

RESEARCH ARTICLE

Dominant malaria vector species in Nigeria: Modelling potential distribution of *Anopheles gambiae* sensu lato and its siblings with MaxEnt

Godwin E. Akpan^{1*}, Kayode A. Adepoju², Olakunle R. Oladosu¹, Samuel A. Adelabu²

1 African Regional Centre for Space Science and Technology Education in English (ARCSSTEE), Obafemi Awolowo University (OAU), Ile-Ife, Osun State, Nigeria, **2** Department of Geography, University of The Free State, Qwaqwa Campus, Qwaqwa, Phuthaditjhaba, South Africa

* profjedidiah@yahoo.com



OPEN ACCESS

Citation: Akpan GE, Adepoju KA, Oladosu OR, Adelabu SA (2018) Dominant malaria vector species in Nigeria: Modelling potential distribution of *Anopheles gambiae* sensu lato and its siblings with MaxEnt. PLoS ONE 13(10): e0204233. <https://doi.org/10.1371/journal.pone.0204233>

Editor: Aysegul Taylan Ozkan, Hitit University, Faculty of Medicine, TURKEY

Received: January 7, 2018

Accepted: September 5, 2018

Published: October 3, 2018

Copyright: © 2018 Akpan et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files. Occurrence data was obtained from Nigeria Anopheles vector database and can be found on: <https://doi.org/10.1371/journal.pone.0028347.s001>.

Funding: The author(s) received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Abstract

Malaria is a major infectious disease that still affects nearly half of the world's population. Information on spatial distribution of malaria vector species is needed to improve malaria control efforts. In this study we used Maximum Entropy Model (MaxEnt) to estimate the potential distribution of *Anopheles gambiae* sensu lato and its siblings: *Anopheles gambiae* sensu stricto, and *Anopheles arabiensis* in Nigeria. Species occurrence data collected during the period 1900–2010 was used together with 19 bioclimatic, landuse and terrain variables. Results show that these species are currently widespread across all ecological zones. Temperature fluctuation from mean diurnal temperature range, extreme temperature and precipitation conditions, high humidity in dry season from precipitation during warm months, and land use and land cover dynamics have the greatest influence on the current seasonal distribution of the *Anopheles* species. MaxEnt performed statistically significantly better than random with AUC approximately 0.7 for estimation of the *Anopheles* species environmental suitability, distribution and variable importance. This model result can contribute to surveillance efforts and control strategies for malaria eradication.

Introduction

Anopheles species have plagued the world with malaria for decades and centuries now. An estimated 3.2 billion people worldwide were at risk of malaria in 2014 [1]. In 2016, ninety-one (91) countries and territories in the world had ongoing malaria transmission with estimated 216 million cases of malaria and 445,000 malaria deaths [2]. Fifteen countries accounted for 80% of all malaria cases and deaths globally. Sub-Saharan Africa region was home to 90% of malaria cases and 91% of malaria deaths, globally. Nigeria accounted for the highest proportion of cases globally (27%), followed by the Democratic Republic of the Congo (10%), India (6%) and Mozambique (4%) [2].

Besides *Anopheles funestus*, *Anopheles gambiae* is the dominant and most efficient vector of human malaria in the Afrotropical Region [3–5], based on its high abundance, longevity, high propensity for humans feeding, and high vectorial capacity [6,7]. *Anopheles gambiae* complex herein referred to as *Anopheles gambiae* sensu lato (s.l.) is made up of eight reproductively isolated species that are almost indistinguishable morphologically: *An. arabiensis*, *An. gambiae* sensu stricto (s.s.), *An. bwambae*, *An. melas*, *An. merus*, *An. quadriannulatus*, *An. coluzzii*, and *An. amharicus* [8,9]. Environmental, social and demographic factors such as climate change/variability drive the distribution of the dominant malaria vector species and their parasite transmission [10,11]. They shift in response to changes in temperature and precipitation [12–15]. Onyabe and Conn [4] explained that climatological factors, especially total annual precipitation strongly influence the range and relative abundance of *An. arabiensis* and *An. gambiae* within forest zones and savannas [16–18], with *An. arabiensis* predominating during the dry season and *An. gambiae* becoming more abundant during the rainy season [19]. In line with recent studies [20,21], Onyabe and Conn [4] observed shifts in species composition of *An. arabiensis* and *An. gambiae* s.s. after two years in four of 10 localities in their study, attributing it to random temporal (seasonal) fluctuations. Also, Umar *et al.* [22] observed low population of female *Anopheles* mosquitoes during the dry seasons (January to June and October to December), while carrying out assessment of indoor resting density of female anopheline mosquitoes in human dwelling at malaria vector sentinel sites in Bauchi State, Nigeria. According to Umar *et al.* [22], higher densities of anopheline mosquitoes during the rainy season to a large extent explains the seasonal pattern of clinical cases of malaria, with peak transmission shortly after maximum annual rainfall [19,23].

An understanding of the temporal and spatial determinants of parasite transmission, its seasonal patterns and the dominant vectors implicated in transmission is crucial [24] for the control of vector species [25]. A reliable risk modelling is one of the precautionary means in the framework of public health and management of malaria, especially in view of climate change [26]. Numerous mathematical models have been applied for disease risk modelling [27–30]. Maximum entropy algorithm (MaxEnt)—a type of machine learning technique of ecological niche modelling has been proved to perform well in modelling the distribution of disease vectors and their possible disease transmissions including that of malaria [28–30]. MaxEnt aims to predict potential distribution of biological species from the observation of species occurrences [31,32]. Moffett *et al.* [28] used MaxEnt to construct niche models for 10 malaria vector species in Africa, and predicted that *An. gambiae* abundance was highest in West Africa followed by *An. arabiensis*, *An. funestus* and *An. melas*; with human population density as the critical factor determining malaria risk. Similarly, Kulkarni *et al.* [30] used MaxEnt in Northern Tanzania when they found seasonality of precipitation and maximum annual temperature to have contributed the most to niche models for *Anopheles arabiensis* and *An. funestus* s.l. with AUC of 0.989 and 0.991 respectively, cold season precipitation and elevation were also found important for *An. gambiae* s.s. with AUC of 0.997.

While studies on MaxEnt modelling of the *Anopheles* species [28–30] predicted potential distribution on all spatial locations within geographic area of interest, without considering ecological zones. The identification of the species distribution based on entomological surveys [18–22] considered the species presence in different ecological zones with respect to absolute locations rather than all spatial locations of interest. In this study, we used MaxEnt for modelling environmental suitability and potential distribution of these dominant malaria mosquitoes in all spatial locations across topographic relief, ecological and regional zones in Nigeria. We also assessed the contributions of bioclimatic and other environmental variables to the occurrence of these *Anopheles* species.

Materials and methods

Study area

Nigeria is a country in West Africa, located approximately between Latitudes 4° and 14° north of the Equator and between Longitudes 2° 2' and 14° 30' east of the Greenwich Meridian (Fig 1) [33]. It has a total area of 923,777 km² characterised by undulating topographic relief, patterned by valleys created by its river systems (Fig 1) [33,34]. Coastal plains in the south have mean elevation of about 150m above sea level. Northern plains rise to about 600-700m, with Jos Plateau (over 1,500m) within Nigeria's geographic centre, and Mambilla plateau (over 2,100m) amongst mountains at the border with Cameroon [35,36]. Temperature varies across ecological zones (Fig 1). Tropical at the coast (within Humid forest and Derived savanna) with 10°C and 37°C extreme low and high temperatures respectively, sub-tropical further inland (within Derived and Guinea savannas), and semi-arid in the far north (within Sudan and Sahel savannas) with 6°C and 44°C extreme low and high temperatures respectively [17,34]. Mid Altitude zone of Jos and Mambilla plateaus has average monthly temperatures range of 21–25°C [34]. Annual rainfall ranges from 500mm to 750mm in the north, and 1,200mm to above 4000mm in the south [34]. This diversity in climate conditions across the country affects the spatial epidemiology of malaria mosquitoes, malaria transmission and human vulnerability [24]. About 90% of over 190 million Nigerians are at the risk of malaria [37,38].

