

CLIMATE MODELLING OF THE GEOGRAPHIC RANGE OF FASCIOLA GIGANTICA PREVALENCE IN SEMI-ARID WEST AFRICA: A CASE STUDY OF SOKOTO STATE, NIGERIA

Isah Hamisu¹, Heiko Balzter^{2,3} & Jörg Kaduk²

¹Department of Geography, Usmanu Danfodio University, Sokoto, Sokoto State, Nigeria

²Centre for Landscape and Climate Research, University of Leicester, Leicester, LE1 7RH, UK

³National Centre for Earth Observation, University of Leicester, Leicester, LE1 7RH, UK

Corresponding Author's email: hamisu.isah@udusok.edu.ng



Abstract

Fasciola gigantica is referred to as a neglected tropical animal and zoonotic disease, which therefore poses severe threats to developing countries like Nigeria where funding on public and veterinary services are grossly inadequate. Data on 210 recorded occurrence locations of *F. gigantica* were obtained from Sokoto State Ministry of Animal Health, Sokoto, Nigeria and were used as training sites. The predictions from these training sites were validated using independent data obtained during field work from fifteen presence localities in ten provinces utilising both threshold-dependent and threshold independent measures of performance. The data on *F. gigantica* occurrence combined with long-term environmental data were inputted into three MaxEnt's modelling scenarios to predict the geographic distribution of the parasite in the study area. The analyses show statistically significant performances across all the three MaxEnt's scenarios. The results further show that precipitation of the wettest quarter, annual mean moisture index and maximum annual NDVI were the most important variables that influenced the geographic distribution of the parasite in Sokoto State. This study presents the first application of CliMond database with MaxEnt models in predicting the distribution range of fascioliasis in semi-arid West Africa.

Keywords: species distribution models, MaxEnt, *F. gigantica*, Sokoto State

Introduction

Fascioliasis is 'a food- or water-borne trematodiasis' (Tolan, 2011) that denotes *F. hepatica* and *F. gigantica* diseases. The former species is endemic in temperate areas and the latter in tropical areas of Africa, hence they have been described as 'epidemiologically separate entities' (Spithill et al., 1999). Moreover, *F. gigantica* is genetically related to *F. hepatica* but diverged many millions of years ago and then dispersed into various parts of Asia and Africa (Irving et al., 2003). Thus, these two species share common characteristics regarding liver infections of the mammalian host and share the same life cycle. Despite these common attributes, these two species differ in the amount of moisture needed by their respective intermediate hosts and also in the interactions between the host and the parasite (Spithill et al., 1999). Differences also exist in the practices of animal management and other development indices between temperate areas where *F. hepatica* thrives and tropical biomes where *F. gigantica* is endemic. Also, *R. natalensis* that transmits *F. gigantica* requires higher temperature and stagnant water bodies whereas *Galba truncatula* needs lower temperature and intermittent water bodies (Mas-Coma, 2004). All these variations account for the epidemiological differences between the two species, which warrants different

studies. In Nigeria, only *F. gigantica* thrives, which reached 60% prevalence rate (Spithill et al., 1999).

Fasciola gigantica is referred to as a neglected tropical animal disease by the Centers for Disease Control and Prevention (CDC, 2010) that is found to be zoonotic by World Health Organisation (WHO, 2006). It poses severe threats to developing countries like Nigeria where funding on public and veterinary services are grossly inadequate (Franklin, 2009). The disease is known to impair the productivity of the infected animals, which leads to the loss of millions of dollars worldwide thereby becoming inimical to trade and human wellbeing indirectly (Mas-Coma et al., 2005). The effects of the disease on humans even though not yet reported in the study area has been captured in various parts of the world and is regarded as important.

Species distribution models have been developed recently to estimate the association between species occurrence locations and climatic characteristics of the locations (Franklin, 2009, Phillips et al., 2006). These models are parameterised with presence-only data and then predict the likely spatial distribution of the species across the whole region (de Leeuw et al., 2002, Arntzen, 2006). Species distribution models apply in many fields based on various objectives: conservation (Kremen et al., 2008, Thorn et al., 2009), ecology and biogeography (Martínez-Meyer et al., 2006, Elith and Leathwick, 2009), risk assessment and identifying disease locations (Peterson, 2003, Andersen et al., 2004). Given these broad applications of species distribution models, various techniques have emerged that include presence/absence and presence only methods (Elith et al., 2011). Random forests boosted regression trees; generalised additive models are an example of techniques that can be applied where both presence and absence data are available (Phillips et al., 2006).

In contrast, some methods require only the presence of a species at a given location to model the geographical distribution of the species across the region. That means that no survey has been carried out to establish the absence of the species, but the model relies on existing presence records from government departments, citizen science or museums and herbariums (Elith et al., 2011). In developing countries presence-only data are more available (Tsoar et al., 2007) while there is a paucity of species absence records which are often impractical to collect, and even if 'available' may not be of satisfactory reliability (Ponder et al., 2001, Anderson et al., 2003).

More studies have been carried out on *F. hepatica* than *F. gigantica* due to its greater global spread (Spithill et al., 1999). As a result of this imbalance, the understanding of fascioliasis has become fragmented due to a paucity of studies from regions of the world outside the temperate biomes, particularly from semi-arid West Africa. This study focuses on the prevalence of *F. gigantica* in Sokoto State in Nigeria. Domestic livestock acts as the main hosts of *F. gigantica* (Tum et al., 2004), and the state is ranked as the second most important in Nigeria regarding animal population (Magaji et al., 2014).

This study opted for maximum entropy as a presence-only technique of species distribution modelling. It uses observed sites of a species occurrence and climatic as well as environmental features of those sites. The aim was to predict the geographic range of the species by delimiting suitable and unsuitable sites for the species (Phillips et al., 2004). This method is appropriate as local government records exist in Sokoto State for locations where cases of *F. gigantica* occurred. The parasite is responsive and highly associated with ecological and environmental conditions in their geographical distribution across the entire globe (Mas-Coma et al., 2009). To date, the application of maximum entropy to the study of fascioliasis has only been limited to a few studies (Valentia-Lopez et al., 2012, Kantzoura et al., 2011b, Yaro et al., 2018). In addition, this study is

the first that utilised BioClim data from CliMond database that incorporates soil moisture variables (BIO1-BIO35) in modelling the geographic range of fascioliasis.

The main objectives of this study were 1) to evaluate the performance of MaxEnt using different combination of variables in scenario 1 (BIO1-19), scenario 2 (BIO1-35) and scenario 3 (BIO1-35 with NDVI and SRTM, 2) to investigate climatic variables that are significant for the prevalence of *F. gigantica* in the study area and, 3) to provide novel methods using GIS analysis for the design of appropriate monitoring to control *F. gigantica* infections of domestic animals in Sokoto State, Nigeria.

Materials and Methods

Study area and *F. gigantica* occurrence data

Sokoto State (Figure 1) is located between longitudes 4° 8' E and 6° 54' E and latitudes 12° N and 13° 58' N. The state shares boundaries with Kebbi state, Zamfara state and the Republic of Niger and has 23 provinces that cover a total land area of 32,000 square kilometres (Bala et al., 2014). The highest temperature is recorded in the dry season and exceeds 40°C between March and May and the mean annual temperature is 34°C (Atedhor, 2015). Rainfall starts in April-May and ends in September with an annual range of 500mm to 1300mm (Bala et al., 2014). River Rima is the major river that trajects the state and extends into Kebbi State and other parts of Nigeria. According to Davis (1982), the landscape of Sokoto State is homogeneously plain with an average height of 300m where the vegetation is short grass savannah. Magaji et al. (2014) report that the majority of the inhabitants in the state practice agriculture and animal husbandry, this is because the state is among the leading producers of animals in Nigeria, second only to Borno State (Mamman, 2005).

This study used both secondary and primary data through field survey for model training and evaluation respectively. The former data on 210 reported presence locations of *F. gigantica* were obtained from the Ministry of Animal Health in Sokoto State, established in 1965. The coordinates of the locations were available from the National Population Commission (NPC, 2006). The use of occurrence data from the mentioned source was due to the availability of only a few (3) *F. gigantica* prevalence records from the Global Biodiversity Information Facility (GBIF) in Nigeria. The duplicates of the species record were removed using a Nearest Neighbour Index (Clark and Evans, 1954) calculated with the R ECOSPAT package (Di Cola et al., 2017). This package was also employed in adjusting the spatial resolution of the sample points to a minimum distance of 1 km in order to conform to that of the climatic variables. These data cleaning and spatial resolution modification were applied to increase model performance and promote proper understanding of environmental conditions occupied by the species (Elith et al., 2011, Hijmans et al., 2011). The number of sites reduced to 177, and 33 were thinned out due to clustering.

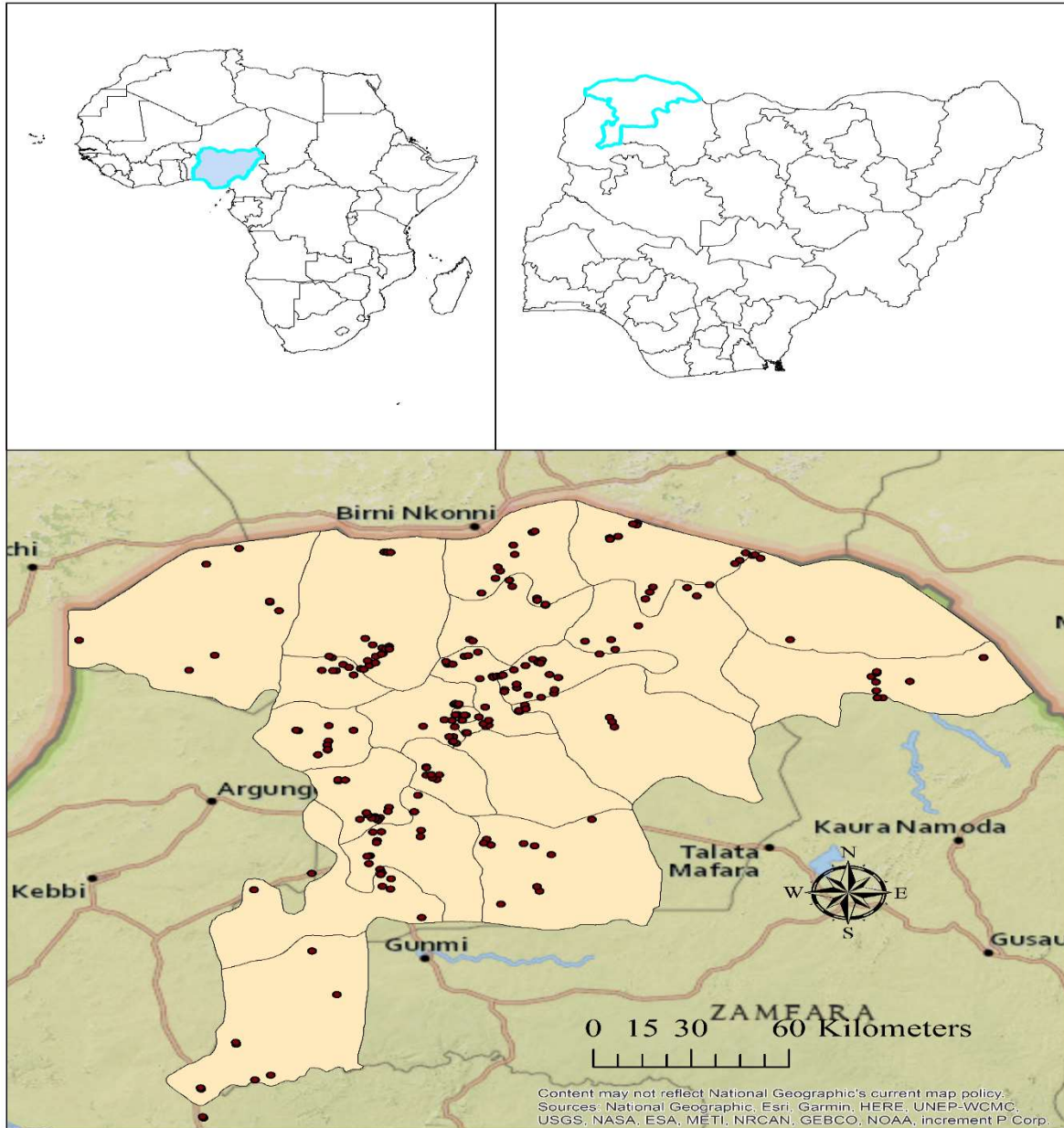


Figure 1: Map of the study area.

