Spatial Temporal Prediction of Malaria Risk in Western Kenya using Bayesian Geostatistical Approach

Gilbert K. Sang[1], Edward H. Waithaka PhD[2]

[1]Jomo Kenyatta University of Agriculture and Technology, Department of Geomatic Engineering and Geospatial Information Systems, P.o. Box 62000 00200 Nairobi, Kenya

[2]Jomo Kenyatta University of Agriculture and Technology, Department of Geomatic Engineering and Geospatial Information Systems, P.o. Box 62000 00200 Nairobi, Kenya

Abstract: ***Malaria is a vector borne disease that occurs in areas where the climatic and environmental conditions are suitable for survival of Anopheles mosquitoes. The environmental and climatic factors that affect malaria transmission are, rainfall, temperature, humidity and vegetation. Deforestation, agricultural activities and population movements are an anthropogenic factor that affects malaria transmission. Despite implementation of several strategies in controlling and management of malaria in western parts of Kenya, high numbers of malaria cases are still being recorded due to changing environmental and climatic factors. The aim of this study was to apply the geostatistical modelling to estimate and map the spatial and temporal changes in malaria risk by using the available time series climatic and environmental data and to estimate the population at risk at different time epochs. The data was prepared using python scripts and different ArcGIS tools. A spatial temporal model based on Bayesian approach was used to estimate malaria risk and was implemented in R using Integrated Nested Laplace Approximation (INLA) package to estimate the malaria risk. INLA was preferred to Monte Carlo Markov Chain (MCMC) due to its efficient computation advantage.***

Keywords: Plasmodium *falciparum* rate, ArcGIS, GIS, Raster, INLA, Population at Risk

1. Introduction

Malaria is a leading cause of mortality and morbidity in the world. The disease which is caused by malaria parasite (Plasmodium falciparum) is the leading cause of deaths in most developing countries especially in Sub-Saharan Africa where rate of transmission is highest in the world [1]. This is a major inhibition to economic development due to large amount of resources mobilized in prevention and control of the disease [1].

The World Health Organization estimates that 3.4 billion people live in areas at risk of malaria transmission in 106 countries and territories [2]. In 2012 alone, Malaria caused approximately 207 million clinical episodes which resulted in 655,000 deaths [3]. 86% of the total deaths were children under the age of 5 years because their bodies have not developed partial immunity to malaria parasite [3].

In Kenya 70% of the population live in malaria prone areas [4]. Depending on intensity of malaria transmission, four malaria zones can be delineated s namely; endemic lake and coastal regions, epidemic-prone highland, seasonal transmission risk district and low risk districts [4]. The malaria prevalence rates vary within zones with endemic areas having a rate of 20% to 40%, 5% to 20% in highland epidemic prone areas. Seasonal malaria transmission areas comprise mainly the arid and semi-arid regions of Kenya is characterized by low prevalence of less than 5%.

Malaria is an environmental disease transmitted by vectors which require optimum climatic and environmental conditions for survival. The climatic and environmental factors vary from one region to another which in turn leads to varying malaria transmissions.

Environmental factors that affect the development and survival of malaria vectors and parasites include; rainfall, temperature, humidity, surface water and vegetation. Any change in these factors has a major impact on malaria transmission [5].

Anthropogenic factors that affect malaria prevalence include deforestation, urbanization, irrigation, population movements among other human induced changes [6]. The vectorial capacity of deforested areas are 77.7% higher than forested areas, therefore deforestation and human caused environmental changes exposes the population with low functional immunity to malaria parasite carrying vectors leading to spread in transmission [7].

2. Problem statement

The government of Kenya and Non-Governmental Organizations (NGOs) have intensified malaria control strategies in malaria prone areas especially in Western Kenya. However, with increased interventions and control strategies being implemented, high number of malaria cases have continued being recorded in Homabay, Siaya, Vihiga and Kisumu counties which border Lake Victoria and also in adjacent counties of Kericho, Nandi and Bomet located in Kenya highlands which were initially had low risk or malaria free transmission [8] [9]. This increase in malaria transmissions may be due to changes in anthropogenic and environmental factors that have occurred over time and space. Therefore it is imperative to develop methods of estimating their effect on malaria risk as well as provide information on the factors that drive the increase and redistribution.

Most studies have not considered the effect of how changing environmental and climatic factors affect spatial and temporal variation in malaria which may have led to formulation of non-effective control strategies.

In this study geostatistical model was adopted. This approach enables a continuous and accurate modelling and prediction of malaria risk taking into consideration environmental, climatic and other factors that directly influence malaria transmission [10].

