) should interpret their results cautiously since both range prediction and accuracy may be imprecise. It is therefore, suggested that model evaluation should use a large number of replications. It is further suggested that where possible, higher resolutions should be used for species distribution modelling.

REFERENCES

Abdulhamid, B., Dawaki, S., & Dau, J. (2017). Diameter at breast height-crown width prediction models for Parkia biglobosa (Jacq) R. Br. ex G. Don. *Current Life Sciences*, *3*(2), 15–21. https://doi.org/10.5281/zenodo.556420

Adams, N. M., Hand, D. J., & Till, R. J. (2001). Mining for classes and patterns in behavioural data. *Journal of the Operational Research Society*, *52*(9), 1017–1024. https://doi.org/10.1057/palgrave.jors.2601202

Adger, W. N., Arnell, N. W., & Tompkins, E. L. (2005). Successful adaptation to climate change across scales. *Global Environmental Change*, *15*(2), 77–86. https://doi.org/10.1016/j.gloenvcha.2004.12.005

ALLOUCHE, O., TSOAR, A., & KADMON, R. (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, *43*(6), 1223–1232. https://doi.org/10.1111/j.1365-2664.2006.01214.x

Anderson, R. P., Araújo, M., Guisan, A., Lobo, J. M., Mart\'\inez-Meyer, E., Peterson, A. T., & Soberón, J. (2016). Final report of the task group on GBIF data fitness for use in distribution modelling. *Global Biodiversity Information Facility*, *March*, 1–27.

Bag, S., Kumar, S. K., & Tiwari, M. K. (2019). An efficient recommendation generation using relevant Jaccard similarity. *Information Sciences*, *483*(May), 53–64. https://doi.org/10.1016/j.ins.2019.01.023

Baur, B. (1992). Population structure , density , dispersal and neighbourhood size in Arianta arbustorum (Linnaeus , 1758) (Pulmonata : Helicidae) Author (s): Bruno Baur Source : Annalen des Naturhistorischen Museums in Wien . Serie B für Botanik und Published by : *Naturhistorisches Museum*, 307–321. https://www.jstor.org/stable/41766894

Bayor, H. (2012). *Diospyros in West Africa: morphology, molecules and climate*. PhD Thesis. University of Reading.

Bean, W. T., Stafford, R., & Brashares, J. S. (2012). The effects of small sample size and sample bias on threshold selection and accuracy assessment of species distribution models. *Ecography*, *35*(3), 250–258. https://doi.org/10.1111/j.1600-0587.2011.06545.x

Bjurström, A., & Polk, M. (2011). Physical and economic bias in climate change research: A scientometric study of IPCC Third Assessment Report. *Climatic Change*, *108*(1), 1–22. https://doi.org/10.1007/s10584-011-0018-8

Br, J. R., & Don, G. (2021). *Parkia biglobosa African Locust Bean PFAF Plant Database*. 1–8.

Browne, N. A. K. (2011). *Model Evaluation for Seasonal Forecasting over Southern Africa*. 1–118.

C. A. Emmanuel Ikpeme, A. O. O. and C. o. A. (2002). Functional properties of Soyabean and Parkia.pdf. *Global Journal of Pure and Applied Science*.

Connor, T., Hull, V., Viña, A., Shortridge, A., Tang, Y., Zhang, J., Wang, F., & Liu, J. (2018). Effects of grain size and niche breadth on species distribution modeling. *Ecography*, *41*(8), 1270–1282. https://doi.org/10.1111/ecog.03416

Dotchamou, F. T., Atindogbe, G., Sode, A. I., & Fonton, H. N. (2016). Density and spatial pattern of Parkia biglobosa under climate change: The case of Benin. *Journal of Agriculture and Environment for International Development*, *110*(1), 173–194. https://doi.org/10.12895/jaeid.20161.447

Drew, C. A., Wiersma, Y. F., & Huettmann, F. (Eds.). (2011). *Predictive Species and Habitat Modeling in Landscape Ecology*. Springer New York. https://doi.org/10.1007/978-1-4419-7390-0

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., … E. Zimmermann, N. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, *29*(2), 129–151. https://doi.org/10.1111/j.2006.0906-7590.04596.x

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, *27*(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010

Fears, D. (2019). One million species face extinction, U.N. report says. And humans will suffer as a result. *Washington Post*. https://www.washingtonpost.com/climate-environment/2019/05/06/one-million-species-face-extinction-un-panel-says-humans-will-suffer-result/?utm_term=.7451a4067bb1

Feeley, K. J., & Silman, M. R. (2011). Keep collecting: Accurate species distribution modelling requires more collections than previously thought. *Diversity and Distributions*, *17*(6), 1132–1140. https://doi.org/10.1111/j.1472-4642.2011.00813.x