Climatic variables

CliMond database is accessible at (<http://www.climond.org/>) for free public use in species distribution modelling, climate studies and niche modelling among others (Hutchinson et al., 2009, Kriticos et al., 2012, Kriticos et al., 2014). It used interpolation in the derivations of climate surfaces at 10' and 30' for the entire land surfaces of the World. It was further highlighted by Kriticos et al. (2012) that this database applied WorldClim and Climate Research Unit (CRU) data sets as a 'baseline climatology' due to their accuracy and broad applications in research.

According to Kriticos et al. (2014), the inadequacy of the core suite of Bioclim variables (BIO1-19) in species distribution modelling of some organisms has led to the development of the soil moisture Bioclim variables (BIO28-35) that are at present only available at CliMond Archive.

The elevation data used in this study was the product of the global 1-arc second Shuttle Radar Topographic Mission (SRTM) via the United States Geological Survey's EarthExplorer website (<http://earthexplorer.usgs.gov>). It was developed from the joint project by NASA, the National Geospatial-Intelligence Agency as well as the German and Italian Space Agencies and hence referred to as most comprehensive (Farr et al., 2007). According to Rodriguez et al. (2006), the Jet Propulsion Laboratory data gathered interferometric radar data for the creation of the elevation data within less than 60° latitudes of the World.

This study also utilised NDVI data from 2005 to 2014 and is available at the National Aeronautic and Space Administration (NASA) website (<http://neo.sci.gsfc.nasa.gov/>). Also, the index was the product of Moderate Resolution Imaging Spectroradiometer (MODIS) described by Wiegand et al. (1991) as capable of capturing the variability in vegetation conditions and the calculation of biophysical parameters from arid areas to dense vegetation of rain forest regions. The NDVI has positive values that show bare soil if it is zero or 0.1 and 0.5 values imply sparse vegetation while greater than 0.6 to 1 show dense vegetation. The negative values always show clouds and water surfaces (Leta et al., 2015).

The sixteen variables chosen for constructing the three MaxEnt's modelling scenarios are annual mean temperature (bio_1), mean diurnal range of temperature (bio_2), isothermality (bio_3), temperature seasonality (bio_4), mean temperature of the coldest month (bio_6), mean temperature of the driest quarter (bio_9), annual precipitation (bio_12), precipitation of the wettest quarter (bio_16), precipitation of the driest quarter (bio_17), precipitation of the warmest quarter (bio_18) and precipitation of the coldest month (bio_19), annual mean moisture index (bio_28), mean moisture index of the coldest quarter (bio_35), maximum annual NDVI (NDVI_36), mean annual NDVI (NDVI_38) and mean elevation (SRTM_39). Highly correlated variables were removed and these aforementioned variables were used. In Sokoto State, according to Magaji et al. (2014) climatic and environmental variables contribute immensely to the prevalence of *F. gigantica*. Previous studies have indicated that temperature and precipitation variables are significant factors in the geographic distribution of fascioliasis (Spithill et al., 1999, Mas-Coma et al., 2009, Fox et al., 2011).

The selected bioclimatic variables were clipped to the extent of Sokoto State using ESRI (Environmental Systems Research Institute) ArcMap 10.3. In addition, the software was used in the modification of the climate layers to the same projection, extent and cell size consistent with the study area prior to modelling operations.

Modelling technique

Maximum entropy is a presence-only machine learning technique that has been developed in 2004 (Elith et al., 2010) that makes use of the known presence of a species at a site and the measured climatic variables across the site in order to model the species geographic dispersion (Phillips et al., 2006). It utilises two data sets that include known sites of species occurrence and the approximated measurement of variables (climatic or environmental) over those sites in order to model the suitability of the habitat (sites) for the species (Franklin, 2009). The main idea is that the dispersion of the species (which is unknown) should have maximum entropy (regarding consistency with the data) (Merow et al., 2013) and as such should be uniform across the area of study under the condition of satisfying some constraints (Phillips et al., 2006). These constraints are that the values of these variables predicted should be equivalent to or near to their observed values which reflects the principle of maximum entropy (Elith et al., 2011).

The relationship between the species and the variables used in fitting the MaxEnt model is complex and not linear (Austin, 2002), which therefore necessitates the variables to be transformed mathematically into features (Elith et al., 2011). These features consist of five classes viz linear, quadratic, product, threshold and hinge (Dudík et al., 2004), and they serve as constraints on the spatial dispersion of species since they are transformations of the original environmental variables (Phillips et al., 2006).

Maximum entropy is a species distribution modelling technique that has been built with a 'guarantee' to produce accurate modelling of species geographical distribution (Phillips et al., 2004, Phillips et al., 2006, Elith et al., 2011). According to Phillips et al. (2006) it uses various steps of mathematical and statistical operations that are 'deterministic' and consequently 'converge to the maximum entropy probability distribution'. In the explanation of how MaxEnt works this study relies strongly on the contributions of (Phillips et al., 2006), (Elith et al., 2011) and (Merow et al., 2013). For illustration, in this study, we use K to represent a collection of pixels that constitute the study site that consist of smaller sub units as occurrence sites q within the unknown distribution D . Each of the study sites q would have a positive probability value in the unknown distribution approximated to 1. The unknown distribution D is estimated as likely distribution $\hat{D}$ whose entropy is interpreted as

$$H(\hat{D}) = - \sum_{q \in K} \hat{D}(q) \ln \hat{D}(q)$$

Where $\ln$ is the natural logarithm, H is the entropy, D is unknown distribution, $\hat{D}$ is the estimate of unknown distribution (Phillips et al., 2006).

The algorithm of maximum entropy is effective (Berger et al., 1996) and is consistent with its principle. This means a homogeneous distribution of the unknown distribution D is expected to be achieved through constraining the mathematical transformation of the environmental features f (Elith et al., 2011) to match the observed values at K the study site. Consequently, each feature in the environmental space allocates a number to every site q in the study area K . The assumption or the averages of each feature f under the unknown distribution D is $\sum_{q \in K} D(q) f_j(q)$ which is simplified as $D[f_j]$.

On the other hand, the empirical or the observed values average is given as $\sum_{i=1}^m f_j(q_i)$ alternately written as $\tilde{D}[f_j]$ ($\tilde{D}$ is representing a homogeneous distribution in the study area). So in line with the principle of maximum entropy, the unknown probability D is constraining the predicted features f_j ($\hat{D}[f_j]$) to have similar average values with the empirical averages $\tilde{D}[f_j]$ (Jaynes, 1957) which is given as

$$\hat{D}[f_j] = \tilde{D}[f_j] \dots \dots \dots \text{for each feature } f_j \dots \dots \text{equation 1}$$

The above proportionality (as shown in equation 1) in the unknown distribution can be performed by application of the mathematical theory of convex duality (Della Pietra et al., 1997). This theory indicates similarity between MaxEnt likelihood distribution and Gibbs distribution (Dudík et al., 2004) and is expressed as

$$q\lambda(D) = \frac{e^{\lambda \cdot f(q)}}{Z_\lambda} \dots \dots \dots \text{equation 2}$$

Equation 2 shows Gibbs distribution with λ indicating a vector of n real-valued coefficient or the transformed environmental variable weights, f indicates the vector of the entire n features while

Z_λ has a unique function of adding probability distributions q_λ to be equal to 1 and hence acts as normalizing constant. It is worth emphasizing that the equality between MaxEnt likelihood distribution D and the Gibbs distribution q_λ as indicated by convex duality ‘maximises’ the probability of all the m (sample) sites (Phillips et al., 2006). Similarly, the negative log probability of all the m (sample) sites is minimised.

$$\tilde{D}[-\ln(q_\lambda)] \dots \dots \dots \text{equation 3}$$

which is alternatively written as $Z_\lambda - \frac{1}{m} \sum_{i=1}^m \lambda_j f(q_i)$ and is referred to as “log loss”

MaxEnt has an inherent tendency to over fit the training data. This is because the predicted values could not match the observed or empirical values precisely but can only estimate them. In that light, as explained by (Phillips et al., 2006) there is a restriction imposed to constrain the average of the expected values under the unknown distribution to be near the observed values. This led to the constraint relaxation in equation (1) with an addendum (Dudik et al., 2004)

$$\tilde{D}[f_j] = \tilde{D}[f_j] \leq \beta_j \dots \dots \dots \text{equation 4}$$

for each feature f_j and constants β_j

Consequently, this leads to e_1 -regularization that altered the two characterization in (4) (Phillips et al., 2006): in (5) MaxEnt’s distribution is indicated to be Gibbs distribution that sets the minimum limits (between observed and empirical values)

$$\tilde{D}[-(q_\lambda)] + \sum_j \beta_j |\lambda_j| \dots \dots \dots \text{equation 5}$$

Where the ‘log loss’ is the first segment and the second segment ‘penalizes’ the utility of higher numbers for the coefficients λ_j (Phillips et al., 2006). As reported by (Williams, 1995) regularization has the advantage of not only reducing over fitting but also in coercing MaxEnt to concentrate on the most significant variables. Furthermore, e_1 -regularization greatly decreases variables to a small numbers thereby reducing the possibility of over fitting and is referred to as “lasso” in Generalised Linear Models (GLM) and Generalised Additive Models [GAM] (Guisan et al., 2002).

The MaxEnt procedure described in the preceding discussion will converge at optimal probability distribution through repetition in changing the coefficients $\lambda = (0 \dots 1)$ in order to minimise regularised log loss (Phillips et al., 2006). The settings used in MaxEnt were 10000 background points and the features include hinge, linear and quadratic.

