Scientific Research Publishing

Open Journal of Statistics, 2017, 7, 901-919
http://www.scirp.org/journal/ojs
ISSN Online: 2161-7198
ISSN Print: 2161-718X

Statistical Modeling of Malaria Incidences in Apac District, Uganda

Ayo Eunice[1*], Anthony Wanjoya[1,2], Livingstone Luboobi[1,3]

[1]Pan African University Institute of Basic Sciences, Technology and Innovation, Nairobi, Kenya
[2]Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya
[3]Strathmore University, Nairobi, Kenya
Email: *ayoodongo@gmail.com

How to cite this paper: Eunice, A., Wanjoya, A. and Luboobi, L. (2017) Statistical Modeling of Malaria Incidences in Apac District, Uganda. *Open Journal of Statistics*, **7**, 901-919.
https://doi.org/10.4236/ojs.2017.76063

Received: October 3, 2017
Accepted: November 13, 2017
Published: November 16, 2017

Abstract

Malaria is a major cause of morbidity and mortality in Apac district, Northern Uganda. Hence, the study aimed to model malaria incidences with respect to climate variables for the period 2007 to 2016 in Apac district. Data on monthly malaria incidence in Apac district for the period January 2007 to December 2016 was obtained from the Ministry of health, Uganda whereas climate data was obtained from Uganda National Meteorological Authority. Generalized linear models, Poisson and negative binomial regression models were employed to analyze the data. These models were used to fit monthly malaria incidences as a function of monthly rainfall and average temperature. Negative binomial model provided a better fit as compared to the Poisson regression model as indicated by the residual plots and residual deviances. The Pearson correlation test indicated a strong positive association between rainfall and malaria incidences. High malaria incidences were observed in the months of August, September and November. This study showed a significant association between monthly malaria incidence and climate variables that is rainfall and temperature. This study provided useful information for predicting malaria incidence and developing the future warning system. This is an important tool for policy makers to put in place effective control measures for malaria early enough.

Keywords

Malaria Incidence, Climate Variables, Poisson Regression, Negative Binomial Regression, Generalized Linear Model, Apac District

1. Introduction

Uganda is one of the Sub-Saharan African countries where malaria is still en-

demic in over 90% of the country's regions [1] [2] [3]. According to the National Malaria Control Pro gramme, malaria alone has shown to contribute to between 30% and 50% of outpatient visits, 15% - 20% of hospital admissions and 20% of hospital deaths with most of this burden found in children under 5 years and pregnant women; severe malarial anaemia is responsible for a case fatality rate of 8.25% among pediatric admissions [4].

Transmission of malaria is very complicated. It can be determined by climatic or non-climatic factors. The impact of climatic variables on malaria patterns still remains controversial. The aim of this study is to model malaria incidences in Apac district, Northern Uganda with respect to climate variables specifically rainfall and temperature.

The greatest burden of malaria, remains in the heart land of Africa, characterized by limited infrastructure to monitor disease trends, large contiguous areas of high transmission, and low coverage of control interventions. The epidemiology of malaria varies widely in Uganda, from highland regions with low prevalence and unstable disease to large regions with dense agricultural settlement and some of the highest recorded malaria intensities in the world [5].

The climate in Uganda allows stable, year round malaria transmission with relatively little seasonal variability in most areas. Malaria is highly endemic in Uganda with some of the highest recorded entomological inoculation rates (EIR, infective mosquito bites per person per year) in the world, including rates of 1586 in Apac district and 562 in Tororo district measured in 2001 to 2002 [5].

Malaria has historically been a very serious health problem and currently poses the most significant threat to the health of the people in malaria prone areas. Uganda show that more than 55 percent of pediatric cases are due to malaria [6].

Malaria remains one of the leading health problems of the developing world, and Uganda bears a particularly large burden from the disease. Our understanding is limited by a lack of reliable data, but it is clear that the prevalence of malaria infection, incidence of disease, and mortality from severe malaria all remain very high. Uganda has made progress in implementing key malaria control measures, in particular distribution of insecticide impregnated bednets, indoor residual spraying of insecticides, utilization of artemisinin-based combination therapy to treat uncomplicated malaria, and provision of intermittent preventive therapy for pregnant women. However, despite enthusiasm regarding the potential for the elimination of malaria in other areas, there is no convincing evidence that the burden of malaria has decreased in Uganda in recent years. Major challenges to malaria control in Uganda include very high malaria transmission intensity, inadequate health care resources, a weak health system, inadequate understanding of malaria epidemiology and the impact of control interventions, increasing resistance of parasites to drugs and of mosquitoes to insecticides, inappropriate case management, inadequate utilization of drugs to prevent malaria, and inadequate epidemic preparedness and response. Despite these challenges, prospects for the control of malaria have improved, and with attention to un-

derlying challenges, progress toward the control of malaria in Uganda can be expected [7].

The relationship between climatic variables and malaria transmission has been reported in many countries [8]. A recent resurgence of malaria in the East African highlands involves multiple factors; climate and land use change, drug resistance, variable disease control efforts, and other socio-demographic factors [9]. Malaria is an extremely climate-sensitive disease [10] common in the tropics, but also reported in mild-to-cold climates [11].

Rainfall and temperature anomalies are widely considered to be a major driver of inter-annual variability of malaria incidence in the semi-arid areas of Africa [12], and recently recorded a warming trend in the East African Highlands that corresponded with concomitant increases in malaria incidences [13].

Based on the background study of malaria above, the impact of weather and environmental factors on dynamics of malaria has attracted considerable attention in recent years, yet uncertainties around future disease trends under environment change remain. The role of climate as a driving force for malaria incidences is still a subject of considerable attention [14]. Assessing the impact of climate variables on malaria incidences is challenging because of a high spatial climate variability and lack of a long term data series on malaria cases from different hospitals. Temperature affects the development rates and survival of malaria parasites and mosquito vectors. Rainfall influences the availability of the mosquito larvae habitats and hence a breeding ground for mosquitoes. Temperature and rainfall may have synergistic effects on the transmission of malaria. Therefore, there is need to analyze the simultaneous effects of rainfall and temperature on malaria incidences. However, the association between climate variables and malaria incidences in Apac district has not been studied.

In this study, the association between malaria incidences and climate variables are modeled using Poisson and negative binomial Regression models respectively. The significance of rainfall and temperature on the malaria incidences are determined. This knowledge is important since it gives clear understanding of malaria incidences predictors. This is necessary for the development of malaria warning systems in Apac district, Northern Uganda and hence enable effective malaria control measures to be put in place in a timely manner.

2. Review of Malaria Models

Several studies have been carried out on malaria incidence. [15] studied the trend of malaria prevalence in Minna, Nigeria, by employing Poisson and Negative binomial regression models. The results revealed that the prevalence of malaria is still on increase by 6% on monthly basis.