The objective of this study was to accurately and precisely estimate the malaria prevalence in western Kenya while taking into account environmental and climatic factors. The resulting population at risk maps will help in understanding the malaria changes over time which will aid in optimization of control strategies.

3. Study Area

The study area in the research covered 9 counties in western Kenya with an area of 17,227.6 km^2. They include Kisumu, Siaya, Homabay, Kisii, Nyamira, Vihiga, Nandi, Bomet and Kericho Counties. The study area lies between longitudes33° 56' 23"Eto 35° 40' 12"E and latitudes between 1° 1' 48S to 0° 34' 11"N.

The area is inhabited by a population of approximately 6.8 million [11]. The main economic activities in lake counties include subsistence farming, fishing, rice farming, sugarcane farming, livestock keeping among others while the highland counties of Kericho, Bomet and Nandi include tea growing and processing, dairy Farming, horticulture and floriculture, wheat, fish farming, commercial businesses.

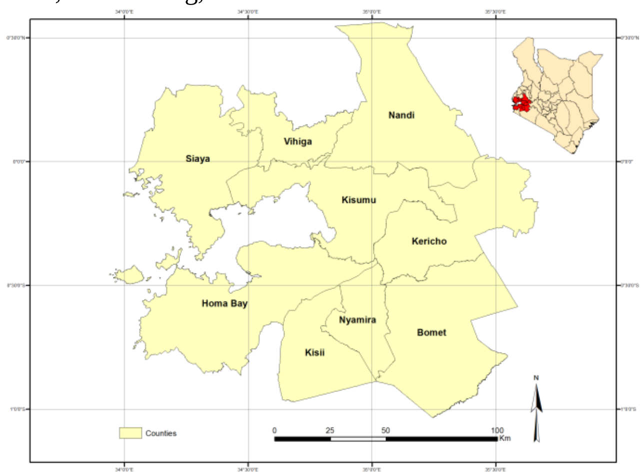

Figure 1: Map of study area

4. Methodology

The data used for this study was identified based on previous literature on the various climatic and ecological determinants which directly affect the development and survival of the plasmodium *falciparum* parasite as well as the malaria transmitting vectors. The data used for this study were population density, temperature suitability Index (TSI), rainfall, deforestation and enhanced vegetation Index. Figure 2 shows the study workflow consisting of several processes. Time series data were downloaded from different data sources and processed using python scripts. Table 1 shows the data used their sources and projection.

Table 1: Data sources and projection

Data	*Source*	*Projection*
Temperature	Global (Land) Precipitation and Temperature (https://climatedataguide.ucar.edu)	WGS84
Rainfall	Fews Net http://earlywarning.usgs.gov/fews/	Africa Albers Equal Area Conic Projected coordinate system
EVI	Land Processes Distributed Active Archive Center (LP DAAC)	WGS 84
Vegetation Cover	Land Processes Distributed Active Archive Center (LP DAAC)	WGS 84
Population	Afripop Project http://www.worldpop.org.uk/	WGS 84
TSI	Malaria Atlas Project (http://www.map.ox.ac.uk/)	WGS 84

The data manipulation processes included projection definition, raster data combination and extraction of values, vector data clipping and masking of raster data. Batch resampling was done to make sure all covariates were 500 meters in pixel resolution