Fifamè Grâce Nadège, V., Abiodoun Olounladé, P., Adam Adenilé, D., Goué Alowanou, G., Virgile Bertrand Azando, E., Pascal Abiodoun Olounladé Assistant Professor, C., Fifamè Grâce Nadège Dedehou, V., & Sylvie Hounzangbé-Adoté, M. (2016). A review on medicinal plants of Parkia Biglobosa (Mimosaceae -Fabaceae) and Pterocarpus Erinaceus (Leguminosae – Papilionoidea). *Journal of Medicinal Plants Studies JMPS*, *132*(46), 132–137. https://pdfs.semanticscholar.org/ea70/a46466db87b96bdfd2b177bfb2a768b189a9.pdf

Franklin, J. (2010). Mapping species distributions. In *Mapping Species*

Distributions: Spatial Inference and Prediction. Cambridge University Press. https://doi.org/10.1017/CBO9780511810602

Franklin, J., Serra-Diaz, J. M., Syphard, A. D., & Regan, H. M. (2016). Global change and terrestrial plant community dynamics. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(14), 3725–3734. https://doi.org/10.1073/pnas.1519911113

Franklin, J., Serra-Diaz, J. M., Syphard, A. D., & Regan, H. M. (2017). Big data for forecasting the impacts of global change on plant communities. *Global Ecology and Biogeography*, *26*(1), 6–17. https://doi.org/10.1111/geb.12501

Gaston, K. J., Blackburn, T. M., & Goldewijk, K. K. (2003). Habitat conversion and global avian biodiversity loss. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*(1521), 1293–1300. https://doi.org/10.1098/rspb.2002.2303

Guisan, A., Zimmermann, N. E., Elith, J., Graham, C. H., Phillips, S., & Peterson, A. T. (2007). What matters for predicting the occurrences of trees: Techniques, data, or species' characteristics? *Ecological Monographs*, *77*(4), 615–630. https://doi.org/10.1890/06-1060.1

Guisan, Antoine, Edward, T. C., Jr, & Hastie, T. (2016). Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological Modelling*, *157*(157), 89–100. www.elsevier.com/locate/ecolmodel

Guisan, Antoine, Graham, C. H., Elith, J., Huettmann, F., Dudik, M., Ferrier, S., Hijmans, R., Lehmann, A., Li, J., Lohmann, L. G., Loiselle, B., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J. M. C., Peterson, A. T., Phillips, S. J., Richardson, K., … Zimmermann, N. E. (2007). Sensitivity of predictive species distribution models to change in grain size. *Diversity and Distributions*, *13*(3), 332–340. https://doi.org/10.1111/j.1472-4642.2007.00342.x

Guisan, Antoine, & Thuiller, W. (2005). Predicting species distribution: Offering more than simple habitat models. *Ecology Letters*, *8*(9), 993–1009. https://doi.org/10.1111/j.1461-0248.2005.00792.x

Hall, W. B., McKeon, G. M., Carter, J. O., Day', K. A., Howden, S. M., Scanlan, J. C., Johnston, P. W., & Burrows, W. H. (1998). Climate change in Queensland's grazing lands: II. An assessment of the impact on animal production from native pastures. *Rangeland Journal*, *20*(2), 177–205. https://doi.org/10.1071/RJ9980177

Hernandez, P. A., Graham, C. H., Master, L. L., & Albert, D. L. (2006). The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*, *29*(5), 773–785. https://doi.org/10.1111/j.0906-7590.2006.04700.x

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*(15), 1965–1978. https://doi.org/10.1002/joc.1276

Hong, T. D., Linington, S., & Ellis, R. H. (1996). Seed Storage Behaviour: a Compendium. Handbooks for Genebanks No. 4. In *Ecology and Classification of North American Freshwater Invertebrates* (Vol. 26, Issue 1). http://linkinghub.elsevier.com/retrieve/pii/B9780123748553000297%5Cnhttp://linkinghub.elsevier.com/retrieve/pii/B9780123748553000285%5Cnhttp://dx.doi.org/10.1016/B978-0-12-374855-3.00024-8%5Cnhttp://dx.doi.org/10.1016/B978-0-12-374855-3.00001-7%5Cnhttp://

Houghton, R. A., Lawrence, K. T., Hackler, J. L., & Brown, S. (2001). The spatial distribution of forest biomass in the Brazilian Amazon: A comparison of estimates. *Global Change Biology*, *7*(7), 731–746. https://doi.org/10.1046/j.1365-2486.2001.00426.x

Houndonougbo, J. S. H. H., Kassa, B., Mensah, S., Salako, V. K., Glèlè Kakaï, R.,