[16], carried out a study to identify the spatial and trends of malaria incidence in Nepal Poisson and negative binomial regression models were used to fit malaria incidence rates as a function of year and location. The study showed a steady decreasing trend in malaria incidence, but the numbers of cases are still very

high.

[17], carried out a study using logistic regression to estimate and assess malaria prevalence and the use of malaria risk reduction measures and their association with selected background characteristics in South Sudan. The results suggest that educational attainment need not be very advanced to affect practices of malaria prevention and treatment. Primary school attendance was a stronger predictor for use of malaria risk reduction measures than any other selected background characteristics.

[18], conducted a study on malaria vector control in South Sudan. The study revealed that the peak of malaria transmission season lasting 7 to 8 months of the year south of the country and 5 to 6 months in the north.

[19], studied malaria incidence over time and its association with temperature and rainfall in four counties of Yunnan province, China. Seasonal trend decomposition was used to examine secular trends and seasonal patterns in malaria incidence, a Poisson regression with Distributed lag non-linear models were used to estimate the weather drivers of malaria seasonality. The study revealed that there was a declining trend in malaria incidence in all four counties.

[20], estimated the effects of climate factors on *P. vivax* malaria transmission using Generalized linear Poisson models and distributed lag non linear models. Their findings suggested that malaria transmission in temperate areas was highly dependent on climate factors.

[21], used Spearman's correlation between weekly climatic variables (temperatures, relative humidity and rainfall) and malaria to analyze the bi variate relationships between types of malaria parasites and potential climatic factors. A discrete poisson model was used to identify purely spatial clusters of malaria incidence in the high risk areas. A poisson regression model combined with distributed lag non-linear model was used to examine the effects of temperature, relative humidity and rainfall on the number of malaria cases. The residuals were checked to evaluate the adequacy of the model. Sensitivity analysis was performed to ensure that the associations between climate variables and malaria incidences did not change substantially when the degrees of freedom for climate variables were changed.

[22], analyzed the temporal correlation between malaria incidence and climatic variables using malaria incidence rates in Kokrajhar district of Assam over the period 2001 to 2010. Linear regressions were used to obtain linear relationships between climatic factors and malaria incidence. Temperature was found to be negatively correlated with non-forest malaria incidence while relative humidity was positively correlated with forest malaria incidence.

[23], used semi-parametric regression models to model the dependence of malaria cases on spatial determinants and climatic covariates including rainfall, temperature and humidity in Burundi. The results obtained suggested that malaria incidence in a given month is strongly associated with minimum temperature of the previous months.

[24] modeled separate meteorological factors, the model with rainfall performed better than the models with other factors respectively. The results showed that the way rainfall influenced malaria incidence was different from other factors, which could be interpreted as rainfall having a greater influence than other factors.

[25], investigated the effects of climate on malaria in Burundi using generalized linear models and generalized additive mixed models. The results suggest a strong positive association between malaria incidence in a given month and minimum temperature of the previous month. In contrast, it was found that rainfall and maximum temperature in a given month have possible negative effect on malaria incidence of the same month.

[26], presented a model for estimation of temperature effects on mortality that is able to capture jointly the typical features of every temperature death relationship, that is, nonlinearity and delayed effect of cold and heat over a few days. Using a segmented approximation along with a doubly penalized spline-based distributed lag parameterization, estimates and relevant standard errors of the cold and heat related risks and the heat tolerance are provided. The model is applied to data from Milano, Italy.

[27], used negative binomial regression model to examine how spatial distribution of the disease changes with inter annual variability of temperature. To analyze the variation in incidence with temperature and altitude, a negative binomial regression to the monthly cases was fitted. Covariates included season, altitude and linearly de-trended temperature (lagged by 3 months).

[28], found that malaria was associated with rainfall and minimum temperature in Ethiopia. Daily average number of cases was modeled using a robust Poisson regression with rainfall, minimum temperature and maximum temperatures as explanatory variables in a polynomial distributed lag model in 10 districts of Ethiopia. To improve reliability and generalizability within similar climatic conditions, the districts were grouped into two climatic zones, hot and cold. In cold districts, rainfall was associated with a delayed increase in malaria cases, while the association in the hot districts occurred at relatively shorter lags. In cold districts, minimum temperature was associated with malaria cases with a delayed effect. In hot districts, the effect of minimum temperature was non-significant at most lags, and much of its contribution was relatively immediate.

[29], compared the level of malaria infection in children from 22 communities in an area of unstable transmission in the Usambara mountains, Tanzania, immediately before and after one of the strongest recorded El Niño Southern Oscillation events. They found strikingly less malaria than in the preceding year despite 2.4 times more rainfall than normal resulted from the event.

[30], used non linear mixed-regression model to investigate the association between auto regression (number of malaria inpatients during previous time period), seasonality and climate variability, and the number of monthly malaria inpatients of the past 10 to 20 years in seven highland sites in East Africa. The

model did not take into consideration other important factors that also impact on malaria incidences for example; topography, human settlement pattern, land use, and drug resistance.

[31], examined the relationship between malaria and environmental and socio-economic variables in the Sudan using health production modified model. The regression results showed significant relationships between malaria and rainfall and water bodies. Other variables including Human Development Index, temperature, population density and percent of cultivated areas were not significant.

[32], investigated temporal associations between weekly malaria incidence in 1993 children < 15 years of age and weekly rainfall. A time series analysis was conducted by using cross-correlation function and autoregressive modeling. The regression model showed that the level of rainfall predicted the malaria incidence after a time lag of 9 weeks (mean = 60 days) and after a time lag between one and two weeks. The analyses provide evidence that high-resolution precipitation data can directly predict malaria incidence in a highly endemic area. Such models might enable the development of early warning systems and support intervention measures.

3. Methodology

In this study, data on monthly Malaria incidence in Apac district for the period January 2006 to December 2016 were obtained from the Ministry of health, Uganda. Climate data were obtained from Uganda National Meteorological Authority. The response variable is the malaria incidence where as the climate variables are the explanatory variables.

4. Data Analysis

Monthly malaria incidences for the period 2007-2016 was used. The data was obtained from the Ministry of health, Uganda. **Table 1** summarizes the Malaria incidences for the period 2007-2016.

From **Table 1**, the minimum value for malaria incidences is 5034, the maximum value is 19,289 cases, the mean value for the whole period is 9746 and the standard deviation is 3017.864.

From **Figure 1**, we observe that the histogram is skewed to the right which is representation of count data. We conclude that malaria incidences is count data and we can model it using a distribution suited for count data which is the Poisson distribution.

The generalized linear models were applied to fit the malaria incidence data as a function of rainfall and average temperature. Count data regression models

Table 1. Descriptive statistics.