Model application

All the three MaxEnt modelling scenarios made use of presence-only of *F. gigantea* occurrence record. However, in order to make up for absence data, background samples were drawn randomly by the models to compare the geographic range of presence data with influential climatic variables. Inside each raster grid of the study area, these presence occurrence data and background points combined with environmental data to project the likelihood of *F. gigantea* occurrence in Sokoto State.

Model evaluation

This study used both threshold-dependent and threshold-independent measures in evaluating the accuracy of the three modelling scenarios. Threshold implies setting a limit for suitability above specific values and unsuitability below those values and is described as an accurate

measure of models ability to produce predictive maps for binary (presence and absence) classification (Liu et al., 2005, Negga, 2007). All the threshold-dependent measures of evaluation used by this study required a choice of a threshold which has much impact on binary (presence and absence maps) classification (Freeman and Moisen, 2008b). In that light, this study applied a threshold based on the 10th percentile training presence. This threshold according to Jarnevig and Reynolds (2010) depends on the choice of a value which if exceeded classifies 90% of the training points as suitable for the species. Morueta-Holme et al. (2010) and Jarnevig and Reynolds (2010) used this threshold in their respective studies where they described it as resulting 'to a more conservative' model than other thresholds criteria that may lead to over-estimate of the model's predictive ability. This study selects this threshold as it gives a correct binary classification of species distribution models that included almost all the known occurrence locations for the prevalence of *F. gigantea* in the study area.

True Skill Statistics TSS is another threshold dependent measure of model performance which correlates highly with an area under the curve (AUC) that does not depends on prevalence (Allouche et al., 2006). The early applications of TSS otherwise known as Hansen-Kuiper's discriminant were in meteorological forecasting and applied in the evaluation of weather forecasts (Accadia et al., 2005). The value scores of TSS ranging between -1 to +1, where zero implies model goodness that is equal to random guessing while +1 implies perfection in agreement.

Likewise, species distribution models are evaluated using threshold-independent measures due to their ability to utilise complete information provided by the model in order to explain the general characteristics of the species distribution (Merow et al., 2013, Fielding and Bell, 1997). These measures assess the performance of a model without relying on any particular threshold (Deleo, 1993, Phillips et al., 2006). Receiver operating characteristic (ROC) has its origin in signal processing, and it indicates the ability of the model in ranking cases (discriminating) into two categories or classes using all possible thresholds (Provost and Fawcett, 1997, Elith, 2002, Merow et al., 2013). It uses the area under the curve (AUC), which is a technique that uses value to show the likelihood of any random selection to be in favour of presence localities than background localities in presence-only models (Fielding and Bell, 1997, Merow et al., 2013).

In interpreting the values of AUC, the random prediction has a score of 0.5 which indicates a bad model performance while a good model has values closer to 1 (Freeman and Moisen, 2008a, Swets, 1988). Furthermore, the species distribution models with an AUC score of 0.7 and above are classified as 'potentially significant' and possessing a good predictive ability, while models with AUC equal to or greater than 0.8 are excellent, and models with equal to or greater than 0.9 are outstanding (Hosmer and Lemeshow, 2000, Elith et al., 2006, Morueta-Holme et al., 2010, Sobek-Swant et al., 2012).

AUC provides a robust estimate of model performance that it is not affected by prevalence (Manel et al., 2001) and hence can be a valid means of comparing different models (Cumming, 2000). The only caveats in the use of AUC as it is ranked-based (Merow et al., 2013) is that validity for comparison among models should ensure that all the models constructed should be similar regarding the study area, background samples, species as well as in the utility of validation data set (Elith et al., 2011).

This study used the AUC score of both the training data and the independent validation data in the evaluation of species distribution models. The former assesses the fitness of the model to the data, and the latter appraises the predictive ability of the model (Araújo and Guisan, 2006, Merow et al., 2013).

Statistical analysis

In testing the statistical significance of the relationship between model predictions and observations, this study used the Wilcoxon rank signed test that is equivalent to a paired two-tailed t-test (Phillips et al., 2006).

Finally, a jack-knife test in MaxEnt was used to evaluate the relative contribution of the climatic variables that were selected in modelling the spatial range of *F. gigantea* in the study area. The jack-knife test determines the contribution of each variable by serially accomplishing the following three tasks;

- Involving all variables in running the model
- Leaving out one variable at a time and running the model again
- Running the model using only one variable, until all variables have been individually used

The most significant variable (s) resulting from these Jack-knife operation is the one that increases highest training gains when used alone in running the model and likewise decreases the training gain than all other variables when left out or isolated from the model. (For more explanation refer to MaxEnt tutorial on (www.cs.princeton.edu)).

Model validation

A field survey approved and funded by the University of Leicester was conducted from July to August, 2016 in order to check the reliability of the prediction made by MaxEnt model in terms of suitable habitats. Fifteen sites representing the four agricultural zones in the study area were surveyed which include: Kuchi, Girkua, Shagari, Gidan Abuzai, Silame, Dan Bara, Sokoto north, Gidan Daji, Rabah, Wurno, Goronyo, Gada, Sarkakarwa, Dagoza and Tamaru. Faecal samples were obtained from the selected cattle slaughtered at abattoirs from each of these localities and subsequently tested at Parasitology Laboratory of Usmanu Danfodio University, Sokoto for *F.gigantica* infection.

The aim of this validation was to ascertain the suitability of the visited sites to *F. gigantea* infection and not otherwise as applied by (Leta et al., 2015) in determining the accuracy of MaxEnt in their study of spatial analysis of tsetse flies in Ethiopia. In addition, failure to detect species during the field work at any of the sites did not guarantee complete absence as it might be due to a random chance (MacKenzie, 2005). Moreover, other contributing factors for the absence of the species could be due to biological, geographic and or human influence (Anderson and Martínez-Meyer, 2004).

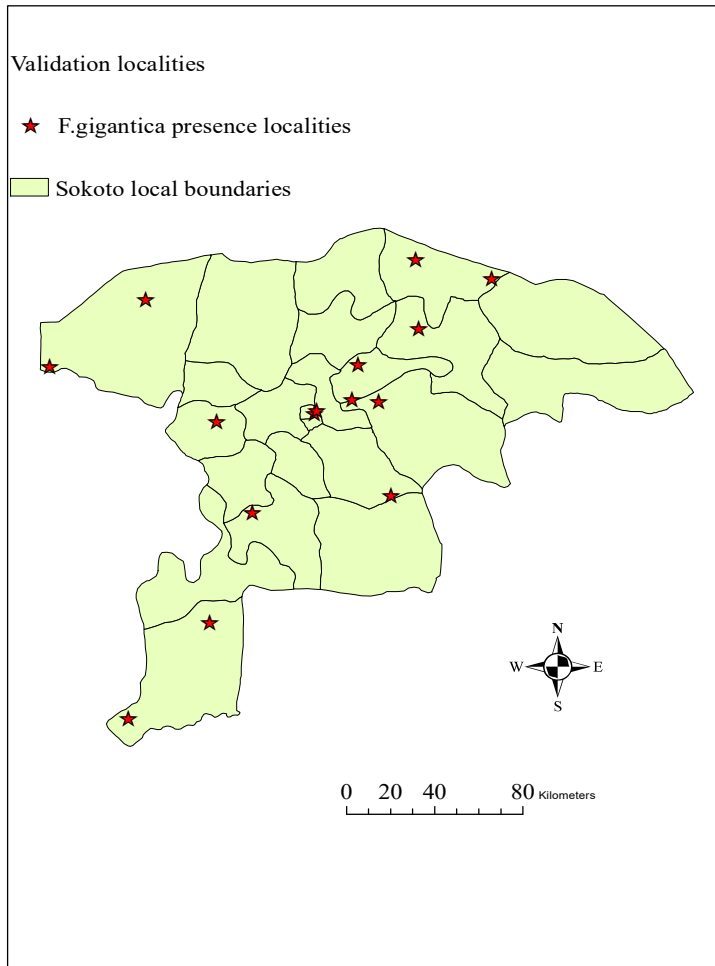


Figure 2: The marked sites are *F.gigantica* presence localities. These sites are visited during field work and used as independent data for model validation

The low number of presence records in this study used for both model training and validation proved to be adequate. Because it was discovered in other studies using GARP alone, and in its comparison with MaxEnt, 50 and 15 records respectively produced acceptable results that were statistically significant (Stockwell and Peterson, 2002, Papes and Baubert, 2007). Also, according to Pearson et al. (2007) as few as five records using MaxEnt alone produced accurate modelling results, which therefore adds confidence in the reliability of the results obtained in the present study.

Results

By reducing the variables that were highly correlated through the utility of Pearson's correlation coefficient, sixteen variables were retained and entered into the three modelling scenarios. These scenarios (Table 1) produced predictions that were better than random. Using the 10th percentile training threshold, the binomial test of omission was significant statistically for all the scenarios (P from 0.003 to 0.03) while their threshold values ranged from 0.2992 to 0.3144. Scenario 1 that

constitutes temperature and precipitation variables only predicted the largest area (0.42). Regarding the omission rate, scenario 2 ranked highest (20%) while scenario 3 has the lowest omission rate (18%). Similarly, the True Statistics Scores (TSS) for each one of the scenarios was better than random chance agreement as the values ranged from 0.30 to 0.41 and were statistically significant (averaged $p=0.01$ at the one-tailed test of binomial probabilities).

Table 1: Results of the threshold-dependent binomial tests of omission based on 10% percentile training presence

Scenarios	Logistic threshold	Fractional predicted area	Test omission rate	P-value	TSS
1	0.31	0.42	0.19	0.01	0.40
2	0.30	0.35	0.20	0.03	0.30
3	0.30	0.41	0.18	0.003	0.41

For all the scenarios (Table 2) using the combination of different climatic and environmental variables, the AUC values based on independent validation data were above 0.5 indicating better than random prediction and were statistically significant ($p=0.03$ at two-tailed Wilcoxon signed-ranked test). The addition of soil moisture in scenario two did not increase the AUC score, but the inclusion of NDVI and SRTM elevation in scenario 3 raised AUC to 0.7511. Regarding the fitness of the training data to the model the performances were very good across all the three MaxEnt modelling scenarios with $AUC \geq 0.8$.

Table 2: Result of threshold-independent measure of modelling methods

Scenarios	Total Number of records for training	Total number of records for testing	Number of variables used	AUC training data	AUC for test data
1	177	15	8	0.84	0.75
2	177	15	8	0.80	0.72
3	177	15	11	0.85	0.75

The jackknife test (Figures 3-5) of variable importance results indicated that precipitation of the wettest quarter (bio_16), annual mean moisture index (bio_28) and maximum annual NDVI (NDVI_36) are the three most influential variables that constrained the geographic range of *F. gigantea* in Sokoto State, which collectively accounted for more than 60% variation (Table 3). Other variables that were very influential in the modelling across the whole scenarios are maximum annual NDVI (NDVI_36), elevation (SRTM_39), precipitation of the warmest quarter (bio_18) and mean temperature of the coldest month (bio_6).