& Assogbadjo, A. E. (2020). A global systematic review on conservation nd domestication of Parkia biglobosa (Jacq.) R. Br. ex G. Don, an indigenous fruit tree species in Sub-Sahara African traditional parklands: current knowledge and future directions. *Genetic Resources and Crop Evolution*, *67*(4), 1051–1066. https://doi.org/10.1007/s10722-020-00892-w

Hulme, P. E. (2005). Adapting to climate change: Is there scope for ecological management in the face of a global threat? *Journal of Applied Ecology*, *42*(5), 784–794. https://doi.org/10.1111/j.1365-2664.2005.01082.x

Hunter, A. F., & Lechowicz, M. J. (2008). Predicting the Timing of Budburst in Temperate Trees Author (s): Alison F . Hunter and Martin J . Lechowicz Published by : British Ecological Society Stable URL : http://www.jstor.org/stable/2404467. *Society*, *29*(3), 597–604.

IPCC. (2008). Glossary of Terms for Working Group II. *2007*, v–6. http://www.ipcc.ch/pdf/glossary/tar-ipcc-terms-en.pdf

Jensen, R. A., Madsen, J., O'connell, M., Wisz, M. S., Tømmervik, H., & Mehlum, F. (2008). Prediction of the distribution of Arctic-nesting pink-footed geese under a warmer climate scenario. *Global Change Biology*, *14*(1), 1–10. https://doi.org/10.1111/j.1365-2486.2007.01461.x

Jiménez-Alfaro, B., Draper, D., & Nogués-Bravo, D. (2012). Modeling the potential area of occupancy at fine resolution may reduce uncertainty in species range estimates. *Biological Conservation*, *147*(1), 190–196. https://doi.org/10.1016/j.biocon.2011.12.030

Jiménez-Valverde, A., Lobo, J., & Hortal, J. (2009). The effect of prevalence and its interaction with sample size on the reliability of species distribution models. *Community Ecology*, *10*(2), 196–205. https://doi.org/10.1556/ComEc.10.2009.2.9

Jiménez-Valverde, Alberto. (2012). Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. *Global Ecology and Biogeography*, *21*(4), 498–507. https://doi.org/10.1111/j.1466-8238.2011.00683.x

Jiménez-Valverde, Alberto, Lobo, J. M., & Hortal, J. (2008). Not as good as they seem: The importance of concepts in species distribution modelling. *Diversity and Distributions*, *14*(6), 885–890. https://doi.org/10.1111/j.1472-4642.2008.00496.x

Kadmon, R., Farber, O., & Danin, A. (2016). A Systematic Analysis of Factors Affecting the Performance of Climatic Envelope Models Author (s): Ronen Kadmon , Oren Farber and Avinoam Danin Published by : Ecological Society of America Stable URL : http://www.jstor.org/stable/4134701 REFERENCES Link. *Ecological Applications*, *13*(3), 853–867.

Karambiri, H., García Galiano, S. G., Giraldo, J. D., Yacouba, H., Ibrahim, B., Barbier, B., & Polcher, J. (2011). Assessing the impact of climate variability and climate change on runoff in West Africa: The case of Senegal and Nakambe River basins. *Atmospheric Science Letters*, *12*(1), 109–115. https://doi.org/10.1002/asl.317

Kelly, R., Leach, K., Cameron, A., Maggs, C. A., & Reid, N. (2014). Combining global climate and regional landscape models to improve prediction of invasion risk. *Diversity and Distributions*, *20*(8), 884–894. https://doi.org/10.1111/ddi.12194

Klutse, N. A. B., Ajayi, V. O., Gbobaniyi, E. O., Egbebiyi, T. S., Kouadio, K., Nkrumah, F., Quagraine, K. A., Olusegun, C., Diasso, U., Abiodun, B. J., Lawal, K., Nikulin, G., Lennard, C., & Dosio, A. (2018). Potential impact of 1.5 °c and 2 °c global warming on consecutive dry and wet days over West Africa. *Environmental Research Letters*, *13*(5), 0–6. https://doi.org/10.1088/1748-9326/aab37b

Kotir, J. H. (2011). Climate change and variability in Sub-Saharan Africa: a review of current and future trends and impacts on agriculture and food security. *Environment, Development and Sustainability*, *13*(3), 587–605. https://doi.org/10.1007/s10668-010-9278-0

Kurukulasuriya, P., Mendelsohn, R., Hassan, R., Benhin, J., Deressa, T., Diop, M., Eid, H. M., Fosu, K. Y., Gbetibouo, G., Jain, S., Mahamadou, A., Mano, R., Kabubo-Mariara, J., El-Marsafawy, S., Molua, E., Ouda, S., Ouedraogo, M., Séne, I., Maddison, D., … Dinar, A. (2006). Will African agriculture survive climate change? *World Bank Economic Review*, *20*(3), 367–388. https://doi.org/10.1093/wber/lhl004