Modelling procedures and data analysis

Data resources. Malaria vectors data (Occurrence data: 1900–2010) was obtained from Nigeria *Anopheles* vector database (S1 Table); a comprehensive review of Okorie *et al.* [39]. Also, 19 bioclimatic variables (1960–1990) with about 1 km² spatial resolution were obtained from WorldClim—Global Climate Data (<http://www.worldclim.org>), global climate models (GCM)—community climate system model version 4 (CCSM4) [40] to model the impact of current climates on malaria vector species distribution in Nigeria. Additional variables used include land use land cover data with 24 classes obtained from U.S. Geological Survey data release [41], and Digital Elevation Model (DEM) derived from Shuttle Radar Topography

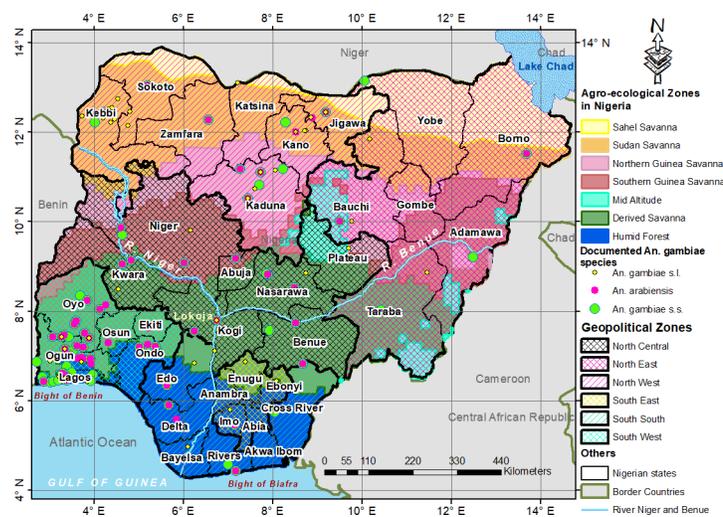


Fig 1. Study area with documented points of *Anopheles gambiae* species. Georeferenced *Anopheles* species locations reprinted for illustrative purposes only from Okorie *et al.* [39] under a CC BY 4.0 license, with permission from PLOS ONE.

<https://doi.org/10.1371/journal.pone.0204233.g001>

Mission (SRTM) 90m obtained from the Consultative Group on International Agricultural Research—Consortium for Spatial Information (CGIAR-CSI) [42].

Model operation. Model operation was carried out as implemented in Maximum entropy algorithm model (MaxEnt) version 3.3.3k, described in detail previously by Phillips *et al.* [43]. All environmental layers were modified in ArcMap to be at the same extent, and a bias layer was created to provide MaxEnt with background samples to guard against bias in datasets [44]. Locations where *Anopheles* species were sampled to occur was defined as 1; the remaining pixels that were not sampled had “no data” in the grid [45]. Taking advantage of all available data without having an independent dataset, occurrence data for each species was split twenty-one times into training (75%) and testing (25%) subsets under sub-sample replicated run type. This was done to test the model performance (robustness and predictiveness) given by Area Under the Receiver Operating Characteristic (ROC) Curve or AUC; a plot of sensitivity against specificity which measures the ability of the model to discriminate between sites where a species is present ($y = 1$) against where it is absent ($y = 0$) [46–48]. Maximum iterations was increased from 500 (default) to 5000, allowing the model to have adequate time for convergence guarding against over-prediction or under-prediction of the relationships by the model [47]. Other setting options were left at default including regularization of 1 that helped reduced model over-fitting.

Probability of species occurrence predicted by all the predictor environmental variables produced a point-wise mean (model images) [47]. This was classified in ArcMap for current distributions of *Anopheles gambiae* s.l. and its siblings. This was also classified into suitable and unsuitable habitats with 10 percentile training presence logistic threshold provided by MaxEnt. Suitable areas within ecological zones for the studied mosquito species were generated by MaxEnt based on the entropy of optimal climatic and other environmental conditions that match the empirical average (threshold value) within documented species records. The 10% minimum threshold meant that suitable habitat was defined to include 90% of the data used to develop the model, considering some errors the data used may likely had [47]. Zonal statistics in ArcGIS was used to determine the average distribution density of predicted *Anopheles* species in all ecological and geopolitical zones, and in each state. Distribution density less than or equal to one (≤ 1) defines the probability of the species not occurring, greater than one (> 1) defines species presence, while 2 represents maximum prevalence of species within a state and each zone. Analysis of percent contribution of each variable and jackknife test of variable importance were used to examine the contributions of environmental variables (Table 1) in defining the *Anopheles* species suitable habitats [30]. Jackknife test shows the training gain of each variable if the model was run in isolation, and compare it to the training gain with all the variables [47]. Comparing three jackknife plots produced for training gain, test gain and AUC gave very informed evaluations of variables contributions [46].

Results

Potential suitable areas for the occurrence and distribution of *An. gambiae* s.l. and its siblings

The result of MaxEnt modelling predicted that the approximately 85,000 square kilometre (km^2) Humid forest and 204,000 km^2 Derived savanna are highly suitable for the occurrence and distribution of *An. gambiae* s.l., *An. gambiae* s.s. and *An. Arabiensis*; with prevalence between 65% and 71% (Fig 2; Table 2). This makes all the states in South West, South East, South South and parts of North Central regions within the two ecological zones highly suitable for the *Anopheles gambiae* species (Fig 2; Table 2). The estimated 25,000 km^2 Mid Altitude zone of Jos and Mambilla plateaus, and highlands along boundary with Cameron (Fig 2) seem

Table 1. Environmental variables used.

Code	Bioclimatic/Ecological variables
BIO1	Annual Mean Temperature
BIO2	Mean Diurnal Range (Mean of monthly (max temp—min temp))
BIO3	Isothermality (BIO2/BIO7)*100
BIO4	Temperature Seasonality (standard deviation *100)
BIO5	Maximum Temperature of Warmest Month
BIO6	Minimum Temperature of Coldest Month
BIO7	Temperature Annual Range (BIO5-BIO6)
BIO8	Mean Temperature of Wettest Quarter
BIO9	Mean Temperature of Driest Quarter
BIO10	Mean Temperature of Warmest Quarter
BIO11	Mean Temperature of Coldest Quarter
BIO12	Annual Precipitation
BIO13	Precipitation of Wettest Month
BIO14	Precipitation of Driest Month
BIO15	Precipitation Seasonality (Coefficient of Variation)
BIO16	Precipitation of Wettest Quarter
BIO17	Precipitation of Driest Quarter
BIO18	Precipitation of Warmest Quarter
BIO19	Precipitation of Coldest Quarter
LULC_NIG	Land use Land cover of Nigeria
DEM_NIG	Digital Elevation Model of Nigeria (Land Surface Terrain)

Note

* denotes multiplication.