Malaria incidences	Mean	Standard deviation	range
	9746	3017.864	5034 - 19,589

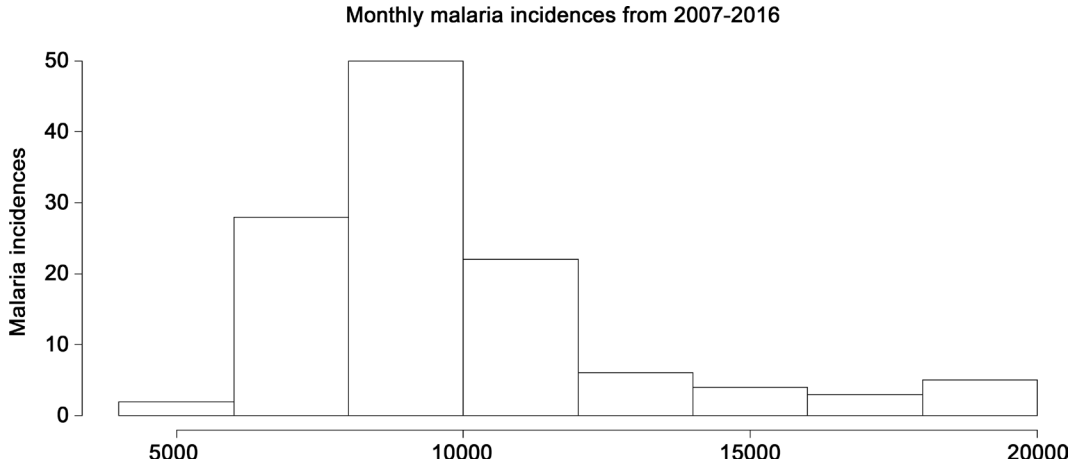

Figure 1. Monthly malaria incidences.

can be represented and understood using generalized linear models (GLM) framework [33]. Poisson regression is commonly used for modeling the number of cases of disease in a specific population within a certain time interval. The Poisson regression is a member of a class of generalized linear models, which is an extension of traditional linear models that allows the mean of a population to depend on a linear predictor through a non linear link function and allows the response probability distribution to be any member of the exponential family distributions. Poisson regression is a special case of (GLM) where the response variable follows Poisson distribution. Poisson models for disease counts are often over-dispersed hence the need for a model which appropriately handles over dispersion in which case negative binomial is considered [34]. The negative binomial model is an extension of Poisson model for incidence rates that allows for the over dispersion that commonly occurs for disease count. The Poisson probability distribution is specifically suited for count data, with density function;

$$f\left(Y_i\right)=\frac{\lambda^{y}\mathrm{e}^{-\lambda}}{y!}, y=0,1,2,\cdots \tag{1}$$

$f\left(Y\right)$ is the probability that the discrete random variable Y takes non-negative integer values, λ is the parameter of the Poisson distribution. It can be proved that;

$$E\left(Y\right)=Var\left(Y\right)=\lambda \tag{2}$$

A unique feature of Poisson distribution is that the mean is equal to the variance [35]. This is called the equidispersion property of the Poisson distribution. In the Poisson regression model, the number of events y has a Poisson distribution with conditional mean that depends on an individual's characteristics:

$$\lambda_i=E\left(y_i/x_i\right)=\exp\left(x_i\beta\right) \tag{3}$$

$$\log\left(\lambda_i\right)=x_i\beta \tag{4}$$

This is the model for analyzing count data. Under this model as λ_i increases,

the conditional variance of y increases. The Poisson regression model can be thought of as a non-linear model [36].

The Negative binomial regression model allows the conditional variance of y to exceed the conditional mean. The mean λ is replaced with the random variable $\tilde{\lambda}$:

$$\tilde{\lambda}_i = \exp\left(x_i\beta + \varepsilon_i\right) \tag{5}$$

where ε is a random error that is assumed to be uncorrelated with x.

$$\tilde{\lambda}_i = \exp\left(x_i\beta\right)\exp\left(\varepsilon_i\right) = \lambda_i \exp\left(\varepsilon_i\right) = \lambda_i\delta_i \tag{6}$$

The assumption is that δ has a gamma distribution with parameters: $E(\delta)=1$ and $Var\left(\frac{1}{v}\right)$.

The expected value of y for the Negative binomial distribution is the same as for Poisson distribution but the conditional variance differs:

$$Var\left(y_i/x_i\right) = \lambda_i\left(1+\frac{\lambda_i}{v_i}\right) = \exp\left(x_i\beta\right)\left(1+\frac{\exp\left(x_i\beta\right)}{v_i}\right) \tag{7}$$

since λ and v are positive, the conditional variance of y must exceed the conditional mean, v is the same for all individuals:

$$v_i = \alpha^{-1} \tag{8}$$

for $\alpha > 0$, α is the dispersion parameter since increasing α increases the conditional variance of y.

$$\begin{aligned} Var\left(y_i/x_i\right) &= \lambda_i\left(1+\frac{\lambda_i}{\alpha^{-1}}\right) = \exp\left(x_i\beta\right)\left(1+\frac{\exp\left(x_i\beta\right)}{\alpha^{-1}}\right) \\ &= \lambda_i\left(1+\alpha\lambda_i\right) = \lambda_i + \alpha\lambda_i^2 \end{aligned} \tag{9}$$

If $\alpha = 0$, the mean and variance are equal [35].

4.1. Estimation of Parameters Using the Maximum Likelihood Estimation

Estimation of parameters in Poisson regression relies on maximum likelihood estimation (MLE) method. Maximum likelihood estimation gives an understanding of the values of the regression coefficients that are more likely to have given rise to the data. The maximum likelihood estimation for Poisson regression is discussed in detail below; let Y_i be the mean for the i^{th} response, for $i = 1, 2, \cdots, p$. The mean response is assumed to be a function of a set of explanatory variables, $X_1, X_2, \cdots, X_p$, the notation $\lambda\left(X_i, \beta\right)$ is used to denote the function that relates the mean response λ_i and X_i (the values of the explanatory variables for case i) and β (the values of the regression coefficients). Let's consider the Poisson regression model in the form below;

$$\lambda_i = \lambda\left(X_i, \beta\right) = \mathrm{e}^{X_i,\beta} \tag{10}$$

Then, from the Poisson distribution;

$$P(Y;\beta)=\frac{\left[\lambda(X_i\beta)\right]^{Y}\mathrm{e}^{-\lambda(X_i\beta)}}{Y!} \tag{11}$$

The likelihood function is given as,

$$L(Y;\beta)=\prod_{i=1}^{N}P(Y;\beta) \tag{12}$$

$$=\prod_{i=1}^{N}\frac{\left[\lambda(X_i\beta)\right]^{Y}\mathrm{e}^{-\lambda(X_i\beta)}}{Y!} \tag{13}$$

$$=\prod_{i=1}^{N}\frac{\left[\mu(X_i\beta)\right]^{Y}}{\mathrm{e}^{\sum_{i=1}^{N}\lambda(X_i\beta)}}\,Y! \tag{14}$$

The next thing to do is taking natural log of the above likelihood function. Then, differentiate the equation with respect to β and equate the equation to zero. The log likelihood function is given as,

$$\log L(Y_i,\beta)=\sum_{i=1}^{N}\left[Y_i\log\lambda(X_i,\beta)\right]-\lambda(X_i,\beta)-\log(Y_i!) \tag{15}$$

$$\frac{\partial}{\partial\beta}\left[\log L(Y;\beta)\right]=0 \tag{16}$$

The solution to the set of maximum likelihood given above must generally be obtained by iteration procedure. One of the procedure is known as iteratively re weighted least squares. This procedure will estimate the values of β. Maximum likelihood estimation produces Poisson parameters that are consistent, asymptotically normal and asymptotically efficient [37].