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159. https://doi.org/10.2307/2529310

Lee, S., Choi, J., & Woo, I. (2004). The effect of spatial resolution on the accuracy of landslide susceptibility mapping: A case study in Boun, Korea. *Geosciences Journal*, *8*(1), 51–60. https://doi.org/10.1007/BF02910278

Leroy, B. (2014). *The virtualspecies R package : a complete tutorial*. *September*, 1–77. https://doi.org/10.13140/2.1.1351.8081

Leroy, B., Delsol, R., Hugueny, B., Meynard, C. N., Barhoumi, C., Barbet-Massin, M., & Bellard, C. (2017). Theoretical solutions for the evaluation of discrimination capacity of species distribution models. *The Reprint Server for Biology*. https://doi.org/10.1101/235770

Leroy, B., Delsol, R., Hugueny, B., Meynard, C. N., Barhoumi, C., Barbet-Massin, M., & Bellard, C. (2018). Without quality presence–absence data, discrimination metrics such as TSS can be misleading measures of model performance. *Journal of Biogeography*, *45*(9), 1994–2002. https://doi.org/10.1111/jbi.13402

Leroy, M. B. (2019). *Package ‘ virtualspecies .’*

Liang, C., Michalk, D. L., & Millar, G. D. (2002). The ecology and growth patterns of Cleistogenes species in degraded grasslands of eastern Inner Mongolia, China. *Journal of Applied Ecology*, *39*(4), 584–594. https://doi.org/10.1046/j.1365-2664.2002.00735.x

Liu, C., White, M., & Newell, G. (2020). Measuring the accuracy of species distribution models: A review. *18th World IMACS Congress and MODSIM 2009 - International Congress on Modelling and Simulation: Interfacing Modelling and Simulation with Mathematical and Computational Sciences, Proceedings*, *July*, 4241–4247.

Liu, Canran, White, M., & Newell, G. (2011). Measuring and comparing the accuracy of species distribution models with presence-absence data. *Ecography*, *34*(2), 232–243. https://doi.org/10.1111/j.1600-0587.2010.06354.x

Liu, X., Zhang, Z., Peterson, J., & Chandra, S. (2007). The effect of LiDAR data density on DEM accuracy. *MODSIM 2007 - International Congress on Modelling and Simulation - Land, Water and Environmental Management: Integrated Systems for Sustainability, Proceedings*, 1363–1369.

Liu, Xiaoping, Liang, X., Li, X., Xu, X., Ou, J., Chen, Y., Li, S., Wang, S., & Pei, F. (2017). A future land use simulation model (FLUS) for simulating multiple land use scenarios by coupling human and natural effects. *Landscape and Urban Planning*, *168*(September), 94–116. https://doi.org/10.1016/j.landurbplan.2017.09.019

Lobo, J. M., Jiménez-Valverde, A., & Hortal, J. (2010). The uncertain nature of absences and their importance in species distribution modelling. *Ecography*, *33*(1), 103–114. https://doi.org/10.1111/j.1600-0587.2009.06039.x

Lobo, J. M., Jiménez-Valverde, A., & Real, R. (2008). AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, *17*(2), 145–151. https://doi.org/10.1111/j.1466-

8238.2007.00358.x

Loiselle, B. A., Jørgensen, P. M., Consiglio, T., Jiménez, I., Blake, J. G., Lohmann, L. G., & Montiel, O. M. (2008). Predicting species distributions from herbarium collections: Does climate bias in collection sampling influence model outcomes? *Journal of Biogeography*, *35*(1), 105–116. https://doi.org/10.1111/j.1365-2699.2007.01779.x

Lompo, D., Vinceti, B., Gaisberger, H., Konrad, H., Duminil, J., Ouedraogo, M., Sina, S., & Geburek, T. (2017). Genetic conservation in Parkia biglobosa (Fabaceae: Mimosoideae)-what do we know? *Silvae Genetica*, *66*(1), 1–8. https://doi.org/10.1515/sg-2017-0001

M. McPherson, J., & Jetz, W. (2007). Effects of species' ecology on the accuracy of distribution models. *Ecography*, *30*(1), 135–151. https://doi.org/10.1111/j.2006.0906-7590.04823.x

Magarey, R., Newton, L., Hong, S. C., Takeuchi, Y., Christie, D., Jarnevich, C. S., Kohl, L., Damus, M., Higgins, S. I., Millar, L., Castro, K., West, A., Hastings, J., Cook, G., Kartesz, J., & Koop, A. L. (2018). Comparison of four modeling tools for the prediction of potential distribution for non-indigenous weeds in the United States. *Biological Invasions*, *20*(3), 679–694. https://doi.org/10.1007/s10530-017-1567-1