<https://doi.org/10.1371/journal.pone.0204233.t001>

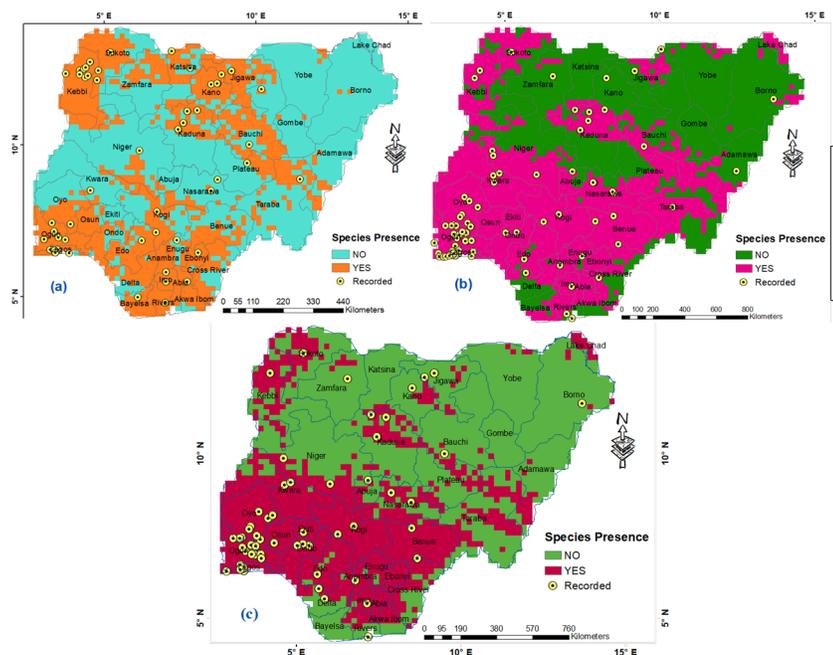


Fig 2. Suitable habitats for *Anopheles* species in Nigeria: (a) *An. gambiae* s.l., (b) *An. gambiae* s.s. and (c) *An. arabiensis*. Georeferenced *Anopheles* species locations reprinted for illustrative purposes only from Okorie *et al.* [39] under a CC BY 4.0 license, with permission from PLOS ONE.

<https://doi.org/10.1371/journal.pone.0204233.g002>

Table 2. Ecological zones suitability and distribution of *Anopheles* species.

Ecological zone	Estimated Area (km ²)	Mean Distribution Density per sq. km			Prevalence (%)		
		<i>An. gambiae</i> s.l.	<i>An. gambiae</i> s.s.	<i>An. arabiensis</i>	<i>An. gambiae</i> s.l.	<i>An. gambiae</i> s.s.	<i>An. arabiensis</i>
Sahel savanna	80,277.79	1.19	1.15	1.12	59.38	57.71	56.12
Sudan savanna	152,500.02	1.34	1.18	1.16	67.21	58.96	58.08
Northern Guinea savanna	91,944.45	1.24	1.20	1.17	61.96	60.22	58.57
Southern Guinea savanna	109,444.46	1.20	1.24	1.20	60.06	61.88	59.82
Mid Altitude	25,000.00	1.17	1.18	1.16	58.45	59.02	57.90
Derived savanna	204,166.69	1.30	1.39	1.36	65.18	69.26	68.24
Humid forest	85,277.79	1.43	1.35	1.36	71.47	67.59	67.79

Note: distribution density = 1km⁻² is equivalent to prevalence = 50% (unsuitable zone and species absence, designated with a green square); >1km⁻² ≡ >50% (suitable zone and species presence); and 2km⁻² ≡ 100% (highly suitable zone with maximum prevalence of species, designated with a red square).

<https://doi.org/10.1371/journal.pone.0204233.t002>

less suitable with prevalence between 57% and 59% (Table 2), just as parts of Mangrove and Fresh water swamp forests within the Humid forest, especially along the deltas within Delta, Bayelsa, Rivers and Akwa Ibom states (Fig 2). Also, highlands within Niger, Kwara, Oyo, Ondo, Ekiti, Edo, Kogi, Enugu and Anambra states appear less suitable for these species especially *An. gambiae* s.l. Moreover, Sudan savanna and parts of Northern Guinea savanna in the North Western region are highly suitable for the occurrence of *An. gambiae* s.l. (Table 2), while North Eastern region landmass seems less suitable (Fig 2A). Sokoto and Kebbi states in the North West within Sudan savanna are highly suitable for the occurrence of *An. gambiae* s. s. and *An. arabiensis*, just as Kaduna South / Kaduna Central and parts of Jigawa, Kano and Zamfara states (with large less suitable landmass). Similar to *An. gambiae* s.l., landmass in Sahel, Sudan, Northern and Southern Guinea savannas within North Eastern region appear less suitable for *An. gambiae* s.s. and *An. arabiensis* (Fig 2; Table 2). However, areas within Lake Chad, Bama and Kala/Balge districts of Borno state appear suitable for *An. gambiae* s.s. and *An. arabiensis*, just as Yobe state boundary with Niger Republic, and considerable landmass from Taraba, Adamawa, Gombe, through Plateau to Bauchi state (Fig 2).

With respect to their environmental suitability, *An. gambiae* s.l., *An. gambiae* s.s. and *An. arabiensis* are widespread in Humid forest, Derived savanna; less distributed in Northern and Southern Guinea savannas, and least distributed in Sahel savanna and Mid Altitude zones (Table 2; Fig 3). While *An. gambiae* s.l. is widespread in Sudan savanna, *An. gambiae* s.s. and *An. arabiensis* record high presence only within west of Sudan savanna (Table 2; Fig 3). Unlike *An. gambiae* s.l., *An. gambiae* s.s. and *An. arabiensis* with similar distribution pattern are widespread within Lake Chad region of Sahel savanna, but record limited presence in Fresh water and Mangrove swamp forests of the Niger Delta region (Fig 3). In terms of regional zones, *An. gambiae* s.l. is more widespread in the South East, followed by South West and South South (Table 3); with highest mean distribution density in Lagos followed by Ogun and Abia States, while Yobe state records the lowest among all states (Figs 3A and 4). North East has the lowest mean distribution density of *An. gambiae* s.l. (Table 3). The boundary line on the mean distribution density graph defines presence and absence condition for each *Anopheles* species from zonal statistics (Fig 4). *An. gambiae* s.s. is highly prevalent in South Western and South Eastern regions of Nigeria (Table 3; Fig 4). It is more widespread in North Central region than South South and North West, and lowest in North East (Table 3). As a dominant *Anopheles* species [6,7], *An. arabiensis* exists in all states in Nigeria, highest in Lagos state and lowest in Bayelsa state (Fig 4). South West records highest prevalence of *An. arabiensis*, followed by South East, South South, North Central, North West, and lowest in North East (Table 3).