4.2. Negative Binomial Regression Analysis

The negative binomial regression model is derived by re writing Poisson regression model such that,

$$\log\lambda=\beta_0+\beta_iX_i+\varepsilon_i \tag{17}$$

where $\mathrm{e}^{\varepsilon_i}$ is a Gamma distributed error-term with mean 1 and variance α^2. This addition allows the variance to differ from the mean as,

$$Var(Y)=\lambda(1+\alpha\lambda)=\lambda+\alpha\lambda^2 \tag{18}$$

α also acts as a dispersion parameter. Poisson regression model is regarded as a limiting model of the negative binomial regression model as α approaches zero, which means that the selection between these two models is dependent upon the value of α. The negative binomial distribution has the form,

$$P(Y=y)=\frac{\Gamma\left(\frac{1}{\alpha}+y\right)}{\Gamma\left(\frac{1}{\alpha}\right)y!}\left[\frac{\frac{1}{\alpha}}{\left(\frac{1}{\alpha}\right)+\lambda}\right]^{\frac{1}{\alpha}}\left[\frac{\lambda}{\left(\frac{1}{\alpha}\right)+\lambda}\right]^{y} \tag{19}$$

where $\Gamma(.)$ is a gamma function. This results in the likelihood function,

$$L(Y_i)=\prod_i \frac{\Gamma\left(\frac{1}{\alpha}+y_i\right)}{\Gamma\left(\frac{1}{\alpha}\right)y_i!}\left[\frac{\frac{1}{\alpha}}{\left(\frac{1}{\alpha}\right)+\lambda_i}\right]^{\frac{1}{\alpha}}\left[\frac{\lambda_i}{\left(\frac{1}{\alpha}\right)+\lambda_i}\right]^{y_i} \tag{20}$$

Maximum likelihood estimation is used to estimate parameters in negative binomial. In addition, the interpretation of regression coefficients for negative binomial regression is the same as for Poisson regression.

4.3. Goodness of Fit

Deviance was used to test the goodness of fit of the model. Deviance is a measure of discrepancy between observed and fitted values. The deviance for Poisson responses takes the form

$$D=2\sum\left\{y_i\log\left(\frac{y_i}{\hat{\mu}_i}\right)-\left(y_i-\hat{\mu}_i\right)\right\}$$

The first term represents 'twice a sum of observed times log of observed over fitted'. The second term, a sum of differences between observed and fitted values, is usually zero, because maximum likelihood estimations in Poisson models have the property of reproducing marginal totals. For large samples of the distribution, the deviance is approximately a chi-square with $n-p$ degrees of freedom, where n is the number of observations and p the number of parameters. Therefore, the deviance can be used directly to test the goodness of fit of the model.

4.4. Residual Analysis

The Poisson regression is a non-normal regression, that is residuals are far from being normally distributed and the variances are non constant. Therefore we assess the model based on quantile residuals which removes the pattern in discrete data by adding the smallest amount of randomization necessary on cumulative probability scale. The quantile residuals are obtained by inverting the distribution function for each response.

Mathematically, let $a_i=\lim_{y\uparrow y_i}F\left(y;\hat{\mu},\hat{\Theta}\right)$ and $b_i=F\left(y_i;\hat{\mu},\hat{\Theta}\right)$ where F is the cumulative function of the probability density function $f\left(y;\mu,\Theta\right)$ then the randomized quantile residuals for y_i is $r_{q,r}=\Phi^{-1}\left(u_i\right)$ with u_i the uniform random variable on $\left(a_i,b_i\right]$. The randomized quantile residuals are distributed normally barring the variability in $\hat{\mu}$ and $\hat{\Theta}$ [38].

5. Results and Discussion

The expected malaria incidences was modeled using Poisson regression and the results are presented in Table 2. The model examines the association between monthly expected malaria incidences with monthly rainfall and monthly average temperature. From Table 2, it was observed that for every unit increase in rainfall, the expected malaria incidences increases by $e^{0.1443}=1.1551$ and for a unit increase in average temperature, the expected malaria incidences increase by

$e^{-0.4419} = 0.6428$ obtained from Equation (21). Based on the P-values, the average temperature and rainfall significantly affect the expected malaria incidences.

The residual deviance for the fitted Poisson regression was given as 89489 on 117 degrees of freedom.

The fitted Poisson model is given as from Table 2

$$\log A = 9.8881 + \left(7.877e^{-04}\right)R - \left(3.265e^{-02}\right)T = 9.888 + 0.1443R - 0.4419T \quad (21)$$

where A is the expected malaria incidences, R stands for rainfall and T stands for average temperature.

To check the goodness of fit of the fitted Poisson model, the value of the residual deviance 89,489 on 117 degrees of freedom was considered which is far greater than the number of degrees of freedom. Therefore it can be concluded that the model has lack of fit. Because the ratio between the residual deviance and the degrees of freedom is far greater than one which implies over dispersion exists. The fitted Poisson model had an AIC value of 90,813 and a null deviance of 100,505 on 119 degrees of freedom. The assumption of mean equal to variance was violated since the dispersion parameter was not approximately equal to the 1, an indication of over dispersion in the data. This meant that the parameters of the model had been over estimated and the standard errors had been under estimated which did not give a true reflection of the model that could provide appropriate expected malaria incidences from 2007 to 2016.

To address this error, Negative Binomial Regression was used to modify the model so that the case of over dispersion in the data was taken care of and the results were presented in Table 3. It was observed that the Negative Binomial was actually the best model which fit the expected malaria incidences because the dispersion parameter given by Poisson Regression Model had been reduced from 770 to 1.03.

Figures 2-5, show plots of the deviance residuals against the normal quantiles based on Poisson model and Negative binomial models respectively. Figure 2

Table 2. Parameter estimates of Poisson regression.