Table 3: Mean percent contribution (out of 500 iterations) of climatic variables to the prediction of geographic range of *F. gigantica* in Sokoto State, Nigeria using MaxEnt's modelling scenarios

Variable	Percent contribution		
	Scenario1	Scenario2	Scenario3
bio_1	10.8		
bio_2	10.7		
bio_3	1.6	0.3	0.4
Bio_4			6.8
Bio_6	21.4	20.3	8.3
Bio_9	5.7	5	1.3
Bio_12			5.8
Bio_16	24.2	5.8	
Bio_17		2.4	4.2
Bio_18	17.8	9.2	2.3
Bio_19	7.8		
Bio_28		49.8	36.5
Bio_33		7.2	
NDVI_36			23.2
NDVI_38			1.1
SRTM_39			10.1

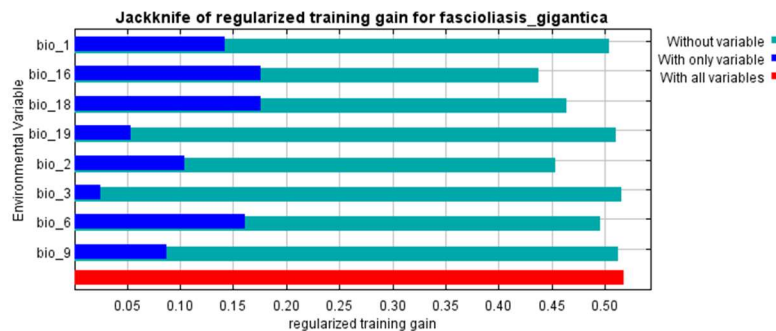


Figure 3: Scenario 1 results of the Jackknife test for the MaxEnt model for Sokoto State showing the gain of each variable to the likelihood map model. It uses temperature and precipitation variables (BIO1-19). Precipitation of the wettest quarter (BIO_16) and precipitation of the warmest quarter (BIO_18) have got the highest contribution in modelling the geographical distribution of *F. gigantica* in Sokoto State in this scenario

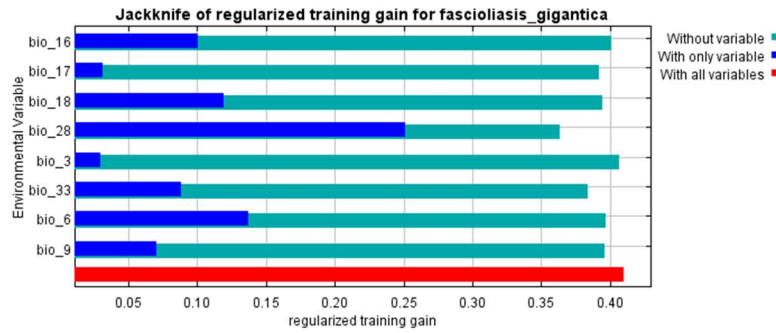


Figure 4: Jackknife test for scenario 2 MaxEnt model for Sokoto state showing the gain of each variable in the predicted probability map model. It contained temperature, precipitation and soil moisture variables only. When used alone in running the model annual mean moisture index (BIO_28) and precipitation of the warmest quarter (BIO_18) have got the highest contribution in modelling *F. gigantica* in Sokoto State in this scenario.

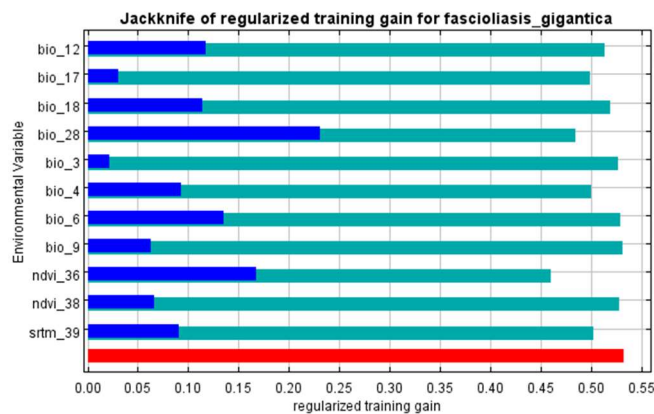


Figure 5: Scenario 3 (that contained temperature, precipitation, soil moisture, NDVI and SRTM elevation) results of the Jackknife test for the MaxEnt model for Sokoto state showing the gain of each variable in the predicted probability map model. In this scenario annual mean moisture index (BIO_28) and mean maximum NDVI (NDVI_36) were the most influential in modelling the geographical distribution of *F. gigantica* in Sokoto State

Geographic distribution of *F. gigantica* prevalence

The predicted areas of suitability for *F. gigantica* prevalence in Sokoto State by scenario 1 and 3 were more extensive than scenario 2 (Figure 6 and 8 respectively). All the MaxEnt modelling scenarios fitted the current distribution significantly and hence there were areas of suitability that are common to all of them. These areas of suitability encompass both central and northern parts of the study area in the provinces of Gada, Yabo, Goronyo and Wamakko. Around the central part of Sokoto State, all the three modelling scenarios indicated availability of favourable climatic conditions for the survival of *F. gigantica* parasite in the provinces of Sokoto north and south, Kware, Wamakko, Yabo and Rabah.

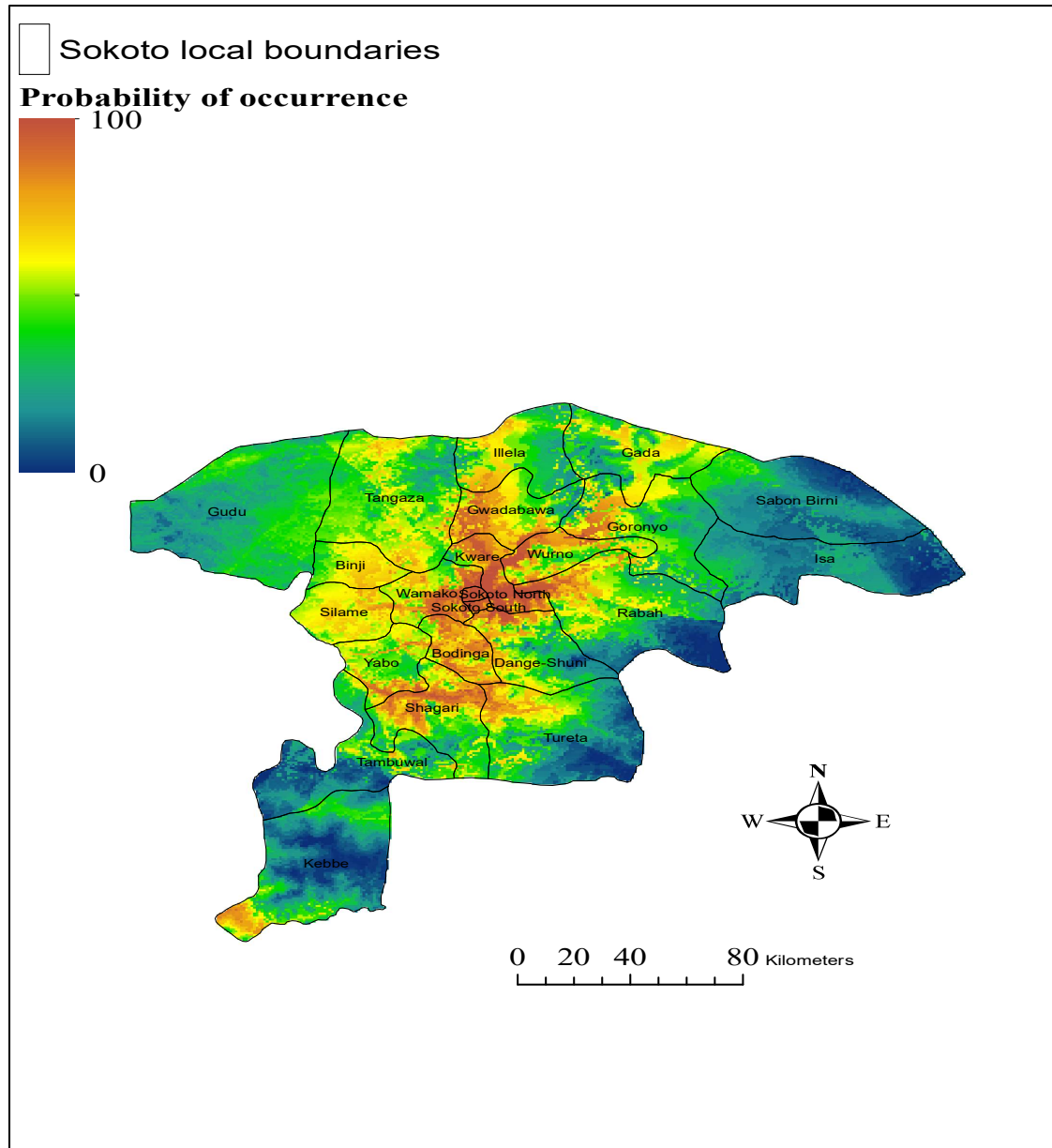


Figure 6: Dark brown colour implies a high probability of suitable conditions for *F. gigantica*, lighter shades of blue implying low predicted the probability of suitable conditions based on scenario 1. That indicates that suitable conditions are predicted to be highly likely through all the four agricultural zones of Sokoto State