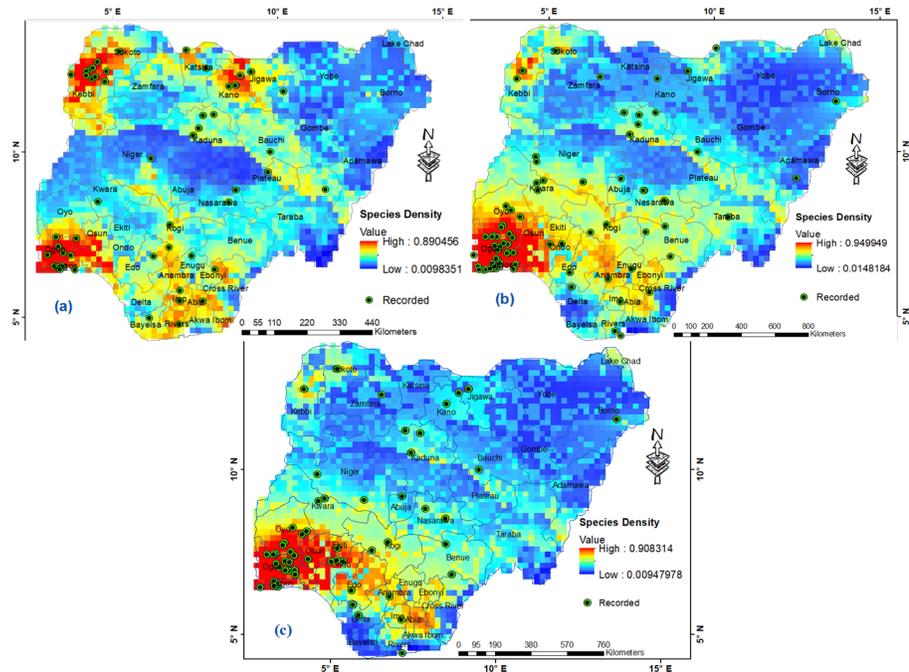


Fig 3. Potential distribution of *Anopheles* species in Nigeria: (a) *An. gambiae* s.l., (b) *An. gambiae* s.s. and (c) *An. arabiensis*. Georeferenced *Anopheles* species locations reprinted for illustrative purposes only from Okorie *et al.* [39] under a CC BY 4.0 license, with permission from PLOS ONE.

<https://doi.org/10.1371/journal.pone.0204233.g003>

Variability exists in the spatial distributions of the *Anopheles* species (Fig 3) among states (Fig 4) across topographic relief, ecological and geopolitical regions based on the diversity in climate conditions across the country, which affects the spatial epidemiology of these *Anopheles* species and malaria transmission [24]. Matching the pattern of malaria parasite prevalence [25], distribution density of the *Anopheles* species increases from the sub-tropical Middle Belt region to the tropical southern regions with high rainfall and coastal plains (Fig 3). Especially South West and South East where Lagos state, smallest by landmass but the most populated, and most urbanised [39,49] records highest prevalence of *An. gambiae* s.l., *An. gambiae* s.s. and *An. arabiensis*, matching about 1.2 million confirmed cases of malaria in 2016 [50].

Table 3. *Anopheles* species prevalence by regions under current climate.

Geopolitical zone	Estimated Area (km ²)	Mean Distribution Density per sq. km.			Prevalence (%)		
		<i>An. gambiae</i> s.l.	<i>An. gambiae</i> s.s.	<i>An. arabiensis</i>	<i>An. gambiae</i> s.l.	<i>An. gambiae</i> s.s.	<i>An. arabiensis</i>
South South	68,888.90	1.38	1.29	1.29	69.10	64.46	64.65
South East	23,611.11	1.44	1.42	1.41	71.94	71.12	70.75
South West	61,666.67	1.43	1.57	1.59	71.29	78.28	79.70
North Central	186,388.91	1.24	1.30	1.25	61.75	64.85	62.51
North East	227,500.02	1.20	1.17	1.14	59.79	58.48	57.05
North West	177,222.24	1.34	1.20	1.18	67.16	60.07	58.95

Note: distribution density = 1km⁻² is equivalent to prevalence = 50% (species absence, designated with a green square); >1km⁻² ≡ >50% (species presence); and 2km⁻² ≡ 100% (maximum prevalence of species, designated with a red square).

<https://doi.org/10.1371/journal.pone.0204233.t003>

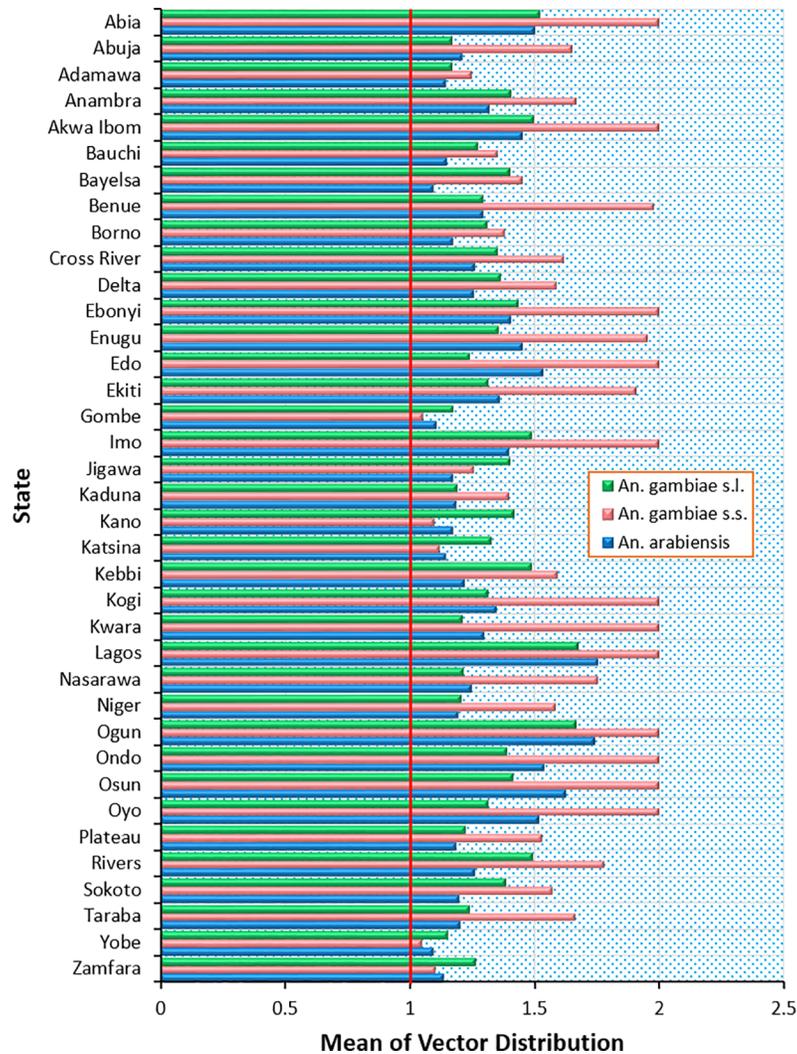


Fig 4. Mean distribution density of *An. gambiae* s.l., *An. gambiae* s.s., and *An. arabiensis* in Each Nigerian State. Note: if 1 = 0 (species do not occur); then >1 = species occur; and 2 = maximum species prevalence.