Makanjuola, V., Omotoso, O., Fadairo, O., Dare, B., Oluwayinka, O., & Adelakun, S. (2016). The Effect of Parkia Leaf Extract on Cadmium-Induced Cerebral Leison in Wistar Rats. *British Journal of Medicine and Medical Research*, *12*(4), 1–7. https://doi.org/10.9734/bjmmr/2016/16017

Manzoor, S. A., Griffiths, G., & Lukac, M. (2018). Species distribution model transferability and model grain size-finer may not always be better. *Scientific Reports*, *8*(1), 1–9. https://doi.org/10.1038/s41598-018-25437-1

Mateo, R. G., Felicísimo, Á. M., & Muñoz, J. (2010). Effects of the number of presences on reliability and stability of MARS species distribution models: The importance of regional niche variation and ecological heterogeneity. *Journal of Vegetation Science*, *21*(5), 908–922. https://doi.org/10.1111/j.1654-1103.2010.01198.x

McPherson, J. M., Jetz, W., & Rogers, D. J. (2004). The effects of species' range sizes on the accuracy of distribution models: Ecological phenomenon or statistical artefact? *Journal of Applied Ecology*, *41*(5), 811–823. https://doi.org/10.1111/j.0021-8901.2004.00943.x

Mensah, C., Amekudzi, L. K., Klutse, N. A. B., Aryee, J. N. A., & Asare, K. (2016). Comparison of Rainy Season Onset, Cessation and Duration for Ghana from RegCM4 and GMet Datasets. *Atmospheric and Climate Sciences*, *06*(02), 300–309. https://doi.org/10.4236/acs.2016.62025

Mertz, O., Halsnæs, K., Olesen, J. E., & Rasmussen, K. (2009). Adaptation to climate change in developing countries. *Environmental Management*, *43*(5), 743–752. https://doi.org/10.1007/s00267-008-9259-3

Miller, J., Franklin, J., & Aspinall, R. (2007). Incorporating spatial dependence in predictive vegetation models. *Ecological Modelling*, *202*(3–4), 225–242. https://doi.org/10.1016/j.ecolmodel.2006.12.012

Millogo, Y., Morel, J., Aubert, J., Millogo, Y., Morel, J., Aubert, J., & Experimental, K. G. (2019). *Experimental analysis of Pressed Adobe Blocks reinforced with Hibiscus cannabinus fibers To cite this version : HAL Id : hal-01850750 Experimental analysis of Pressed Adobe Blocks reinforced with Hibiscus cannabinus fibers*.

New, M., Lister, D., Hulme, M., & Makin, I. (2002). A high-resolution data set of surface climate over global land areas. *Climate Research*, *21*(1), 1–25. https://doi.org/10.3354/cr021001

Odoemene, A. (2017). Climate change and land grabbing. In P. R. Shukla, J. Skea,

E. Buendia Calvo, V. Masson-Delmotte, H.-O. Pörtner, D. C. Roberts, P. Zhai, R. Slade, S. Connors, R. van Diemen, M. Ferrat, E. Haughey, S. Luz, S. Neogi, M. Pathak, J. Petzold, J. P. Pereira, P. Vyas, E. Huntley, … J. Malley (Eds.), *Research Handbook on Climate Change and Agricultural Law* (pp. 423–448). Edward Elgar Publishing. https://doi.org/10.4337/9781784710644.00020

Ojewumi, M. E., Omoleye, J., & Ajayi, A. A. (2016). Optimum Fermentation Temperature for the Protein Yield of Parkia biglobosa Seeds (Iyere). *3rd International Conference on African Development Issues*, *May*, 584–587.

Olajuyigbe S. O., Jimoh S. O., Adegeye A. O., M. R. B. (2013). Drought stress on early growth of Diospyros mespiliformis Hochst ex A. Rich in Jega, Northern Nigeria. *Nigerian Journal of Ecology*, *12*(2012), 71–76.

Oyerinde, O. V., Olusola, J. A., & Adebo, A. A. (2018). Variation in morphometric traits of trees, pods and seeds of Parkia biglobosa (Jacq) G. in Southwestern, Nigeria. *International Journal of Conservation Science*, *9*(1), 185–192.