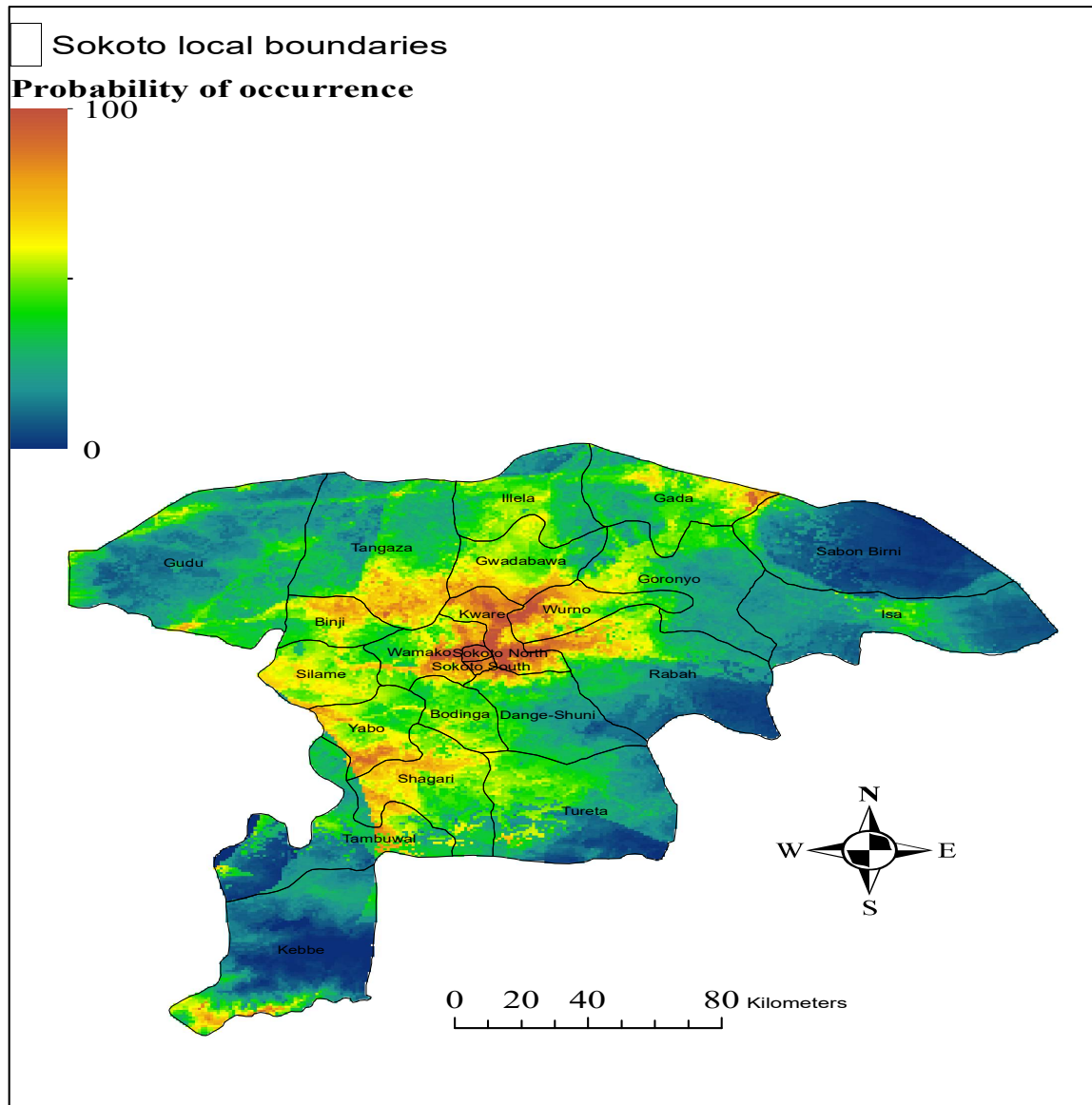


Figure 7: Scenario 2 MaxEnt's prediction using dark brown colour indicating the probability of suitable conditions and lighter shades of blue implying low predicted the probability of suitable conditions. That prediction is not as extensive as scenario 1 and more concentrated around the centre of the state.

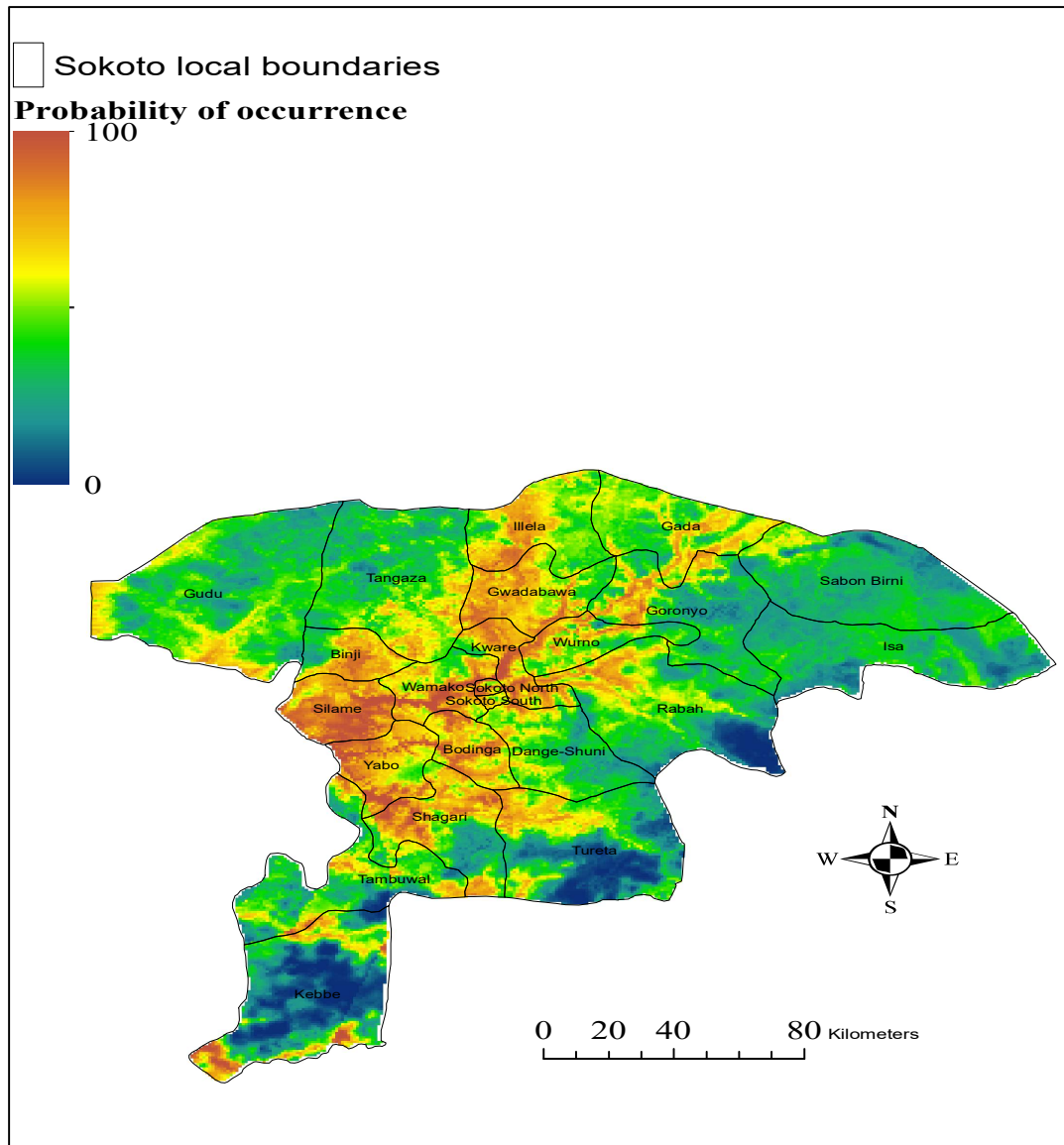


Figure 8: Showing wider spread of suitable condition based on Scennario_3 using dark brown colour in implying a high probability of suitable conditions and lighter shades of blue indicating low predicted probability of suitable conditions for *F.gigantica*. That gives more extensive coverage of Sokoto State than the other two scenarios.

The provinces with high densities of livestock population (Table 3) and with high probabilities of *F.gigantica* occurrence include Sokoto north and Sokoto south, Kware, Wamakko and Silame. The data on livestock population densities was obtained from Sokoto State Ministry of Animal Health and Fisheries Development.

Table 4. Distribution of livestock population density in Sokoto State in decreasing order.

Provinces	Density per km ²
Sokoto south	5585.614
Sokoto north	4634.171
Sabon-Birni	1130.196
Isa	806.3161
Kware	767.073
Bodinga	745.148
Wamakko	584.484
Silame	580
Wurno	471
Dange shuni	451.781
Binji	400.906
Yabo	397.143
Gwadabawa	273.166
Tambuwal	254
Gada	249.698
Shagari	235.931
Goronyo	200.595
Illela	190.628
Kebbe	163.534
Rabah	153.789
Tureta	111.322
Gudu	76.781
Tangaza	57

Discussion

In Nigeria, the prevalence of *F. gigantica* is high due to favourable prevailing climatic and environmental conditions. Also, there is the availability of extensive areas of low-elevation referred to as floodplains or Fadamas that are more common in northern Nigeria that support the existence of water bodies in lakes, ponds and streams (Magaji et al., 2014). Sokoto state is part of northwestern Nigeria and is among the leading producers of animals in the country (Mamman, 2005) with a population of over three million people (NPC, 2006). The state has an estimated 4 million cattle, sheep and goats (Bala et al., 2014) where the majority of the inhabitants practice animal rearing as an alternative to crop farming due to sparse arable land (Magaji et al., 2014). All the known studies on fascioliasis in Nigeria have focused on reporting prevalence using abattoir records (Bunza et al., 2008b, Magaji et al., 2014, Danbirni et al., 2015, Elelu et al., 2016a). Studies that applied species distribution models to examine the geographic range of fascioliasis in Nigeria or any part of West Africa in response to climatic variables are very rare (Yaro et al., 2018). Application of maximum entropy using different combinations of variables is a recommended approach in modelling potential distributional range of species (Phillips et al., 2006).

MaxEnt algorithm in all the three scenarios performed better than random predictions that were statistically significant. The threshold-dependent binomial test (Table 1) indicated independent test sites omission rates and fractional predicted area test that were statistically significant for all the three scenarios. Similarly, the AUC being the threshold-independent measure indicated better than random scores for all the scenarios. That reflected the suitability of Bioclim variables in species distribution modelling due to their association with the biological mechanism of different

species of animals and plants (Ramirez and Jarvis, 2010). In addition, the predictions of the potential distribution of *F. gigantica* by all the scenarios were reasonable as they all indicated almost the same 'core' areas of suitability. These probability maps, therefore, show the location of most of the most suitable areas around the central part of Sokoto State, in the provinces of Sokoto north and south extending to Goronyo, Wamakko, Shagari, Silame, Binji, Tangaza and Gwadabawa. These core areas of suitability are cut across by river Rima with large expanse of 'Fadama' which is a Hausa name that denotes area liable to flooding or 'floodplains' that can support irrigation due to being 'low-lying and underlined by shallow aquifers' mostly found adjacent to significant Rivers in Nigeria (Dan-Azumi, 2010).

It is explicitly clear regarding the visual investigation that the MaxEnt algorithm produced a prediction of potential distribution for *F. gigantica* across all the three scenarios that fit the data significantly. This is because MaxEnt integrates and sums the contributions of all the climatic and environmental variables at each of the pixels that constitute the study area which reflects its advantage of 'additivity' (Phillips et al., 2006). This has enabled the production of more continuous predictions. MaxEnt also exhibits the capability to show continuity in prediction through differentiating between various levels of probability of species occurrence (Phillips et al., 2006).

Given the preceding discussion, it is evident and reasonable for MaxEnt to identify soil moisture and rainfall as the dominant predictor variables in all the MaxEnt's modelling scenarios in the study area. That, therefore, reflects the reality on the ground as the study area lies in the semi-arid ecological zone where the main source of moisture was rainfall that lasts for only 3 to 4 months of a year (Abdulrahim et al., 2013). In addition, other secondary sources of moisture for the survival of the parasite were the existing extensive floodplains or fadamas, irrigation sites and areas adjoining perennial lakes and streams. Similar to the findings of this study Yilma and Malone (1998) developed the geographic information system (GIS) forecast model to predict fascioliasis using rainfall and temperature variables in Ethiopia, East Africa. The results revealed that areas of high soil moisture content were more risky areas for fascioliasis prevalence.