<https://doi.org/10.1371/journal.pone.0204233.g004>

Environmental variables contributions in defining *Anopheles gambiae* species distributions

Mean temperature of wettest quarter (bio_8) is the environmental variable with highest gain to the MaxEnt model of *An. gambiae* s.l. when used in isolation, which therefore appears to have the most useful information by itself (Fig 5A). Bio_8 demonstrates how mean temperatures (28°C minimum and 30°C maximum) [51] during the wettest three months (June—August) of the year may affect seasonal distributions of *An. gambiae* s.l. Other environmental variables that influence the occurrence and distribution of *An. gambiae* s.l. when used together with all other environmental variables are minimum temperature of coldest month (bio_6) (23°C, August), precipitation of coldest quarter (bio_19) (June—August, about 211mm in the arid north to above 2000mm in the coastal south), annual mean temperature (bio_1) 33°C, and precipitation of driest quarter (bio_17) (December—February, 0mm in arid north to 240mm in coastal south) (Fig 5A) [52]. Also, mean diurnal range (bio_2) (7–16°C) [52] is the major environmental variable which defines suitable habitats for *An. gambiae* s.s. in isolation, and

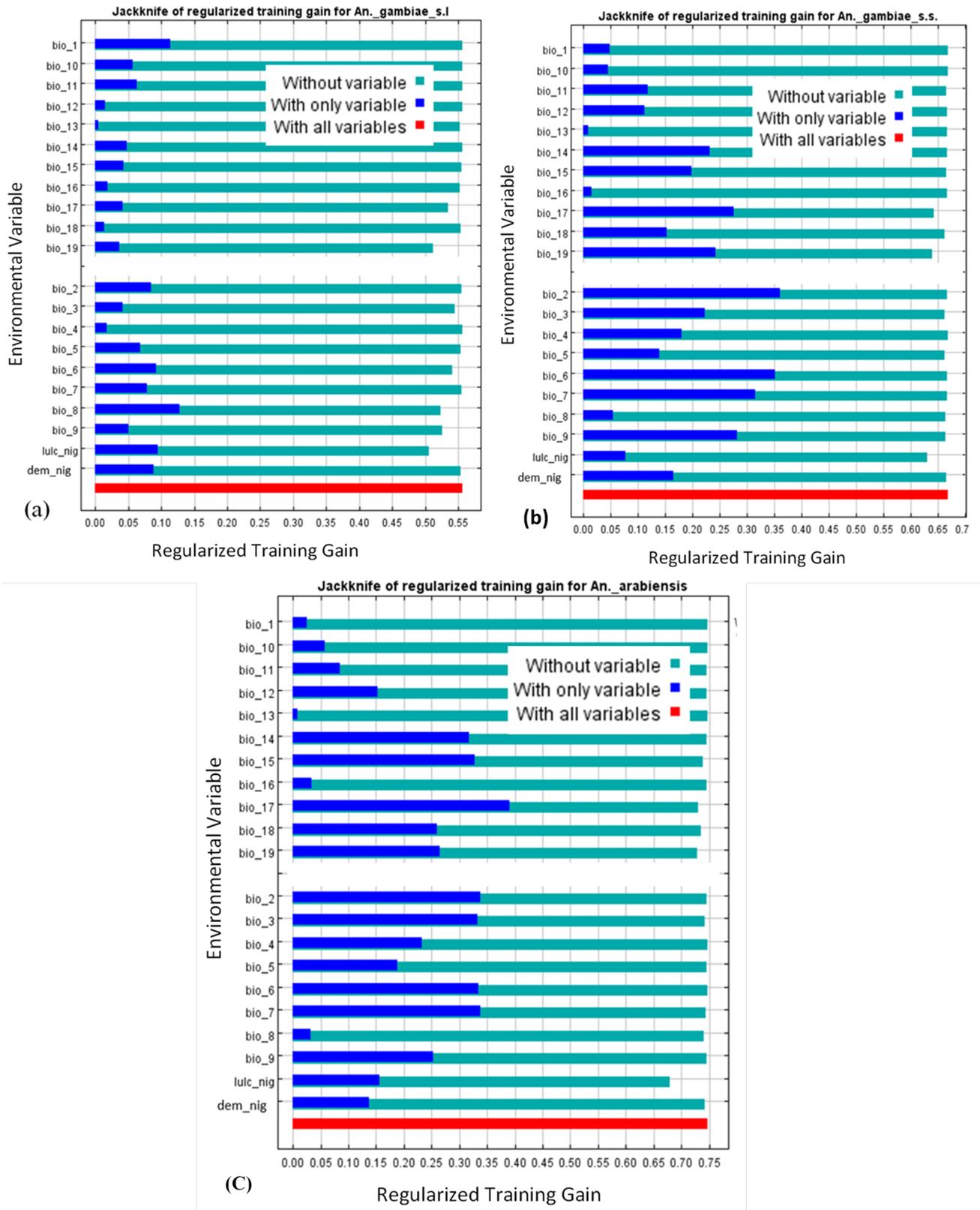


Fig 5. Jackknife test for (a) *An. gambiae* s.l. (b) *An. gambiae* s.s. (c) *An. Arabiensis*.

<https://doi.org/10.1371/journal.pone.0204233.g005>

constitutes the relevance of temperature fluctuation on spatial distribution of *An. gambiae* s.s. Other variables which appear very pivotal in the occurrence and distribution of *An. gambiae* s.s. when used alongside all other environmental variables are minimum temperature of coldest month (bio_6), temperature annual range (bio_7) (28–36°C), mean temperature of driest quarter (bio_9) (20°C minimum and 35°C maximum), precipitation of driest quarter (bio_17), precipitation of coldest quarter (bio_19) and precipitation of driest month (bio_14) (January, not exceeding 27mm in the wet southern coast) (Fig 5B) [51,52]. Precipitation of driest quarter (bio_17) is the major environmental variable that influences the occurrence and seasonal distributions of *An. arabiensis* in isolation (Fig 5C). Precipitation of coldest quarter (bio_19), precipitation seasonality (bio_15), mean diurnal range (bio_2), mean temperature of driest quarter (bio_9) and precipitation of warmest quarter (bio_18) (March to May, about 10mm in the arid north to 700mm in the coastal south) combine with all other environmental variables to delineate suitable habitats for the occurrence and distribution of *An. arabiensis* (Fig 5C). However, land use land cover map reflecting high urbanisation, increased population density and anthropogenic activities is the environmental variable that would decrease the gain the most for all three *Anopheles* species if omitted. It has the most information that is not present in the other variables. The critical influence of land use land cover dynamics in the occurrence and distribution dynamics of *An. gambiae* s.l., *An. gambiae* s.s. and *An. arabiensis* is expressed in their high distribution density in highly populated/urbanised states with increased anthropogenic activities including Abia, Akwa Ibom, Anambra, Enugu, Imo, Kano, Lagos, Ogun, Ondo, Osun, Oyo, Rivers and Sokoto states (Fig 4) [13].

MaxEnt model performance

MaxEnt recorded a fair performance for 21 replicate runs of *An. gambiae* s.l., *An. gambiae* s.s., and *An. arabiensis* with average test AUC of 0.713, 0.699, and 0.713 respectively. The value of AUC determines the performance of the model; AUC of 0.5 implies that the model was no better than random, while an AUC of 1 indicates a perfect prediction. In essence, AUC values tend to be higher for species with narrow ranges relative to the study area described by the environmental data. A behaviour that is an artifact of the AUC statistic, but does not necessarily mean that the models are better [46].