Pearson, R. G., & Dawson, T. P. (2003). Predicting the impacts of climate change on the distribution of species: Are bioclimate envelope models useful? *Global Ecology and Biogeography*, *12*(5), 361–371. https://doi.org/10.1046/j.1466-822X.2003.00042.x

Pearson, R. G., Thuiller, W., Araújo, M. B., Martinez-Meyer, E., Brotons, L., McClean, C., Miles, L., Segurado, P., Dawson, T. P., & Lees, D. C. (2006). Model-based uncertainty in species range prediction. *Journal of Biogeography*, *33*(10), 1704–1711. https://doi.org/10.1111/j.1365-2699.2006.01460.x

Pena, J. C. de C., Kamino, L. H. Y., Rodrigues, M., Mariano-Neto, E., & de Siqueira, M. F. (2014). Assessing the conservation status of species with limited available data and disjunct distribution. *Biological Conservation*, *170*, 130–136. https://doi.org/10.1016/j.biocon.2013.12.015

Peterson, A. T., Papeş, M., & Soberón, J. (2008). Rethinking receiver operating characteristic analysis applications in ecological niche modeling. *Ecological Modelling*, *213*(1), 63–72. https://doi.org/10.1016/j.ecolmodel.2007.11.008

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, *190*(3–4), 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Phillips, S. J., Dudík, M., Elith, J., Graham, C. H., Lehmann, A., Leathwick, J., & Ferrier, S. (2009). Sample selection bias and presence-only distribution models: Implications for background and pseudo-absence data. *Ecological Applications*, *19*(1), 181–197. https://doi.org/10.1890/07-2153.1

Polak, A., Kucio, C., Kloth, L. C., Paczula, M., Hordynska, E., Ickowicz, T., Blaszczak, E., Kucio, E., Oleszczyk, K., Ficek, K., & Franek, A. (2016). Adaptation to Climate Change and Variability in Rural West Africa. In J. A. Yaro & J. Hesselberg (Eds.), *Ostomy/wound management* (Vol. 64, Issue 2). Springer International Publishing. https://doi.org/10.1007/978-3-319-31499-0

Ræbild, A., Larsen, A. S., Jensen, J. S., Ouedraogo, M., de Groote, S., van Damme, P., Bayala, J., Diallo, B. O., Sanou, H., Kalinganire, A., & Kjaer, E. D. (2011). Advances in domestication of indigenous fruit trees in the West African Sahel. *New Forests*, *41*(3), 297–315. https://doi.org/10.1007/s11056-010-9237-5

Raes, N., & Ter Steege, H. (2007). A null-model for significance testing of presence-only species distribution models. In *Ecography* (Vol. 30, Issue 5, pp. 727–736). https://doi.org/10.1111/j.2007.0906-7590.05041.x

Raes, N., Ter Steege, H., Steege, H. ter, Ter Steege, H., & Steege, H. ter. (2007). A null-model for significance testing of presence-only species distribution models. *Ecography*, *30*(5), 727–736. https://doi.org/10.1111/j.2007.0906-

7590.05041.x

Roudier, P., Ducharne, A., & Feyen, L. (2014). Climate change impacts on runoff in West Africa: A review. *Hydrology and Earth System Sciences*, *18*(7), 2789–2801. https://doi.org/10.5194/hess-18-2789-2014

Segurado, P., & Araújo, M. B. (2004). An evaluation of methods for modelling species distributions. *Journal of Biogeography*, *31*(10), 1555–1568. https://doi.org/10.1111/j.1365-2699.2004.01076.x

Shabani, F., Kumar, L., & Ahmadi, M. (2016). A comparison of absolute performance of different correlative and mechanistic species distribution models in an independent area. *Ecology and Evolution*, *6*(16), 5973–5986. https://doi.org/10.1002/ece3.2332

Shabani, F., Kumar, L., & Ahmadi, M. (2018). Assessing accuracy methods of species distribution models: AUC, Specificity, Sensitivity and the True Skill Statistic. *Global Journal of Human Social Science*, *18*(1), 6–18. https://socialscienceresearch.org/index.php/GJHSS/article/download/2469/2358

Somodi, I., Lepesi, N., & Botta-Dukát, Z. (2017). Prevalence dependence in model goodness measures with special emphasis on true skill statistics. *Ecology and Evolution*, *7*(3), 863–872. https://doi.org/10.1002/ece3.2654

Soultan, A., & Safi, K. (2017). The interplay of various sources of noise on reliability of species distribution models hinges on ecological specialisation. *PLoS ONE*, *12*(11), 1–19. https://doi.org/10.1371/journal.pone.0187906

Stocker, T. F., Clarke, G. K. C., Le Treut, H., Lindzen, R. S., Meleshko, V. P., Mugara, R. K., Palmer, T. N., Pierrehumbert, R. T., Sellers, P. J., Trenberth, K. E., & Willebrand, J. (2001). Physical Climate Processes and Feedbacks. *Climate Change 2001: The Scientific Bases. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate*

Change, 881.