The second most important variable as predicted by MaxEnt is precipitation of the wettest quarter that influenced the distribution of fascioliasis in Sokoto State, through seasonal supply of moisture needed for the dispersal, reproduction and free-living stages of fascioliasis (Andrews, 1999, Altizer et al., 2006, Mas-Coma et al., 2009). Rainfall also influences the activities of infected snails in the distribution and transmission of fascioliasis. This illustrates the significance of moisture availability in affecting every stage of the fascioliasis lifecycle. Desiccation or moisture deficit conditions and dryness for extensive period will lead to mortality of the parasite (Spithill et al., 1999).

The MaxEnt model also predicted mean diurnal temperature as the most important determinant of *F. gigantica* distribution in the study area. Due to the location of the State in the tropics, the prevailing mean temperature condition hardly falls below the minimum value required for the completion of the life cycle of *F. gigantica* throughout the year at optimum range of 16°C to 45°C (Saleha, 1991, Andrews, 1999, Graczyk and Fried, 1999, Fairweather et al., 1999, Spithill et al., 1999, Mas-Coma et al., 2009, Fox et al., 2011). According to Andrews (1999), the processes of embryonation and evolution in the life cycle of fascioliasis are significantly influenced by temperature within tolerable limits between 23°C to 30°C.

Normalised Difference Vegetation Index (NDVI) is an index that assesses the density and greenness of vegetation and was found by this study to be a very significant predictor of *F. gigantica* distribution. The locations of most suitable areas lie in the Fadama of Sokoto State that supports vegetation around lakes, ponds and streams. Thus vegetation provides pasture for animals and hence NDVI 'proved' to be an essential determinant of risk due to fascioliasis (Afshan

et al., 2014). In a similar development, Fuentes et al. (2001) and (Kantzoura et al., 2011b) described NDVI as an instrumental variable in the risk assessment due to fascioliasis.

Elevation made a significant contribution to the modelling of fascioliasis distribution in this study. That is because of the relationships between elevation and hydrological, geomorphological as well as biological processes (Moore et al., 1991). It also affects the temperature regime, precipitation formation and in the flow and accumulation of rainfall water at lowland areas of the landscape which subsequently determines soil moisture availability (Franklin, 2009). Moreover, soil moisture affects the growth of grass that animals feed on and if (the grass is) contaminated with cercariae aids the transmission of *F. gigantica*. In a related development, Zumaquero-Ri'os et al. (2013) and Rahman et al. (2017) report that low-elevation was found to be very suitable 'hotspots' and 'clusters' for the prevalence of *F. gigantica* specifically in Asia and Africa.

Sokoto North and South, and some peripheral provinces such as Sabon Birni, Silame, Wamakko and Bodinga were the main areas of suitability as predicted by the three MaxEnt's scenarios probability maps in this study that also supports a high density of animals (Table 4). This result is consistent with the findings by (Tum et al., 2004) that applied a geographic information system to create a model for mapping risk of fascioliasis in cattle in Cambodia where cattle density was confirmed to be a risk factor for fascioliasis transmission. Fabiyi and Adeleye (1982) corroborated that fact with a report that in Nigeria the morphology of fascioliasis prevalence is consistent with zones of high animal density among others.

Conclusions

This study, compared the performance of three MaxEnt modelling scenarios in determining the geographic range of *F. gigantica* in Sokoto State. The findings revealed that the differences in the combination of variables across the scenarios did not affect their accuracy performance greatly. In addition, the central part of the state was at higher risk of *F. gigantica* infection than other parts of the state due to more favourable climatic conditions. Precipitation of the wettest quarter (bio_16), annual mean moisture index (bio_28) and maximum annual NDVI (NDVI_36) are the most important climatic and environmental variables that affected the geographic range of fascioliasis parasite in the study area.

However, it is very significant to highlight that by modelling the geographic range of *F. gigantica* in Sokoto state in Nigeria in the present study; it does not indicate the precise limits of *F. gigantica* distribution in the whole of Nigeria. Instead, the modelling techniques applied have identified provinces that share the same climatic conditions with the known occurrence sites of *F. gigantica*. These modelling results can support biogeographic information and regard as a first effort to estimate the geographic range of *F. gigantica* in Sokoto State, Nigeria.

The results suggest that future research should focus on increasing the extent of the study area as well as the spatial resolution of the environmental variables which may affect the predictive abilities of different modelling techniques (Karl et al., 2000, Hernandez et al., 2006). Species distribution modelling cannot be a perfect match to field surveying (Hernandez et al., 2006, Tognelli et al., 2009) but can provide a necessary guide in the planning of appropriate target areas for the control of *F. gigantica* prevalence. Moreover, all the modelling scenarios in our study made use of climatic variables but did not integrate other factors that affect livestock management in the study area. These factors according to (Kantzoura et al., 2011a) also affect the modelling of the geographic range of fascioliasis. However, using this approach, the models of species distribution can be evaluated and MaxEnt model can be more confidently applied by

animal and public health planners in the design of the field survey for the control of *F. gigantica* prevalence or any parasitic pathogens with a similar pattern of transmission in the study area.

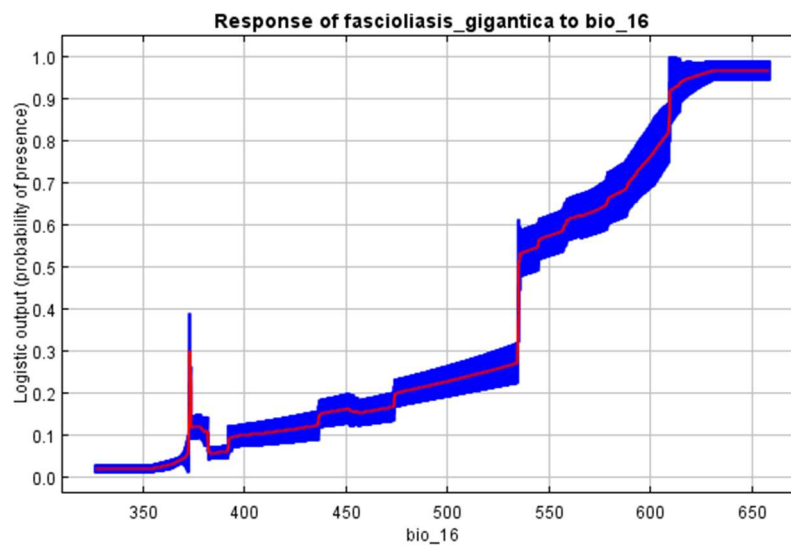
Supplementary Materials: The datasets used and analysed during the current study are included within the article and its additional file.

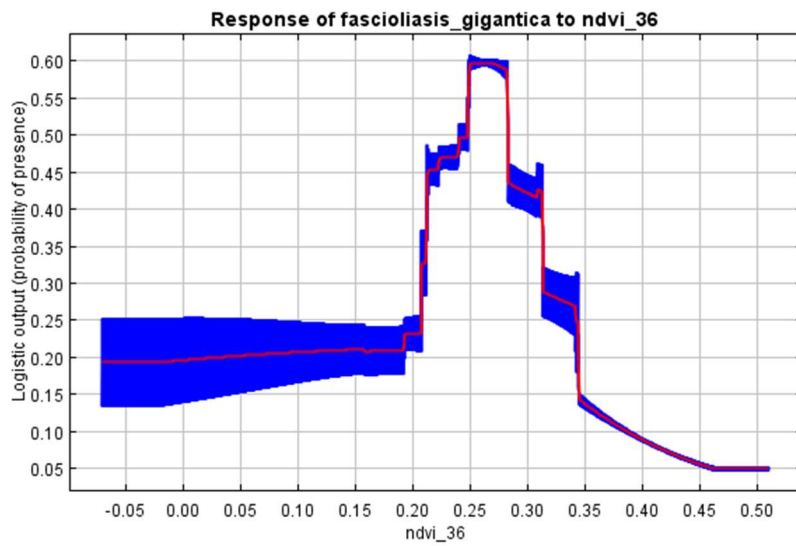
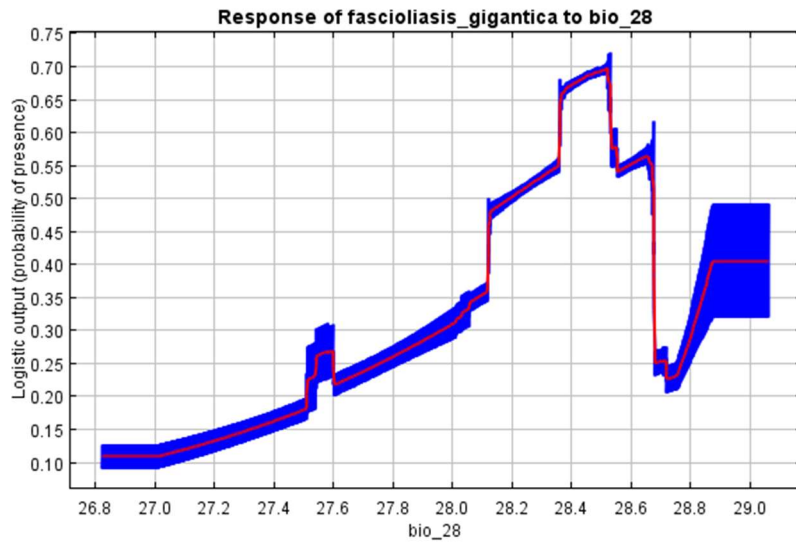
Author Contributions: conceptualization, IH. And HB. Methodology, IH. Software, IH. Validation, IH. HB. And JK. Formal analysis, IH. Investigation, IH. Resources, HB. Data curation, IH. Writing—original draft preparation, IH. Writing—review and editing, HB, and JB.; visualization, IH. Supervision, HB and JB.; project administration, HB and JB.; funding acquisition, IH.

Funding: The PhD tuition was awarded to IH being the corresponding author by the government of the federal republic of Nigeria Tertiary Education Trust Fund (TETFUND) and a one year living expenses by Petroleum Education Trust Fund (PTDF).

Acknowledgments: We thank the Ministry of Animal health Sokoto for providing the data for the spatial analysis in this study. We are also grateful to the University of Sokoto, Sokoto state, Nigeria and the University of Leicester, the UK for assisting with the PhD study.

Conflicts of Interest: The authors declare that they have no competing interest.