Discussion

In agreement with previous studies [4,17,23,39], the model results suggested that *An. gambiae* s.l., *An. gambiae* s.s., and *An. arabiensis* are widespread across all ecological zones in Nigeria, where they co-exist in sympatric relationship [24]. The combinations of soil, landform and climatic characteristics within ecological zones define distinct distribution of the modelled *Anopheles* species [36], and they predominantly occur in Humid forest, Guinea savannas and Sudan savanna regions [4,39,53–56]. The high environmental suitability of the Derived savanna and Humid forest within southern and parts of North Central regions [28,39] is influenced by human settlement patterns, topographical and climatic conditions of the regions [11,28,57]. Total annual precipitation, random temporal fluctuations, climate seasonality and land use land cover dynamics strongly influenced the range, relative abundance and ecological adaptability of the dominant members of the *An. gambiae* complex, in line with previous findings [4,16,28,30]. The highest mean distribution density of *An. gambiae* s.s. amongst other species corroborates with the results of Bruce-Chwatt [17] and Okwa *et al.* [18] who reported *An. gambiae* s.s. as the most efficient and most widespread within the *gambiae* complex [58]. Its high abundance is highly associated with the mean diurnal temperature range that increases the species sensitivity to changes in climates, leading to widespread presence of the species

[52,59]. In accordance with Oyewole *et al.* [19], the combined contributions of environmental variables favour higher distribution of *An. gambiae* s.s. in wet season than *An. arabiensis*, while precipitation during dry and warm months (high humidity in dry season) favour higher distribution of *An. arabiensis* than *An. gambiae* s.s. in dry season [20,21,30]. *An. arabiensis* preference of warmer climates [19,30] possibly impacts its limited presence in the cold swamps within the Humid forest, and the high suitability of Sahel savanna localities, especially Lake Chad basin area of Borno state enhanced by high relative humidity from the lake [53].

The low suitability of areas within Mid Altitude zone may be attributed to the comparatively cold weather of the highlands associated with average monthly temperatures 21–25°C [14,34], relatively below the model optimum temperatures of 23–35°C for rapid population expansion of the *An. gambiae* species [12]. As derived from the model, the high suitability of Lagos state for the malaria vector species can also be attributed to high temperature and precipitation in that area of the Humid forest [12]. The complex nature of the society, poor planning and lack of infrastructure in expanding slum areas, rapid population expansion and industrial activities make the climate warmer and create conditions highly suitable for malaria vector reproduction, survival and increased biting rates, exacerbating malaria transmission in Lagos state [11,14,28]. The influence of seasonal rainfall variability and high tropical temperatures on the extent and unbalanced distribution of the modelled mosquito species observed in most part of the country also agrees with the findings of Oyewole *et al.* [19] and Umar *et al.* [22], in relation to unbalanced and seasonal malaria transmission. The observed gradient in the distribution density from coastal south to arid north, shows that vector abundance is greatest in areas with consistently high temperatures and in any case, small mean diurnal temperature range and consistent precipitation. This is in line with the observation of Dimitrov and Morton [52] who reported that entomological inoculation rate was highest in the coastal areas and lowest in the northeast.

MaxEnt performance was better than random [45] with AUC values less than those obtained in similar studies [28,30]. This may be attributable to large ranges of the documented species [46] relative to the study area (especially in the North Eastern part of the country), resulting in increased sampling bias [44,60,61], which may influence the model performance [46,47]. However, according to Lobo *et al.* [62], an accurate model for widespread species (just as the ones modelled in this study) where the probability of presence increases steadily with predictor values have low AUC values, denoting the true generalist nature of the species distribution. It is important to note that suitable areas with low distribution density of the modelled *Anopheles gambiae* species may likely experience widespread prevalence with high distribution density, species migration and invasion [63], if there is a change in any of the environmental variables identified in this study as crucial to their distribution pattern. This will lend credence to the prediction of ecological models, that the distribution of world biomes is likely to shift as a result of changes in climate system associated with increased warming [64], since *An. gambiae* s.l., *An. gambiae* s.s., and *An. arabiensis* highly flourish with warm climate [12,65]. Thus, the propensity of future malaria transmission in Nigeria is expected to be higher with seasonal spatial shifts due to climate change and altered weather patterns; influencing the range (both latitude and altitude), intensity, and seasonality of vectors [11–15,65,66].

Conclusions

In this paper we used Maxent in modelling environmental suitability and distribution of dominant *Anopheles gambiae* species in Nigeria. We also assessed the contributions and importance of bioclimatic and other environmental variables to the model. Results showed that the species are more prevalent within the Humid forest and the Derived savanna, but most prevalent in

South Western and South Eastern geopolitical zones within the two ecological zones. This is particularly worrisome in highly populated and urbanised Lagos state which recorded highest distribution density of all three species. Our results also showed that land use dynamics become very critical for the occurrence and distribution of the three dominant species of *Anopheles*, while seasonal rainfall, temperature fluctuations and high humidity during warm weather (dry season) drive the occurrence and seasonal distribution of the *Anopheles* species and potential malaria transmission. The derived MaxEnt model was successful in defining potential suitable habitats for the occurrence and distribution of the *Anopheles* species, and estimated variable importance. This result might be useful in predicting the variability of malaria vector distribution across ecological gradients and in understanding the potential causes of its severity from an environmental point of view in tropical regions.

Supporting information

S1 Table. Nigeria *Anopheles* vector database. Reused from Okorie *et al.* [39] under a CC BY 4.0 license, with permission from PLOS ONE. (XLSX)

Author Contributions

Conceptualization: Godwin E. Akpan, Kayode A. Adepoju.

Data curation: Godwin E. Akpan.

Formal analysis: Godwin E. Akpan, Kayode A. Adepoju.

Investigation: Godwin E. Akpan, Kayode A. Adepoju.

Methodology: Godwin E. Akpan, Kayode A. Adepoju, Olakunle R. Oladosu, Samuel A. Adelabu.

Project administration: Kayode A. Adepoju.

Supervision: Kayode A. Adepoju, Olakunle R. Oladosu.

Validation: Godwin E. Akpan, Kayode A. Adepoju.

Visualization: Godwin E. Akpan.

Writing – original draft: Godwin E. Akpan, Kayode A. Adepoju, Olakunle R. Oladosu.

Writing – review & editing: Godwin E. Akpan, Kayode A. Adepoju, Olakunle R. Oladosu, Samuel A. Adelabu.

References

1. World Health Organization. World Malaria Report 2015. Geneva: World Health Organization; 2015. 280p.
2. World Health Organization. World Malaria Report 2017. Geneva: World Health Organization; 2017. 196p.
3. CDC. Anopheles Mosquitoes. Centers for Disease Control and Prevention: Global Health. Division of Parasitic Diseases and Malaria 2015, October 21. Available from: <http://www.cdc.gov/malaria/about/biology/mosquitoes/>
4. Onyabe DY, Conn JE. The Distribution of Two Major Malaria Vectors, *Anopheles gambiae* and *Anopheles arabiensis*, in Nigeria. Mem Inst Oswaldo Cruz, Rio de Janeiro. 2001; 96(8):1081–1084.
5. Sinka ME, Bangs MJ, Manguin S, Coetzee M, Mbogo CM, Hemingway J, et al. The dominant *Anopheles* vectors of human malaria in Africa, Europe and the Middle East: occurrence data, distribution maps