Sylla, M. B., Faye, A., Klutse, N. A. B., & Dimobe, K. (2018). Projected increased risk of water deficit over major West African river basins under future climates. *Climatic Change*, *151*(2), 247–258. https://doi.org/10.1007/s10584-018-2308-x

Thornton, P. K., Jones, P. G., Owiyo, T., Kruska, R. L., Herrero, M., Orindi, V., Bhadwal, S., Kristjanson, P., Notenbaert, A., Bekele, N., & Omolo, A. (2008). Climate Change and Poverty in Africa: Mapping Hotspots of Vulnerability. *African Journal of Agricultural and Resource Economics*, *2*(1), 24–44.

van Proosdij, A. S. J., Sosef, M. S. M., Wieringa, J. J., & Raes, N. (2016). Minimum required number of specimen records to develop accurate species distribution models. *Ecography*, *39*(6), 542–552. https://doi.org/10.1111/ecog.01509

van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G. C., Kram, T., Krey, V., Lamarque, J.-F. F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S. J., & Rose, S. K. (2011). The representative concentration pathways: an overview. *Climatic Change*, *109*(1), 5–31. https://doi.org/10.1007/s10584-011-0148-z

Veloz, S. D. (2009). Spatially autocorrelated sampling falsely inflates measures of accuracy for presence-only niche models. *Journal of Biogeography*, *36*(12), 2290–2299. https://doi.org/10.1111/j.1365-2699.2009.02174.x

Vizy, E. K., & Cook, K. H. (2012). Mid-twenty-first-century changes in extreme events over northern and tropical Africa. *Journal of Climate*, *25*(17), 5748–5767. https://doi.org/10.1175/JCLI-D-11-00693.1

Wisz, M. S., Hijmans, R. J., Li, J., Peterson, A. T., Graham, C. H., Guisan, A., Elith, J., Dudík, M., Ferrier, S., Huettmann, F., Leathwick, J. R., Lehmann, A., Lohmann, L., Loiselle, B. A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J. M. C., … Zimmermann, N. E. (2008). Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, *14*(5), 763–773. https://doi.org/10.1111/j.1472-4642.2008.00482.x

Yates, C. J., Elith, J., Latimer, A. M., Le Maitre, D., Midgley, G. F., Schurr, F. M., & West, A. G. (2010). Projecting climate change impacts on species distributions in megadiverse South African Cape and Southwest Australian Floristic Regions: Opportunities and challenges. *Austral Ecology*, *35*(4), 374–391. https://doi.org/10.1111/j.1442-9993.2009.02044.x

Yesson, C., Brewer, P. W., Sutton, T., Caithness, N., Pahwa, J. S., Burgess, M., Gray, W. A., White, R. J., Jones, A. C., Bisby, F. A., & Culham, A. (2007). How global is the global biodiversity information facility? *PLoS ONE*, *2*(11). https://doi.org/10.1371/journal.pone.0001124

Yira, Y., Diekkrüger, B., Steup, G., & Yaovi Bossa, A. (2017). Impact of climate change on hydrological conditions in a tropical West African catchment using an ensemble of climate simulations. *Hydrology and Earth System Sciences*, *21*(4), 2143–2161. https://doi.org/10.5194/hess-21-2143-2017

Zurell, D., Thuiller, W., Pagel, J., Cabral, J. S., Münkemüller, T., Gravel, D., Dullinger, S., Normand, S., Schiffers, K. H., Moore, K. A., & Zimmermann, N. E. (2016). Benchmarking novel approaches for modelling species range dynamics. *Global Change Biology*, *22*(8), 2651–2664. https://doi.org/10.1111/gcb.13251

APPENDICES

Appendix 1: Parkia biglobosa plant (picture taken in Nyankpala/Tamale – Ghana)