References:

- ABDULRAHIM, M. A., IFABIYI, I. P. & ISMAILA, A. 2013. Time series analyses of mean monthly rainfall for drought management in Sokoto, Nigeria. *Ethiopian Journal of Environmental Studies and Management*, 6.
- ACCADIA, C., MARIANI, S., CASAIOLI, M., LAVAQNINI, A. & SPERANZA 2005. Verification of precipitation forecasts from two limited-area models over Italy and comparison with ECMWF forecasts using a resampling technique. *Weather and Forecasting*, 20, 276-300.
- AFSHAN, K., FORTES-LIMA, C. A., ARTIGAS, P., VALERO, M. A., QAYYUM, M. & S, M.-C. 2014. Impact of climate and Man-made irrigation systems on the transmission risk, long-trem trend and seasonality of human and animal fascioliasis in Pakistan. *Geospatial Health*, 8, 317-334.
- 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, 1223-1232.
- ALTIZER, S., DOBSON, A., HOSSEINI, P., HUDSON, P., PASCUAL, M. & ROHANI, P. 2006. Seasonality and the dynamics of infectious diseases. *Ecology letters*, 9, 467-484.
- ANDERSEN, M. C., ADAMS, H. & HOPE, B., & POWELL, M. 2004. Risk analysis for invasive species: general framework and research needs. *Risk Analysis*, 24, 893-900.
- ANDERSON, R. P., LEW, D. & PETERSON, A. T. 2003. Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecol. Model.*, 211-232.
- ANDERSON, R. P. & MARTÍNEZ-MEYER, E. 2004. Modeling species' geographic distributions for preliminary conservation assessments: an implementation with the spiny pocket mice (*Heteromys*) of Ecuador. *Biol. Conser.*, 116.
- ANDREWS, S. J. 1999. The Life Cycle of *Fasciola hepatica*. In: J.P.DALTON (ed.) *Fasciolosis*. Wellington, Oxon, UK: CABI INTERNATIONAL.
- ARAUJO, M. & GUISAN, A. 2006. Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, 33, 1677-1688.
- ARNTZEN, J. W. 2006. From descriptive to predictive distribution models: a working example with Iberian amphibians and reptiles. *Frontiers in zoology*, 3, 38.
- ATEDHOR, G. 2015. Agricultural Vulnerability to Climate Change in Sokoto State, Nigeria. *African Journal of Food, Agriculture Nutrition and Development*, 15, 9855-9871.
- AUSTIN, M. 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecol. Model.*, 157, 101-121.
- BALA, A., SULEIMAN, N., JUNAIDU, A. U., SALIHU, M. D., IFENDE, V. I. & SAULAWA, M. A. 2014. Detection of Lead (Pb), Cadmium (Cd), Chromium (Cr) Nickel (Ni) and Magnesium Residue in Kidney and Liver of Slaughtered Cattle in Sokoto Central Abattoir, Sokoto State, Nigeria. *International Journal of Livestock Research* 1, 74-80.
- BERGER, A. L., DELLA PIETRA, S. A. & DELLA PIETRA, V. J. 1996. A maximum entropy approach to natural language processing. *Comput. Linguist.*, 22, 39-71.
- BUNZA, M. D. A., ABUBAKAR, U. I., ADAMU, T. & DANEJI, A. 2008b. Prevalence of Ruminant Fascioliasis in Kebbi State, Nigeria. *Nigerian Journal of Basic And Applied Sciences*, 16, 28-35.

- CDC. 2010. *Laboratory Identification of Parasites of Public Health Concern*. [Online]. Available: www.dpd.cdc.gov/dpdx/HTML/ImageLibrary/Fascioliasis_il.htm. [Accessed July 23, 2010].
- CLARK, P. J. & EVANS, F. C. 1954. Distance to Nearest Neighbor as a Measure of Spatial Relationships in Populations. *Ecology*, 35, 445-453.
- CUMMING, G. S. 2000. Using between-model comparisons to fine-tune linear models of species ranges. *Journal of Biogeography*, 27, 441-455.
- DAN-AZUMI, J. 2010. Agricultural sustainability of <I>fadama</I> farming systems in northern Nigeria: the case of Karshi and Badeggi. *International Journal of Agricultural Sustainability*, 8, 319-330.
- DANBIRNI, S., ZIYAUHAQ, H., ALLAM, L., OKAIYETO S.O & SACEY, A. K. B. 2015. Prevalence of Liver Comdemnation Due to Fascioliasis in Slaughtered Cattle and its Financial Losses at Kano Old Abattoir, Nigeria. *J.Vet. Adv*, 5.
- DAVIS, G. 1982. Rainfall and Temperature: Sokoto State in Maps. *An Atlas Physical and Human Resources*. John Wiley and Sons. Inc., New York.
- DE LEEUW, A. J., OTTICHILO, W. K., TOXOPEUS, A. G. & PRINS, H. H. T. 2002. Application of remote sensing and geographic information systems in Wildlife mapping and modelling. In: A.SKIDMORE (ed.) *Environmental modelling with GIS and remote sensing (pp.121-145)*. London: Taylor & Francis.
- DELEO, J. M. Receiver operating characteristic laboratory (ROCLAB): software for developing decision strategies that account for uncertainty. Proceedings of the Second International Symposium on Uncertainty Modelling and Analysis,, 1993 College Park, MD:. IEEE Computer Society Press., pp. 318–25.
- DELLA PIETRA, S., DELLA PIETRA, V. & LAFFERTY, J. 1997. Inducing features of random fields. *Pattern Anal. Mach. Intell.*, 19, 1–13.
- DI COLA, V., BROENNIMANN, O., PETITPIERRE, B., RANDIN, C., ENGLER, R., DUBUIS, A., D'AMEN, M., PELLISSIER, L., POTTIER, J., PIO, D., SALAMIN, N., BREINER, F., MATEO, R., HORDIJK, W. & GUISAN, A. 2017. *ecospat: An R package to support spatial analyses and modeling of species niches and distributions*.
- DUDIŁ, M., PHILLIPS, S. J. & SCHAPIRE, R. E. Performance guarantees for regularized maximum entropy density estimation. Proceedings of the 17th Annual Conference on Computational Learning Theory,, 2004 2004 New York,. ACM Press, .
- ELELU, N., AMBALI, A., COLES, G. C. & EISLER, M. C. 2016a. Cross-sectional study of Fasciola gigantica and other trematode infections of cattle in Edu Local Government Area, Kwara State, north-central Nigeria. *Parasit Vectors*, 9, 470.
- ELITH, J. 2002. Quantitative methods for modeling species habitat: comparative performance and an application to Australian plants. In: FERSON, S., BURGMAN, M. (EDS.), (ed.) *Quantitative Methods for Conservation Biology*. Springer-Verlag, New York.
- ELITH, J., GRAHAM, C. H., ANDERSON, R. P., DUDIŁ, M., FERRIER, S. & GUISAN, A., ET AL 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29, 129-151.
- ELITH, J., KEARNEY, M. & PHILLIPS, S. J. 2010. The art of mod-elling range-shifting species. *Methods in Ecology and Evolution*, 1, 330–342.

- ELITH, J. & LEATHWICK, J. R. 2009. Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution and Systematics*, 40, 677-697.
- ELITH, J., PHILLIPS, S. J., HASTIE, T., DUDÍK, M., CHEE, Y. E. & YATES, C. J. 2011. A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, 17, 43-57.
- FABIYI, J. P. & ADELEYE, G. A. 1982. Bovine Fascioliasis on the Jos plateau, Northern Nigeria with particular reference to economic importance. *Bulletin of Animal Health and Production in Africa*, 30, 41-43.
- FAIRWEATHER, I., THREADGOLD, L. T. & HANNA, R. E. B. 1999. Development of *Fasciola hepatica* in the Mammalian Host. In: DALTON, J. P. (ed.) *Fasciolosis*. Wellington, Oxon, UK: CABI International.
- FARR, T. G., ROSE, P., CARO, E., CRIPPEN, R., DUREN, R., HENSLEY, S., KOBRICK, M., PELLER, M., RODRIGUEZ, E., ROTH, L., SEAL, D., SHAFFER, S., SHIMADA, J., UMLAND, J., WERNER, M., OSKIN, M., BURBANK, D., ALSDORF, D. & 2007. The Shuttle Radar Topography Mission. *Reviews of Geophysics*, 45.
- FIELDING, A. H. & BELL, J. F. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ. Conservation*, 24, 38-49.
- FOURCADE, Y., ENGLER, J. O., RODDER, D. & SECONDI, J. 2014. Mapping species distributions with MAXENT using a geographically biased sample of presence data: a performance assessment of methods for correcting sampling bias. *PLoS One*, 9, e97122.
- FOX, N. J., WHITE, P. C., MCCLEAN, C. J., MARION, G., EVANS, A. & HUTCHINGS, M. R. 2011. Predicting impacts of climate change on *Fasciola hepatica* risk. *PLoS One*, 6, e16126.
- FRANKLIN, J. 2009. *Mapping Species Distributions: Spatial Inference and Prediction*, Cambridge University Press, Cambridge, UK.
- FREEMAN, E. A. & MOISEN, G. G. 2008a. A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. *Ecological Modelling*, 217, 48-58.
- FREEMAN, E. A. & MOISEN, G. G. 2008b. A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. *Ecol Modell*, 48-58.
- FUENTES, M. V., MALONE, J. B. & MAS-COMA, S. 2001. Validation of a mapping and prediction model for human fasciolosis transmission in Andean very high altitude endemic areas using remote sensing data. *Acta Tropica*, 79, 87-95.
- GRACZYK, T. & FRIED, B. 1999. Development of fasciola in the intermediate host. In: DALTON, J. P. (ed.) *Fasciolosis*. Wellington, Oxon, UK: CABI INTERNATIONAL.
- GUISAN, A., EDWARDS JR., T. C. & HASTIE, T. 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecol. Model.*, 157, 147-186.
- 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, 773-785.