- and bionomic précis. *Parasites and Vectors*. 2010; 3(117). <https://doi.org/10.1186/1756-3305-3-117> PMID: 21129198
6. Takken W, Scott TW. *Ecological Aspects for Application of Genetically Modified Mosquitoes*. Dordrecht. Kluwer Academic Publishers. 2003; 75–90.
 7. Autino B, Noris A, Russo R, Castelli F. Epidemiology of Malaria in Endemic Areas. *Mediterr J Hematol Infect Dis*. 2012; 4(1):e2012060. <https://doi.org/10.4084/MJHID.2012.060> PMID: 23170189
 8. Gillies MT, Coetzee M. A supplement to the Anophelinae of Africa South of the Sahara (Afrotropical Region). In Publications of the South African Institute for Medical research Johannesburg. 1987. p. 55.
 9. Gillies MT, De Meillon B. The Anophelinae of Africa south of the Sahara (Ethiopian Zoogeographical Region). In Publications of the South African Institute for Medical Research Johannesburg. 1968. p. 54.
 10. Adepoju KA, Akpan GE. Historical Assessment of Malaria Hazard and Mortality in Nigeria—Cases and Deaths: 1955–2015. *Int J Environ. Bioener*. 2017; 12(1): 30–46.
 11. Tolulope O. Spatio–Temporal Clustering of Malaria Morbidity in Nigeria (2004–2008). *J Sci Res*. 2014; 13:99–113.
 12. Martens WJ, Niessen LW, Rotmans J, Jetten TH, McMichael AJ. Potential Impact of Global Climate Change on Malaria Risk. *Environ Health Perspect*. 1995; 103(5): 458–464. <https://doi.org/10.1289/ehp.95103458> PMID: 7656875
 13. Khasnis AA, Nettleman MD. REVIEW ARTICLE: Global Warming and Infectious Disease. *Arch Med Res*. 2005; 36: 689–696. <https://doi.org/10.1016/j.arcmed.2005.03.041> PMID: 16216650
 14. Pascual M, Ahumada JA, Chaves LF, Rodo X, Bouma M. Malaria resurgence in the East African highlands: temperature trends revisited. *Proceedings of the National Academy of Sciences of the United States of America*. 2006; 103(15): 5829–5834. <https://doi.org/10.1073/pnas.0508929103> PMID: 16571662
 15. Ostfeld RS. Climate change and the distribution and intensity of infectious diseases. *Ecol*. 2009; 90(4): 903–905.
 16. Lindsay SW, Parson L, Thomas CJ. Mapping the ranges and relative abundance of the two principal African malaria vectors, *Anopheles gambiae* sensu stricto and *An. arabiensis*, using climate data. *Proc R Soc London B*. 1998; 265:847–854.
 17. Bruce-Chwatt LJ. Malaria in Nigeria. *Bull. World Hlth Org*. 1951; 4:301–327.
 18. Okwa OO, Akinmolayan FI, Carter V, Hurd H. Transmission dynamics of malaria in four selected ecological zones of Nigeria in the rainy season. *Ann Afr Med*, 2009; 8(1):1–9. [10.4103/1596-3519.55756](https://doi.org/10.4103/1596-3519.55756) PMID: 19762999
 19. Oyewole IO, Ibadapo CA, Oduola AO, Obansa JB, Awolola TS. Anthropophilic mosquitoes and malaria transmission in a tropical rain forest area of Nigeria. *acta SATECH*. 2005; 2(1):6–10.
 20. Awolola TS, Oyewole IO, Koekemoer LL, Coetzee M. Identification of three members of *Anopheles funestus* (Diptera: Culicidae) group and their role in malaria transmission in two ecological zones in Nigeria. *Roy Soc Trop Med Hyg*. 2005; 99:525–531.
 21. Amadi EC, Harry SI, Ogbalu OK. Preliminary Survey on The Abundance of Mosquito Species and Transmission of Plasmodium to Humans in Bille, Oporoama, Sama and Tema Communities of Rivers State, Nigeria. *IOSR J. of Phar. and Bio. Sci*. 2015; 10(6 Ver. II):82–86.
 22. Umar A, Kabir BG, Abdullahi MB, Barde A, Misau AA, Sambo ML, et al. Assessment of Indoor Resting Density of Female Anopheline Mosquitoes in Human Dwelling at Malaria Vector Sentinel Sites in Bauchi State, Nigeria. *Adv Stud Biol*. 2015; 7(7):323–333. Available from: <http://dx.doi.org/10.12988/asb.2015.5210>
 23. Okorie PN, Popoola K, Awobifa OM, Ibrahim KT, Ademowo GO. Species composition and temporal distribution of mosquito populations in Ibadan, Southwest Nigeria. *J Entomol Zoo Stud* 2014; 2(4):164–169.
 24. National Malaria Control Programme, suNMaP, World Health Organization and the INFORM Project. A description of the epidemiology of malaria to guide the planning of control in Nigeria. A report prepared for the Federal Ministry of Health, Nigeria, the Roll Back Malaria Partnership and the Department for International Development, UK. November, 2013.
 25. World Health Organization. *World Malaria Report 2016*. Geneva: World Health Organization; 2016. 186p.
 26. Dalitz MK. *Autochthone Malaria im mitteleuropäischen Raum*. Dissertation, Martin-LutherUniversität, Halle-Wittenberg Universität. Halle-Wittenberg. 2005.
 27. Siettos CI, Russo L. Mathematical modeling of infectious disease dynamics. *Virulence*. 2013; 4(4):295–306. <https://doi.org/10.4161/viru.24041> PMID: 23552814

28. Moffett A, Shackelford N, Sarkar S. Malaria in Africa: Vector Species' Niche Models and Relative Risk Maps. PLoS ONE. 2007; 2(9):e824. <https://doi.org/10.1371/journal.pone.0000824> PMID: 17786196
29. Foley DH, Klein TA, Kim HC, Sames WJ, Wilkerson RC, Rueda LM. Geographic Distribution and Ecology of Potential Malaria Vectors in the Republic of Korea. J Med Entomol. 2009; 46(3):680–692. PMID: 19496442
30. Kulkarni MA, Desrochers RE, Kerr JT. High Resolution Niche Models of Malaria Vectors in Northern Tanzania: A New Capacity to Predict Malaria Risk? PLoS ONE. 2010; 5(2): p. e9396. <https://doi.org/10.1371/journal.pone.0009396> PMID: 20195366
31. Peterson AT. Ecologic Niche Modeling and Spatial Patterns of Disease Transmission. Emerg Infect Dis. 2006; 12(12): p1822–1826.
32. Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ. A statistical explanation of MaxEnt for ecologists. Diversity Distrib. 2011; 17:43–57.
33. Amadi SO, Udo SO, Ewona IO. Trends and Variations of Monthly Mean Minimum and Maximum Temperature Data over Nigeria for the Period 1950–2012. IJPAP. 2014; 2(4):1–27.
34. Metz HC (Ed.). Nigeria: A Country Study. Washington: GPO for the Library of Congress. 1991. Available from: <http://countrystudies.us/nigeria/33.htm>
35. USAID. Nigeria Environmental Analysis Final Report. Biodiversity and Sustainable Forestry (BIOFOR) Indefinite Quantity Contract (IQC); 2002. USAID BIOFOR IQC No.: LAG-I-00-99-00013-00.
36. FAO Soils Bulletin 73. Agro-Ecological Zoning Guidelines. Soil Resources, Management and Conservation Service. FAO Land and Water Development Division. Food and Agriculture Organization of the United Nations, Rome. 1996. Available from: http://www.fao.org/docrep/w2962e/w2962e-03.htm#P229_15540
37. United Nations, Department of Economic and Social Affairs, Population Division (UN-DESA). World Population Prospects: The 2017 Revision, Key Findings and Advance Tables. 2017; Working Paper No. ESA/P/WP/248.
38. Owoseye A. World Malaria Day: Nigerians Warned to Stop Using Chloroquine for Malaria. Treatment. Premium Times. 25 Apr 2017. Available from: <http://allafrica.com/stories/201704260043.html> Cited 9 Nov 2017.
39. Okorie PN, McKenzie FE, Ademowo OG, Bockarie M, Kelly-Hope L. Nigeria *Anopheles* Vector Database: An Overview of 100 Years' Research. PLoS ONE. 2011; 6(12):e28347. <https://doi.org/10.1371/journal.pone.0028347> PMID: 22162764
40. WorldClim—Global Climate Data: Free climate data for ecological modelling and GIS. Download. Available from: <http://www.worldclim.org/download.html>.
41. Tappan GG, Cushing WM, Cotillon SE, Mathis ML, Hutchinson JA, Dalsted KJ. West Africa Land Use Land Cover Time Series: U.S. Geological Survey data release. 2016. Available from: <http://dx.doi.org/10.5066/F73N21JF>
42. CGIAR-CSI. SRTM 90m Digital Elevation Data. The CGIAR Consortium for Spatial Information (CGIAR-CSI). Jan 2017. Available from: <http://srtm.csi.cgiar.org/SELECTION/inputCoord2.asp>
43. Phillips SJ, Anderson RP, Schapire RE. Maximum entropy modeling of species geographic distributions. Ecol Model. 2006; 190: 231–259. Available from: <https://doi.org/10.1016/j.ecolmodel.2005.03.026>
44. Fourcade Y, Engler JO, Rödder D, Secondi J. Mapping Species Distributions with MAXENT Using a Geographically Biased Sample of Presence Data: A Performance Assessment of Methods for Correcting Sampling Bias. PLoS ONE. 2014; 9(5): e97122. <https://doi.org/10.1371/journal.pone.0097122> PMID: 24818607
45. Altamiranda-Saavedra M, Arboleda S, Parra JL, Peterson AT, Correa MM. Potential distribution of mosquito vector species in a primary malaria endemic region of Colombia. PLoS ONE. 2017; 12(6): e0179093. Available from: <https://doi.org/10.1371/journal.pone.0179093> PMID: 28594942
46. Phillips SJ. A Brief Tutorial on Maxent. 2017. Available from: http://biodiversityinformatics.amnh.org/open_source/maxent/
47. Young N, Carter L, Evangelista P. A MaxEnt Model v3.3.3e Tutorial (ArcGIS v10). Natural Resource Ecology, Colorado State University, USA. 2011.
48. Hanley JA, McNeil BJ. The meaning and use of the area under a Receiver. Operating Characteristic (ROC) curve. Radiology, 1982; 143:29–36. <https://doi.org/10.1148/radiology.143.1.7063747> PMID: 7063747
49. Sunday O, Ajewole A. Implications of the Changing Pattern of Land cover of the Lagos Coastal Area of Nigerian. AEJSR, 2006; (1):31–37. IDOSI Publications. Available from: [www.idosi.org/aejsr/1\(1\)06/7.pdf](http://www.idosi.org/aejsr/1(1)06/7.pdf)