	Estimate	Standard Errors	z value	Pr (>|z|)
Intercept	9.888	$3.087e^{-02}$	320.35	$<2e^{-16}$***
Rainfall	$7.877e^{-04}$	$1.493e^{-05}$	52.76	$<2e^{-16}$***
Average Temperature	$-3.265e^{-02}$	$1.187e^{-03}$	−27.51	$<2e^{-16}$***

Table 3. Parameter estimates for negative binomial regression model for rainfall and average temperature.

	Estimate	Standard Errors	z value	Pr (>|z|)
Intercept	10.1940861	0.7788004	13.089	$<2e^{-16}$***
Rainfall	0.0008147	0.0003849	2.116	0.0343*
Average Temperature	−0.0451456	0.0298939	−1.510	0.1310

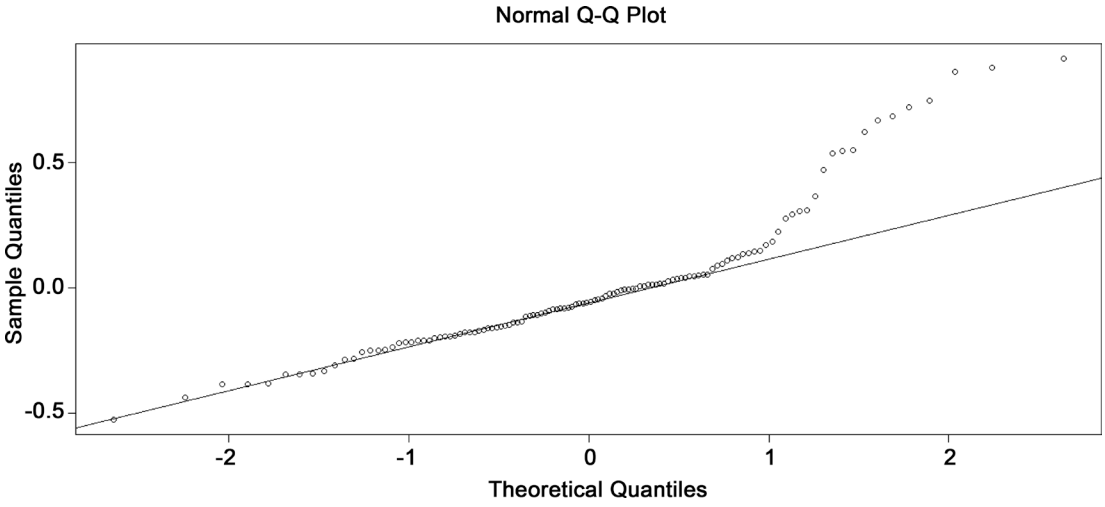

Figure 2. Normal Q-Q plot for Poisson regression.

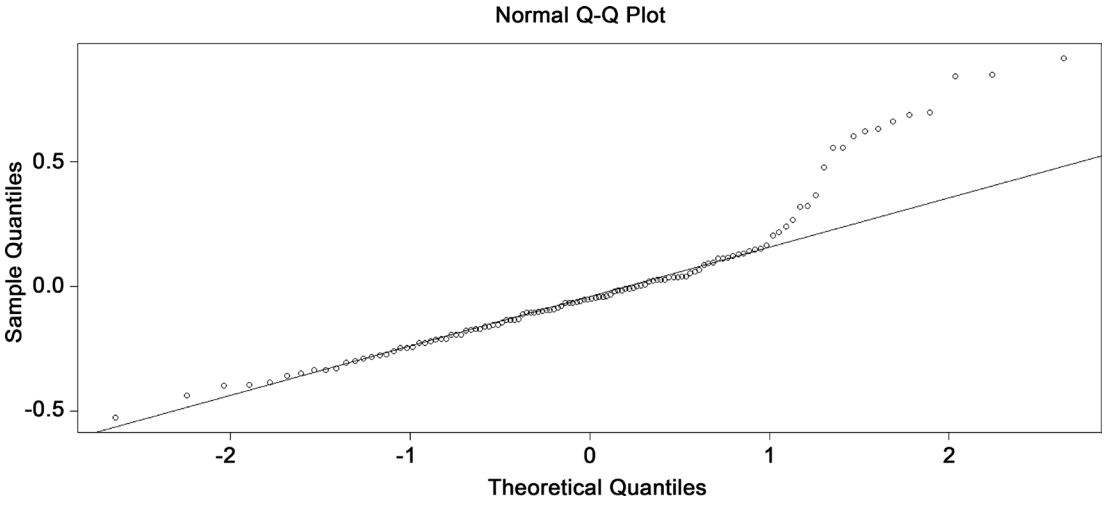

Figure 3. Normal Q-Q plot for negative binomial regression model between rainfall and expected malaria incidences.

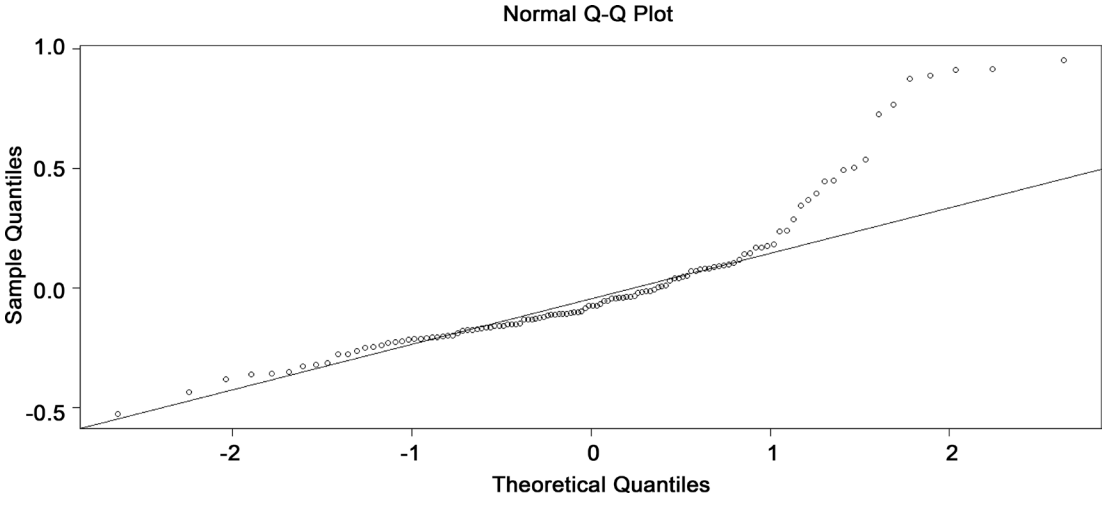

Figure 4. Normal Q-Q plot for negative binomial regression model between average temperature and expected malaria incidences.