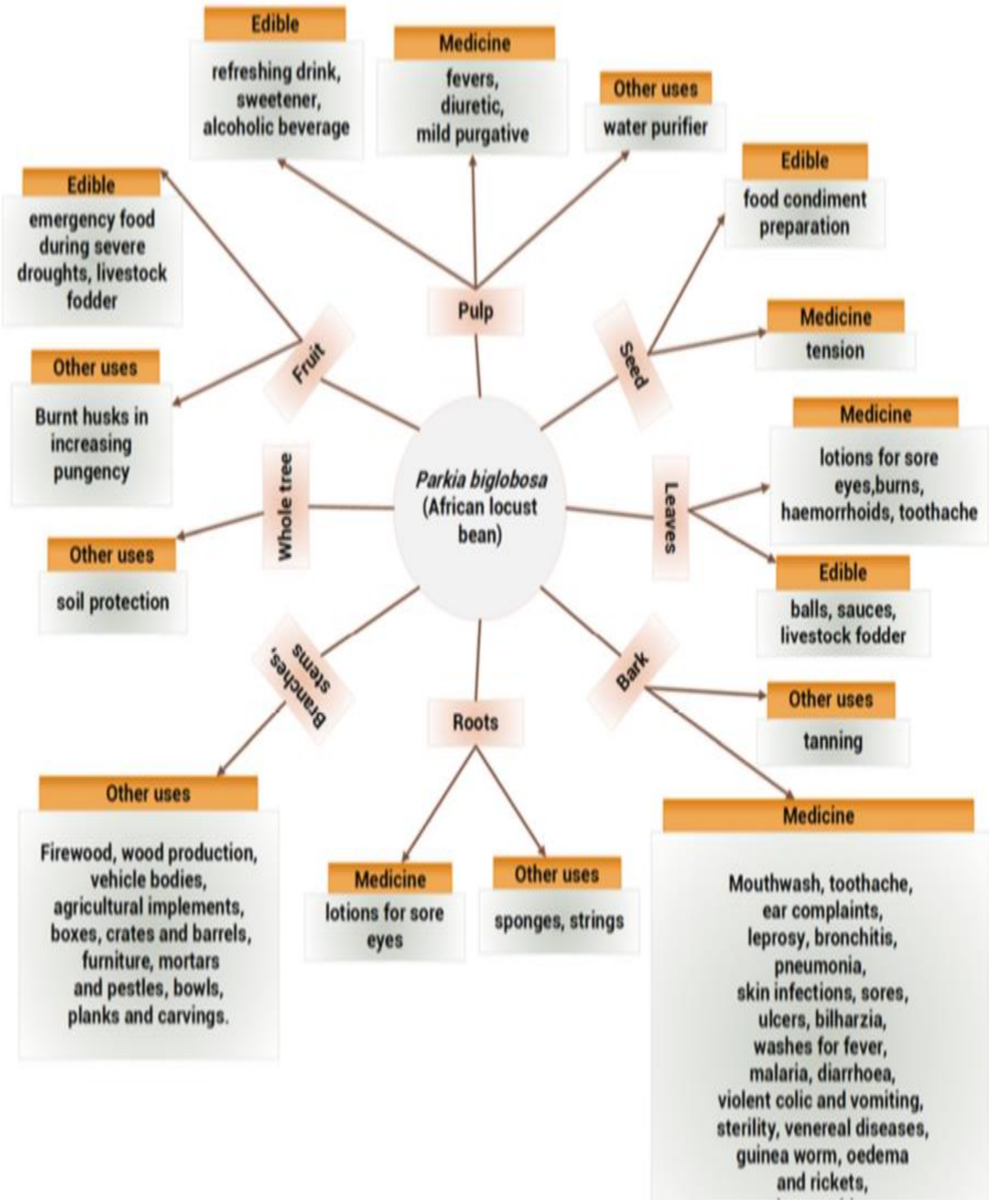
A: Fresh Parkia biglobosa fruits on the tree

B: Dried Parkia biglobosa fruits

Appendix 2: Picture of the locust bean, both unripe and dried (in Tamale - Ghana)

Appendix 3: Africa Locust bean plant, parts and uses

The Flower (pollinated by bees in their quest for nectar for the production of honey), Bean (brown covering when ripped, yellow pulp and sour, indication of vital nutrients and vitamins like vitamin A and ascorbic acid) Pulp (used as a feed supplement in pig production) and “Dawadawa” – a protein rich condiment made from fermented seeds of the bean, and the back being chopped for medicinal purpose. (Pictures taken in Ghana and derived from Alamy Stock photos on the internet)



Appendix 4: Summary of some uses of *Parkia biglobosa* (African Locust bean).

Source: Resources et al., (2020)

Appendix 5: Countries with the greatest multi-hazard population exposure, with country rankings for cyclones, drought, and floods. Source: Christenson et al. (2014)

Country	Cyclone rank	Drought rank	Flood rank	Multi-hazard exposure rank
Hong Kong	5	139	3	1
Philippines	11	74	22	2
Macao	10	132	1	3
Guatemala	63	10	5	3
South Korea	22	118	15	5
Bangladesh	53	29	2	4
Veitnam	36	80	12	7
Saint Kitts and Nevis	20	6	181	8
Guadeloupe	17	65	83	9
Guam	1	68	132	10
Lebanon	93	2	42	11
Ecuador	93	27	17	12
Nepal	93	44	6	13
Japan	7	182	64	14
British Virgin Island	8	45	181	15
Thailand	73	35	20	16
Puerto Rico	14	193	48	17
Antigua and Barbuda	9	70	134	18
New Caledonia	6	66	166	19
Mozambique	40	31	73	20