- HERNENDEZ, 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, 773-785.
- HIJMAN, R. J., PHILLIPS, S., LEATHWICK, J. & ELITH, J. 2011. dismo: Species distribution modeling. R package version.
- HOSMER, D. W. & LEMESHOW, S. 2000. Applied Logistic Regression. Wiley, New York.
- IRVING, J. A., SPITHILL, T. W., PIKE, R. N., J.C.WHISSTOCK, P. M. & SMOOKER, P. 2003. The evolution of enzyme specificity in *Fasciola* spp. *Journal of Molecular Evolution*, 57, 1-15.
- JARNEVICH, C. S. & REYNOLDS, L. V. 2010. Challenges of predicting the potential distribution of a slow-spreading invader: a habitat suitability map for an invasive riparian tree. *Biological Invasions*, 13, 153-163.
- JAYNES, E. T. 1957. Information theory and statistical mechanics. *Phys. Rev.*, 106,, 620–630.
- KANTZOURA, V., KOUAM, M. K., DEMIRIS, N., FEIDAS, H. & THEODOROPOULOS, G. 2011a. Risk factors and geospatial modelling for the presence of *Fasciola hepatica* infection in sheep and goat farms in the Greek temperate Mediterranean environment. *Parasitology*, 138, 926-38.
- KANTZOURA, V., KOUAM, M. K., FEIDAS, H., TEOFANOVA, D. & THEODOROPOULOS, G. 2011b. Geographic distribution modelling for ruminant liver flukes (*Fasciola hepatica*) in south-eastern Europe. *Int J Parasitol*, 41, 747-53.
- KARL, J. W., HEGLUND, P. J., GARTON, E. O., SCOTT, J. M., WRIGHT, M. N. & HUTTO, R. L. 2000. Sensitivity of species habitat-relationship model performance to factors of scale. *Ecological Applications*, 10, 1690-1705.
- KREMEN, C., CAMERON, A. & MOILANEN, A. E. A. 2008. Aligning conservation priorities across taxa in Madagascar with high-resolution planning tools. *Science*, 320, 222-226.
- KRITICOS, D. J., JAROŠIK, V. & N., A. O. 2014. Extending the suite of Bioclim variables: a proposed registry system and case study using principal components analysis. *Methods in Ecology and Evolution*.
- KRITICOS, D. J., WEBBER, B. L., LERICHE, A., OTA, N., MACADAM, I., BATHOLS, J. & SCOTT, J. K. 2012. CliMond: global high resolution historical and future scenario climate surfaces for bioclimatic modelling. *Methods in Ecology and Evolution*, 3, 53-64.
- LETA, S., HABTAMU, Y., ALEMAYEHU, G., AYELE, B., CHANIE, M., TESFAYE, S. & MESELE, F. 2015. Spatial analysis of the distribution of tsetse flies in Ethiopia using high resolution environmental datasets and Maxent...
- LETA, S., HABTAMU, Y., ALEMAYEHU, G., SHIGUTE, B., CHANIE, M., TESFAYE, S. & MESELE, F. 2015. *Spatial analysis of the distribution of tsetse flies in Ethiopia using high resolution environmental datasets and Maxent modeling technique*.
- LIU, C., BERRY, P. M., DAWSON, P. T. & PEARSON, R. G. 2005. Selecting thresholds of occurrence in the prediction of species distributions. *ECOGRAPHY* 28, 385-393,.
- MACKENZIE, D. 2005. Was it there? Dealing with imperfect detection for species presence/absence data. *Aust.N.Z.J Stat.*, 47, 65-74.

- MAGAJI, A. A., IBRAHIM, K., SALIHU, M. D., SAULAWA, M. A., MOHAMMED, A. A. & MUSAWA, A. I. 2014. Prevalence of Fascioliasis in Cattle Slaughtered in Sokoto Metropolitan Abattoir, Sokoto, Nigeria. *Advances in Epidemiology*, 2014, 1-5.
- MAMMAN, A. B. 2005. Transport aspect of livestock marketing at Achida and Sokoto Kara Markets. . Paper prepared on a network supported by UK Department of International Development (DFID) Sokoto.
- MANEL, S., WILLIAMS, H. C. & ORMEROD, S. J. 2001. Evaluating presence-absence models in ecology: the need to account for prevalence. - *J. Appl.*, 38, 921-931.
- MARTÍNEZ-MEYER, E., PETERSON, A. T., SERVIN, J. I. & KIFF, L. F. 2006. Ecological niche modelling and prioritizing areas for species reintroductions. *Oryx*, 40, 411-418.
- MAS-COMA, S. 2004. Chapter 19: Human fascioliasis. *In*: COTRUVO, J. A., DUFOUR, A., REES, G., BARTRAM, J., CARR, R., CLIVER, D. O., CRAUN, G. F. & FAYER, R. G., V.P.J. (eds.) *World Health Organization (WHO), Waterborne Zoonoses: Identification, Causes and Control*. London, UK:: IWA Publishing,.
- MAS-COMA, S., BARGUES, M. D. & VALERO, M. A. 2005. Fascioliasis and other plant-borne trematode zoonoses. *International Journal for Parasitology*, 35, 1255-1278.
- MAS-COMA, S., VALERO, M. A. & BARGUES, M. D. 2009. Climate change effects on trematodiasis, with emphasis on zoonotic fascioliasis and schistosomiasis. *Vet Parasitol*, 163, 264-80.
- MCCANN, C. M., BAYLIS, M. & WILLIAMS, D. J. 2010. The development of linear regression models using environmental variables to explain the spatial distribution of *Fasciola hepatica* infection in dairy herds in England and Wales. *Int J Parasitol*, 40, 1021-8.
- MEROW, C., SMITH, M. J. & SILANDER, J. A. 2013. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography*, 36, 1058-1069.
- MOORE, I. D., GRAYSON, R. & LADSON, A. 1991. Digital terrain modelling: a review of hydrological, geomorphological, and biological applications. *Hydrological Processes*, 5, 3-30.
- MORUETA-HOLME, N., CAMILLA, F. & JENS-CHRISTIAN, S. 2010. Climate Change Risks and Conservation Implications for a Threatened Small-Range Mammal Species. *PLoS ONE*, 5.
- NEGGA, H. E. 2007. *Predictive Modelling of Amphibian Distribution Using Ecological Survey Data: a case study of Central Portugal*. MSc, International Institute for Geo-Information Science and Earth Observation, Enschede, The Netherlands.
- NPC 2006. National Population Commission. *Census Data*.
- PAPES, M. & BAUBERT, P. 2007. Modeling ecological niches from low numbers of occurrences: assessment of the conservation status of poorly known viverrids (Mammalia, Carnivora) across two continents. *Diversity Distribution*, 13, 890-902.
- PEARSON, R. G., RAXWORTHY, C. J., NAKAMURA, M. & PETERSON, A. T. 2007. Predicting species' distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *J. Biogeogr.*, 34, 102-117.
- PETERSON, A. T. 2003. Predicting the geography of species invasions via ecological niche. *Quarterly Review of Biology*, 78, 419-433.

- PHILLIPS, S. J., ANDERSON, R. P. & SCHAPIRE, R. E. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190, 231-259.
- PHILLIPS, S. J., MIROSLAV, D. & E.SCHAPIRE., R. A Maximum Entropy Approach to Species Distribution Modeling. Proceedings of the Twenty-First International Conference on Machine Learning, . 2004
- 2004.
- PONDER, W. F., CARTER, G. A., FLEMONS, P. & CHAPMAN, R. R. 2001. Evaluation of museum collection data for use in biodiversity assessment. *Conserv. Biol.*, 15, 648-657.
- PROVOST, F. J. & FAWCETT, T. 1997. Analysis and visualization of classifier performance: comparison under imprecise class and cost distributions. *Knowledge Discovery and Data Mining. ACM Press, New York*,, 43-48.
- RAHMAN, A., ISLAM, S. K. S., TALUKDER, M. H., HASSAN, M. K., DHAND, N. K. & WARD, M. P. 2017. Fascioliasis risk factors and space-time clusters in domestic ruminants in Bangladesh. *Parasit Vectors*, 10, 228.
- RAMIREZ, J. & JARVIS, A. 2010. Disaggregation of Global Circulation Model Outputs Decision and Policy Analysis Working Paper No. 2. Cali, Colombia: International Center for Tropical Agriculture.
- RODRIGUEZ, E., MORRIS, C. S. & BELZ, J. E. 2006. A global assessment of the SRTM performance., *Photogramm. Eng. Rem. Sens.*,, 72, 249-260.
- SALEHA, A. A. 1991. Liverfluke disease (Fascioliasis):Epidemiology and public health significance. *Southeast Asian J Trop Med Public Health*.
- SOBEK-SWANT, S., KLUZA, D. A., CUDDINGTON, K. & LYONS, D. B. 2012. Potential distribution of emerald ash borer: What can we learn from ecological niche models using Maxent and GARP? *Forest Ecology and Management*, 281, 23-31.
- SPITHILL, T., SMOOKER, P. & COPEMAN, D. 1999. Fasciola gigantica: epidemiology, control, immunology and molecular biology In: Dalton JP (Ed.), . Fasciolosis. CABI Publishing, Oxon, UK.
- STOCKWELL, D. R. B. & PETERSON, A. T. 2002. Effects of sample size on accuracy of species distribution models. *Ecol. Model.*, 148, 1-13.
- SWETS, J. A. 1988. Measuring the accuracy of diagnostic systems. *Science*, 240,, 1285–1293.
- THORN, J. S., NIJMAN, V., SMITH, D. & NEKARIS, K. A. I. 2009. Ecological niche modelling as a technique for assessing threats and setting conservation priorities for Asian slow lorises (Primates: Nycticebus). *Diversity and Distributions*,, 15, 289-298.
- TOGNETTI, M. F., ROIG-JUNENT, S. A., MARVALDI, A. E., FLORES, G. E. & LOBO, J. M. 2009. An evaluation of methods for modelling distribution of Patagonian insects. *Revista Chilena De Historia Natural*,, 82,, 347–360.
- TOLAN, R. W. 2011. Fascioliasis Due to Fasciola hepatica and Fasciola gigantica Infection: An Update on This 'Neglected' Neglected Tropical Disease. *Laboratory Medicine*, 42, 107-116.
- TSOAR, A., ALLOUCHE, O., STEINITZ, O., ROTEM, D. & KADMON, R. 2007. A comparative evaluation of presence-only methods for modelling species distribution. *Diversity and Distributions*, 13, 397-405.

- TUM, S., PUOTINEN, M. L. & COPEMAN, D. B. 2004. A geographic information systems model for mapping risk of fasciolosis in cattle and buffaloes in Cambodia. *Veterinary Parasitology*, 122, 141-149.
- VALENTIA-LOPEZ, N., MALONE J.B, CARMONA C.G & VELASQUEZ, L. E. 2012. climate-based risks model for fasciolosis hepatica in Colombia. *Geospatial Health*, 6, 75-85.
- WHO. 2006. *Report of the WHO Informal Meeting on Use of Triclabendazole in Fascioliasis Control*. [Online]. WHO Headquarters, Geneva, Switzerland:October 17-18, 2006. Available: www.who.int/neglected_diseases/preventive_chemotherapy/WHO_CDS_NTD_PCT_2007.1.pdf. [Accessed October 17-18,2006.
- WIEGAND, C. L., RICHARDSON, A. J., ESCOBAR, D. E. & GEBERMANN, A. H. 1991. Vegetation indices in crop assessments,. *Remote Sens. Environ.*, 35, 105-119.
- WILLIAMS, P. M. 1995. Bayesian regularization and pruning using a Laplace prior. *Neural Comput.*, 7, 117-143.
- YARO, C. A., KOGI, E. & O, I. F. 2018. Bovine fasciolosis in Niger state, Nigeria: effects of climatic and elevation factors on its distribution. *MOJ Public Health.*, 7, 275-280.
- YILMA, J. M. & MALONE, J. B. 1998. A Geographic information system forecast model for strategic control of fasciolosis in Ethiopia. *Veterinary Parasitology*, 78.
- ZUMAQUERO-RÍOS, J., SARRACENT-PE´ REZ, J., ROJAS-GARCÍA, R., ROJAS-RIVERO, L. Z., MARTÍNEZ-TOVILLA, Y., VALERO, M. A. & MAS-COMA, S. 2013. Fascioliasis and Intestinal Parasitoses Affecting Schoolchildren in Atlixco, Puebla State, Mexico: Epidemiology and Treatment with Nitazoxanide. *Neglected Tropical Diseases*, 7.