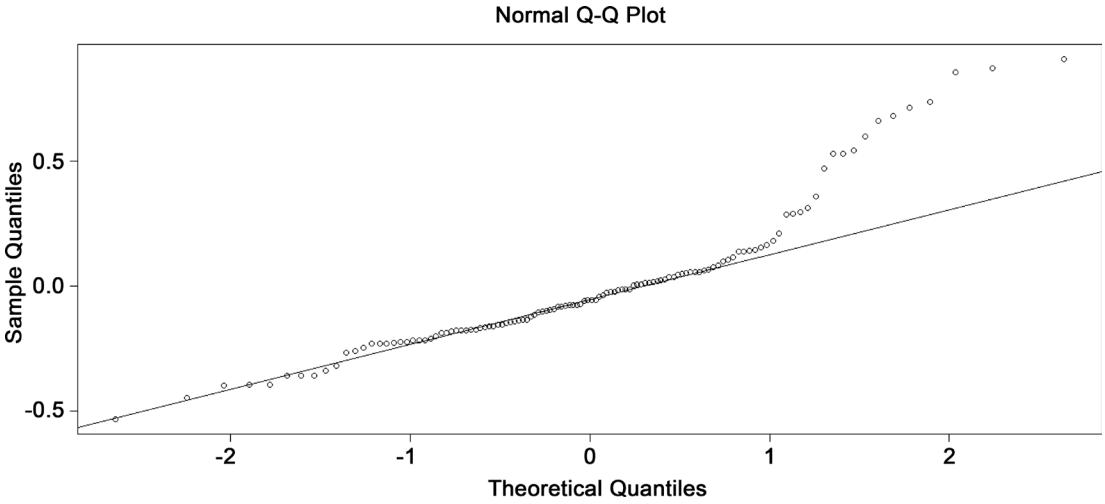

Figure 5. Normal Q-Q plot for negative binomial regression model between average temperature, rainfall and expected malaria incidences.

Table 4. Parameter estimates for negative binomial regression model for average temperature.

	Estimate	Standard Errors	z value	Pr (>|z|)
Intercept	11.37160	0.54154	21	$<2e^{-16}$***
Average Temperature	−0.08822	0.02178	−4.05	$5.12e^{-05}$***

Table 5. Parameter estimates for negative binomial regression model for rainfall.

	Estimate	Standard Errors	z value	Pr (>|z|)
Intercept	9.0252866	0.0436207	206.904	$<2e^{-16}$***
Rainfall	0.0011832	0.0002779	4.258	$2.06e^{-05}$***

for Poisson regression, the plot was not approximately linear just as for **Figure 4** and **Figure 5**. This indicated poor fit of the models. **Figure 3**, for the Negative Binomial model relating Malaria incidences and rainfall, the plot was approximately linear. This gave the best fit compared to the rest of the plots.

From the **Figure 2**, the points form a curve instead of a straight line. Normal Q-Q plots that look like this usually imply the model has a lack of fit.

Three models for Negative Binomial regression were considered and compared using the Akaike information criterion (AIC). The results for the two Negative Binomial Regression models without collinearity respectively are given in **Table 4** and **Table 5** respectively. The model from **Table 5** gave the lowest AIC value that is 2225.4 compared to the other model from **Table 4** which gave higher AIC value that is 2227.7. Hence Negative Binomial model for the relationship between rainfall and expected malaria incidences whose results were presented in **Table 5** was considered the model with better fit since it had lower AIC value. The lower the AIC value, the better the model.

The Negative Binomial regression model whose results were presented in **Ta-**

ble 3, showed null deviance to be 140.31 on 119 degrees of freedom, residual deviance to be 121.36 on 117 degrees of freedom and AIC to be 2225.3.

The Negative Binomial regression model whose results were presented in Table 4, showed null deviance to be 135.34 on 119 degrees of freedom, residual deviance to be 121.41 on 118 degrees of freedom and AIC to be 2227.7

The Negative Binomial regression model whose results were presented in Table 5 showed null deviance to be 137.88 on 119 degrees of freedom, residual deviance to be 121.38 on 118 degrees of freedom and AIC to be 2225.4.

The Model whose results were presented in Table 3 was not considered because of a negatively strong correlation between the explanatory variables that is rainfall and average temperature which was reported in Table 6 as −0.6985212. Therefore one of the independent variables was removed from the model and two Negative Binomial regression models were developed and results for the parameter estimates presented in Table 4 and Table 5 respectively. Based on P-value, the results in Table 4 and Table 5 showed that average temperature and rainfall significantly affected the expected malaria incidences respectively.

Pearson correlation test was performed and the results were presented in Table 7. The results showed that malaria and rainfall were strongly positively associated based on the P-value which is significant since it was found to be less than 5% significance level and a positive confidence interval. Malaria and Average temperature were found to be negatively associated and not significant given the P-value was greater than 5% significance level.

5.1. Interpretation of Coefficients

From Table 5, we observe that rainfall is very significant at 5% significance level with their significance value equal to 0.01388. For every one unit increase in amount of rainfall, the expected malaria incidences increases by $e^{0.0011832} = 1.0011839003$ times. From Table 3, we observe that rainfall is slightly significant with significance value 0.0343* at 5% significance level while temperature is not significant at 5% significance level with P-value of 0.1310. This indicated the presence of collinearity between rainfall and average temperature.

Table 6. Correlation between rainfall and average temperature.

Climate Variables	Correlation value
Rainfall and Average Temperature	−0.6985212

Table 7. Pearson correlation test results.

	Malaria and Average temperature	Malaria and rainfall
t-value	−3.0311	3.4594
P-value	0.9985	0.0003771
confidence interval	[−0.4033168,1.0000000]	[0.1598574,1.0000000]
degrees of freedom	118	118

5.2. Discussion

Malaria is transmitted by the female Anopheles mosquito. The female Anopheles mosquito go through four stages in their life cycle that is egg, larva pupa and adult [39]. The first three stages are aquatic and also depend on the temperature. The adult stage is when the female Anopheles mosquito acts as malaria vector [39]. Once adult mosquitoes have emerged, the temperature, humidity and rainfall determine their chances of survival. To transmit malaria successfully, female Anopheles must survive long enough after they have become infected to allow the parasites they harbor to complete their growth cycle [31]. A conducive climatic environment will also shorten the time required for the parasite development in the mosquito [32]. The climate variables can affect the malaria incidences by affecting the life cycle of the mosquito development and the parasite in the mosquitoes.

Pearson correlation between rainfall and average temperature showed a strong negative correlation. This highlights the importance of removing one of the climate variables from the model to avoid invalid association due to collinearity. In this study, rainfall was the only climate variable considered in the Negative Binomial Regression model since it presented the best fit. Negative Binomial regression model relating expected malaria incidences, rainfall and temperature was not selected as the final model due to high correlation between rainfall and average temperature which affected the significance of individual climate variables to expected malaria incidences. The model results showed that average temperature was not significant in the model while rainfall was weakly significant. This result was seen to contradict the biology of mosquito development.The model relating malaria incidences and average temperature showed a significant positive relationship though it was not the model selected since it had a higher AIC value. Modeling has shown that optimal malaria transmission occurs at 25˚C and malaria transmission decreases at temperature above 28˚C [40]. Temperatures below 16˚C are also detrimental for survival of mosquitoes [41]. Results from previous study showed a strong positive association between malaria incidence in a given month and the minimum temperature of the previous month. This indicates that minimum temperature also affects malaria transmission [23].