50. Vanguard. Malaria: Lagos State records 1.2m cases in 2016 –Commissioner. Vanguard Media Limited, Nigeria. 4 May 2017. Available from: <https://www.vanguardngr.com/2017/05/malaria-lagos-state-records-1-2m-cases-2016-commissioner/> Cited 10 Nov 2017.
51. World Guides. Nigeria Weather, When to Go and Climate Information (Nigeria, NG, West Africa). TravelSmart Ltd. 2016. Available from: http://www.world-guides.com/africa/west-africa/nigeria/nigeria_weather.html Cited 7 June 2018.
52. Dimitrov NB, Morton DP. Combinatorial Design of a Stochastic Markov Decision Process. In: Saltzman MJ, Chinneck JW, Kristjansson B, editors. Operations Research and Cyber-Infrastructure. New York: Springer; 2009. P. 167–193. <https://doi.org/10.1007/978-0-387-88843-9>
53. Coluzzi M, Sabatini A, Petrarca V, Di Deco MA. Chromosomal differentiation and adaptation to human environments in the *Anopheles gambiae* complex. *Trans R Soc Trop Med Hyg.* 1979; 73:483–497. PMID: [394408](https://pubmed.ncbi.nlm.nih.gov/394408/)
54. Amajoh CN. A review of Malaria Vector behaviour in Nigeria. Abstracts of two days National symposium on malaria in Nigeria held at Nigerian, 4th to 5th November 1997. Yaba, Lagos: Institute of Medical Research. 1997.
55. Coetzee M, Craig M, le Sueur D. Distribution of African malaria mosquitoes belonging to the *Anopheles gambiae* complex. *Parasitol Today.* 2000; 16:74–77. PMID: [10652493](https://pubmed.ncbi.nlm.nih.gov/10652493/)
56. Onyabe DY, Vajime CG, Nock IH, Ndams IS, Akpa AU, Alaribe AA, et al. The distribution of M and S molecular forms of *Anopheles gambiae* in Nigeria. *Trans R Soc Trop Med Hyg.* 2003; 97:605–608. PMID: [15307438](https://pubmed.ncbi.nlm.nih.gov/15307438/)
57. Siteti MC, Injete SD, Wanyonyi WA. Malaria parasite species prevalence and transmission dynamics at selected sites in the Western highlands of Kenya. *CHRISMED J Health Res.* 2016; 3:45–50.
58. Olayemi IK, Ande AT, Odeyemi MO, Ibemesi G Emmanuel R. Temporal Ecologic Adaptability of the Principal Vector of Malaria, *Anopheles gambiae* s.l. (Diptera: Culicidae), in Northcentral Nigeria. *App Sci Report.* 2014; 5(3):110–117. <https://doi.org/10.15192/PSCP.ASR.2014.1.3.110117>
59. Briga M, Verhulst S. Large diurnal temperature range increases bird sensitivity to climate change. *Sci Report.* 2015; 5:16600. <https://doi.org/10.1038/srep16600> PMID: [26563993](https://pubmed.ncbi.nlm.nih.gov/26563993/)
60. Syfert MM, Smith MJ, Coomes DA. The Effects of Sampling Bias and Model Complexity on the Predictive Performance of MaxEnt Species Distribution Models. *PLoS ONE.* 2013; 8(2): e55158. <https://doi.org/10.1371/journal.pone.0055158> PMID: [23457462](https://pubmed.ncbi.nlm.nih.gov/23457462/)
61. Phillips SJ, Dudík M. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography.* 2008; 31:161–175.
62. Lobo JM, Jiménez-Valverde A, Real R. AUC: a misleading measure of the performance of predictive distribution models. *Global Ecol. Biogeogr.* 2007. <https://doi.org/10.1111/j.1466-8238.2007.00358.x>
63. Peterson AT. Shifting suitability for malaria vectors across Africa with warming climates. *BMC Infect Dis.* 2009; 9(59). <https://doi.org/10.1186/1471-2334-9-59> PMID: [19426558](https://pubmed.ncbi.nlm.nih.gov/19426558/)
64. IPCC. The Regional Impacts of Climate Change: An Assessment of Vulnerability, (Eds Watson R. T., Zinyowera M. C., and Moss R. H.), Cambridge University Press, Cambridge, UK. 1998.
65. IPCC. Climate Change 1995 Impacts, Adaptations and Mitigation of Climate Change: Scientific-Technical Analyses. Contribution of Working Group II to the Second Assessment Report of the Intergovernmental Panel on Climate Change. (RT Watson, MC Zinyowera, RH Moss, DJ Dokken, Eds.) New York, USA: Press Syndicate of the University of Cambridge. 1995.
66. Rogers DJ, Randolph SE. The global spread of malaria in a future, warmer world. *Science.* 2000; 289:1763–5. PMID: [10976072](https://pubmed.ncbi.nlm.nih.gov/10976072/)