In previous studies of climatic effects on malaria incidence, different results on the effect of rainfall on malaria incidence were found. [29], found rainfall to be significant when precipitation was 2.4 times higher than the normal level. Rainfall plays an important role in the survival of mosquitoes, since water pools from the rain provide a habitat for mosquito larvae to develop. [19], found there was no significant effect when rainfall was less than 100 mm per month in Yunnan, China. A study in Ethiopia found that rainfall had a significant effect on malaria incidence in hot districts with an altitude lower than 1650 mm, but not in cold districts with an altitude higher than 1650 mm [28].

In the present study, there was a positive significant effect between rainfall

and malaria incidences, similar to previous findings [20] [28]. The positive effects were reasonable because rain water forms water pools which provide a breeding ground for mosquitoes, hence increasing the mosquito density which in turn leads to increase in malaria incidences. To our knowledge, no study has investigated the association between climate variables and malaria incidences in Apac district, Uganda.

The study had it's own limitations such as short data length and not being able to include non-climatic variables in the models. The relationship between malaria incidences and climate variables a period of 10 years was not found to be sufficient enough to predict future occurrences. Malaria incidence is associated with socio-economic conditions of the people as well as malaria control measures. These factors were not incorporated in the models.

6. Conclusion and Recommendation

Malaria remains an important public health problem in Apac district, Northern Uganda. The objective of this study was to model the factors associated with malaria incidences in Apac District. The study used monthly data for the period January 2007 to December 2016 in Apac district. The Poisson regression did not accurately fit the data on malaria incidences due to over dispersion in the data. The Negative Binomial Model was a better fit. The result obtained suggested that rainfall was positively significant on monthly malaria incidences whereas average temperature was not a significant predictor for malaria incidences based on results from Pearson correlation test in Apac District. A positive relationship between rainfall and expected malaria incidences was observed based on the coefficient value of parameter estimates in **Table 5**. The findings provide better insight of climate effects on malaria and provide important information for malaria prediction. It is observed that rainfall is a strong predictor of malaria incidences in Apac District. We recommend that in future studies, these factors should be incorporated in the models and more lengthy data set should be used.

Acknowledgements

We are grateful to all the authors who contributed generously to this work.

Competing Interests

The authors declare that they have no competing interests.

Authors' Contributions

All the authors have read and approved the final manuscript.

References

[1] Ministry of Health (2014) Uganda Malaria Reduction Strategic Plan 2014-2020. Ministry of Health, Uganda.

[2] Talisuna, A.O., Okello, P.E., Erhart, A., Coosemans, M., Alessandro, U.D., *et al.*

(2007) Intensity of Malaria Transmission and the Spread of Plasmodium Falciparum-Resistant Malaria: A Review of Epidemiologic Field Evidence. *The American Journal of Tropical Medicine and Hygiene*, **77**, 170-180.

[3] O'Meara, W.P., Mangeni, J.N., Steketee, R. and Greenwood, B. (2010) Changes in the Burden of Malaria in Sub-Saharan Africa. *The Lancet Infectious Diseases*, **10**, 545-555.

[4] Lubanga, R.G.N., Norman, S., Ewbank, D. and Karamagi, C. (1997) Maternal Diagnosis and Treatment of Children's Fever in an Endemic Malaria Zone of Uganda: Implications for the Malaria Control Programme. *Acta Tropica*, **68**, 53-64.

[5] Van Bortel, W., Byaruhanga, A.M., Correwyn, A., Roelants, P., Talisuna, A., D'Alessandro, U., Coosemans, M. and Okello, P.E. (2006) Variation in Malaria Transmission Intensity in Seven Sites throughout Uganda. *The American Journal of Tropical Medicine and Hygiene*, **75**, 219-225.

[6] Martens, P. and Hall, L. (2000) Malaria on the Move: Human Population Movement and Malaria Transmission. *Emerging Infectious Diseases*, **6**, 103.

[7] Yeka, A., Gasasira, A., Mpimbaza, A., Achan, J., Nankabirwa, J., Nsobya, S., Staedke, S.G., Donnelly, M.J., Wabwire-Mangen, F., Talisuna, A., *et al.* (2012) Malaria in Uganda: Challenges to Control on the Long Road to Elimination: I. Epidemiology and Current Control Efforts. *Acta Tropica*, **121**, 184-195.

[8] Aribodor, D.N., Ugwuanyi, I.K. and Aribodor, O.B. (2016) Challenges to Achieving Malaria Elimination in Nigeria. *American Journal of Public Health Research*, **4**, 38-41.

[9] Pascual, M., Ahumada, J.A., Chaves, L.F., Rodo, X. and Bouma, M. (2006) Malaria Resurgence in the East African Highlands: Temperature Trends Revisited. *Proceedings of the National Academy of Sciences*, **103**, 5829-5834. https://doi.org/10.1073/pnas.0508929103

[10] Rogers, D.J. and Randolph, S.E. (2000) The Global Spread of Malaria in a Future, Warmer World. *Science*, **289**, 1763-1766.

[11] Hulden, L. and Hulden, L. (2009) The Decline of Malaria in Finland—The Impact of the Vector and Social Variables. *Malaria Journal*, **8**, 94. https://doi.org/10.1186/1475-2875-8-94

[12] Connor, S.J., Thomson, M.C. and Molyneux, D.H. (1999) Forecasting and Prevention of Epidemic Malaria: New Perspectives on an Old Problem. *Parassitologia*, **41**, 439-448.

[13] Thomson, M.C., Mason, S.J., Phindela, T. and Connor, S.J. (2005) Use of Rainfall and Sea Surface Temperature Monitoring for Malaria Early Warning in Botswana. *The American Journal of Tropical Medicine and Hygiene*, **73**, 214-221.

[14] Shanks, G.D., Hay, S.I., Stern, D.I., Biomndo, K. and Snow, R.W. (2002) Meteorological Influences on Plasmodium Falciparum Malaria in the Highland Tea Estates of Kericho, Western Kenya. *Emerging Infectious Diseases*, **8**, 1404-1408. https://doi.org/10.3201/eid0812.020077

[15] Patience, E.O. and Osagie, A.M. (2014) Modeling the Prevalence of Malaria in Niger State: An Application of Poisson Regression and Negative Binomial Regression Models. *International Journal of Physical Sciences*, **4**, 061-068.

[16] Kakchapati, S. and Ardkaew, J. (2011) Modeling of Malaria Incidence in Nepal. *Journal of Research in Health Sciences*, **11**, 7-13.

[17] Drebel, T., Kueil, B.G. and Meyrowitsch, D.W. (2013) Prevalence of Malaria and Use of Malaria Risk Reduction Measures among Resettled Pregnant Women in

South Sudan. *International Health*, **5**, 211-216. https://doi.org/10.1093/inthealth/iht008

[18] Chanda, E., Doggale, C., Pasquale, H., Azairwe, R., Baba, S. and Mnzava, A. (2013) Addressing Malaria Vector Control Challenges in South Sudan: Proposed Recommendations. *Malaria Journal*, **12**, 59. https://doi.org/10.1186/1475-2875-12-59

[19] Wardrop, N.A., Barnett, A.G., Atkinson, J.-A. and Clements, A.C.A. (2013) Plasmodium Vivax Malaria Incidence over Time and Its Association with Temperature and Rainfall in Four Counties of Yunnan Province, China. *Malaria Journal*, **12**, 452. https://doi.org/10.1186/1475-2875-12-452

[20] Kim, Y.-M., Park, J.-W. and Cheong, H.-K. (2012) Estimated Effect of Climatic Variables on the Transmission of Plasmodium Vivax Malaria in the Republic of Korea. *Environmental Health Perspectives*, **120**, 1314-1319. https://doi.org/10.1289/ehp.1104577

[21] Sriwattanapongse, W., Me-ead, S. and Khanabsakdi, S. (2011) Forecasting Malaria Incidence Based on Monthly Case Reports and Climatic Factors in Ubon Ratchathani Province, Thailand, 2000-2009. *Science Journal Ubon Ratchathani University*, **2**, 17-24.

[22] Nath, D.C. and Mwchahary, D.D. (2013) Association between Climatic Variables and Malaria Incidence: A Study in Kokrajhar District of Assam, India: Climatic Variables and Malaria Incidence in Kokrajhar District. *Global Journal of Health Science*, **5**, 90.

[23] Nkurunziza, H., Gebhardt, A. and Pilz, J. (2011) Geo-Additive Modeling of Malaria in Burundi. *Malaria Journal*, **10**, 234. https://doi.org/10.1186/1475-2875-10-234

[24] Huang, F., Zhou, S., Zhang, S., Zhang, H. and Li, W. (2011) Meteorological Factors-Based Spatio-Temporal Mapping and Predicting Malaria in Central China. *The American Journal of Tropical Medicine and Hygiene*, **85**, 560-567. https://doi.org/10.4269/ajtmh.2011.11-0156

[25] Nkurunziza, H., Gebhardt, A. and Pilz, J. (2010) Bayesian Modelling of the Effect of Climate on Malaria in Burundi. *Malaria Journal*, **9**, 114. https://doi.org/10.1186/1475-2875-9-114

[26] Muggeo, V.M.R. (2008) Modeling Temperature Effects on Mortality: Multiple Segmented Relationships with Common Break Points. *Biostatistics*, **9**, 613-620. https://doi.org/10.1093/biostatistics/kxm057

[27] Zhou, G., Minakawa, N., Githeko, A.K. and Yan, G. (2004) Association between Climate Variability and Malaria Epidemics in the East African Highlands. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 2375-2380. https://doi.org/10.1073/pnas.0308714100

[28] Teklehaimanot, H.D., Lipsitch, M., Teklehaimanot, A. and Schwartz, J. (2004) Weather-Based Prediction of Plasmodium Falciparum Malaria in Epidemic-Prone Regions of Ethiopia I. Patterns of Lagged Weather Effects Reflect Biological Mechanisms. *Malaria Journal*, **3**, 41. https://doi.org/10.1186/1475-2875-3-41

[29] Lindsay, S.W., Bodker, R., Malima, R., Msangeni, H.A. and Kisinza, W. (2000) Effect of 1997-98 El-Nino on Highland Malaria in Tanzania. *The Lancet*, **355**, 989-990. https://doi.org/10.1016/S0140-6736(00)90022-9

[30] FMOH (2000) Federal Ministry of Health, Situational Analysis of Malaria Control in Nigeria, Abuja. Federal Ministry of Health Report.

[31] Martens, P., Kovats, R.S., Nijhof, S., De Vries, P., Livermore, M.T.J., Bradley, D.J., Cox, J. and McMichael, A.J. (1999) Climate Change and Future Populations at Risk

of Malaria. *Global Environmental Change*, **9**, S89-S107. https://doi.org/10.1016/S0959-3780(99)00020-5

[32] Kres, A.C., Schwarz, N.G., Kruger, A., Fobil, J., Nkrumah, B., Acquah, S., Loag,, W., Sarpong, N., Adu-Sarkodie, Y., Ranft, U., *et al.* (2011) Modeling the Relationship between Precipitation and Malaria Incidence in Children from a Holoendemic Area in Ghana. *The American Journal of Tropical Medicine and Hygiene*, **84**, 285-291. https://doi.org/10.4269/ajtmh.2011.10-0381

[33] Nelder, A.J. and McCullagh, P. (1992) Generalized Linear Models. In Breakthrough in Statistics. Springer, New York.

[34] Venables, W.N. and Ripley, B.D. (2002) Modern Applied Statistics with S. Springer, New York, 271-300. https://doi.org/10.1007/978-0-387-21706-2

[35] Gujarati, D.N. (2009) Basic Econometrics. Tata McGraw-Hill Education, New York.

[36] Williams, R. (2016) Models for Count Outcomes. University of Notre Dame, Notre Dame.

[37] Agresti, A. (2002) Inference for Contingency Tables. *Categorical Data Analysis, Second Edition*, **7**, 70-114. https://doi.org/10.1002/0471249688.ch3

[38] Dunn, P.K. and Smyth, G.K. (1996) Randomized Quantile Residuals. *Journal of Computational and Graphical Statistics*, **5**, 236-244.

[39] Hunter, P.R. (2003) Climate Change and Waterborne and Vector-Borne Disease. *Journal of Applied Microbiology*, **94**, 37-46. https://doi.org/10.1046/j.1365-2672.94.s1.5.x

[40] Mordecai, E.A., Paaijmans, K.P., Johnson, L.R., Balzer, C., Ben-Horin, T., Moor, E., McNally, A., Pawar, S., Ryan, S.J., Smith, T.C., *et al.* (2013) Optimal Temperature for Malaria Transmission Is Dramatically Lower than Previously Predicted. *Ecology letters*, **16**, 22-30. https://doi.org/10.1111/ele.12015

[41] Patz, J.A. and Olson, S.H. (2006) Malaria Risk and Temperature: Influences from Global Climate Change and Local Land Use Practices. *Proceedings of the National Academy of Sciences*, **103**, 5635-5636. https://doi.org/10.1073/pnas.0601493103