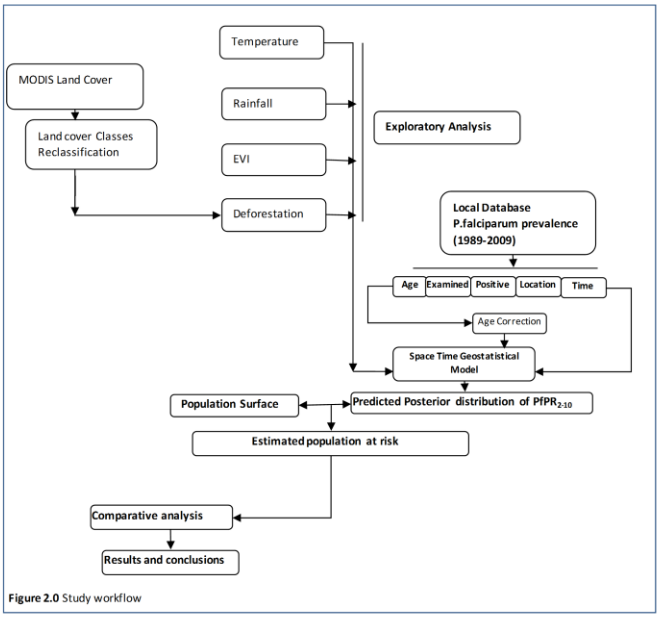

Figure 2: Study methodology

5. Spatial-Temporal Analysis

The plasmodium falciparum surveys had different age ranges for sampled populations, to ensure similarity between different surveys data they were age corrected to standardized range of 2-10 years (PfPR2-10). Since the survey prevalence data was unevenly distributed in both time (years) and space (longitudes and latitudes) a Bayesian geostatistical generalized linear mixed model (BGGLM) with spatial and temporal random effects was used to predict PfPR2-10 in the study region. The model assumes that the number of children from 2 years to below 10 years of age that are positive in the study region at any time is a binomial random variable. The logistic model incorporates the covariates effects, spatial autocorrelations and time as a second order autoregressive effects. Bayesian inference was achieved using intergrated nested laplace approximation (INLA) in R [12]. Model posterior outputs included the predicted mean and standard deviation of $PfPR_{2\text{-}10}$.

6. Results and Discussion

Deforestation

Figure 3 Shows areas in which deforestation occurred between 2000 and 2009.

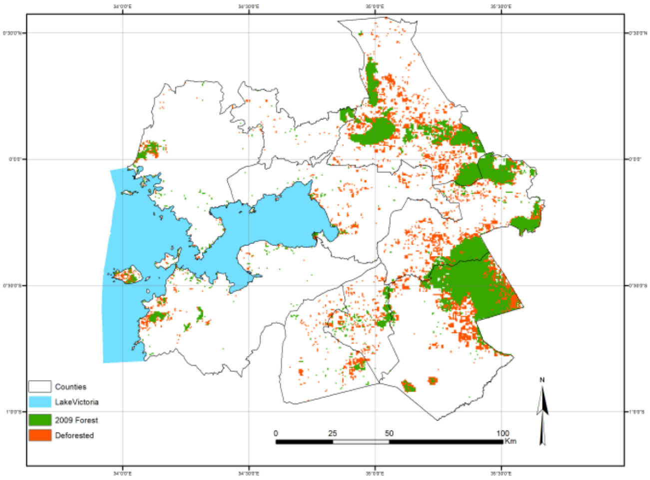

Figure 3: Map of deforested areas

Rainfall

The rainfall patterns in the area as shown in figure 4 depicts increased rainfall amounts between 1998 and 1999 when the El nino rains were experienced in western Kenya. However, the total rainfall remained relatively low from 2001 to 2008 followed by an increase in 2009

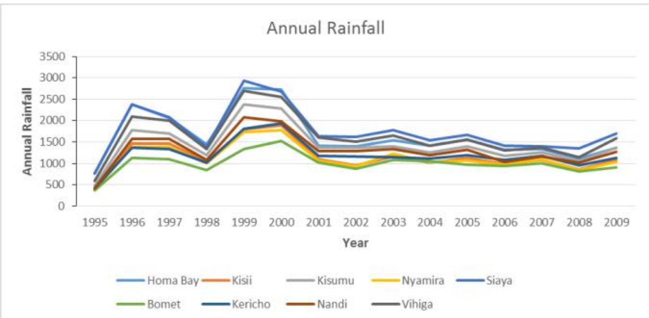

Figure 4: Long Term Rainfall Trend

Temperature

Figure 5 shows the temporal changes of the mean annual temperature in each of the counties under study. Siaya County had the highest average temperature while Kericho and Nandi had low temperatures throughout the study period. However uniform increase in temperature of up to 1°C is experienced between1996-2000, 2002-2005 and 2007-2009. Year 2001 was the coldest in all the counties under study

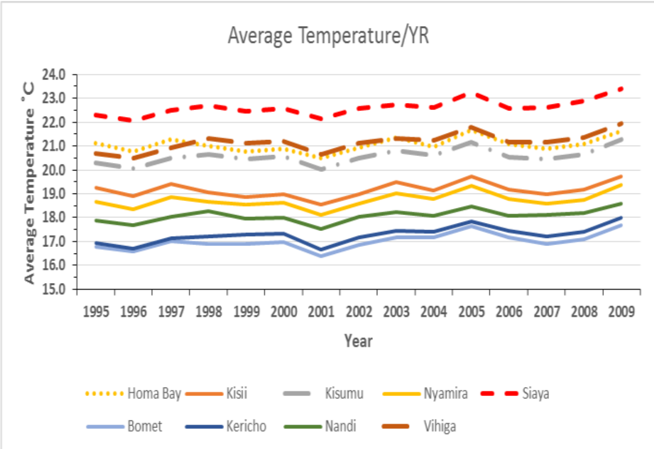

Figure 5 long term rainfall trend

Average EVI.

High average EVI values were recorded in all counties in 2001 with a gradual decline until 2005 when they started to rise achieving peak values in 2005 followed by continous decline upto 2009. There was increase in EVI values from 2009 to 2010.

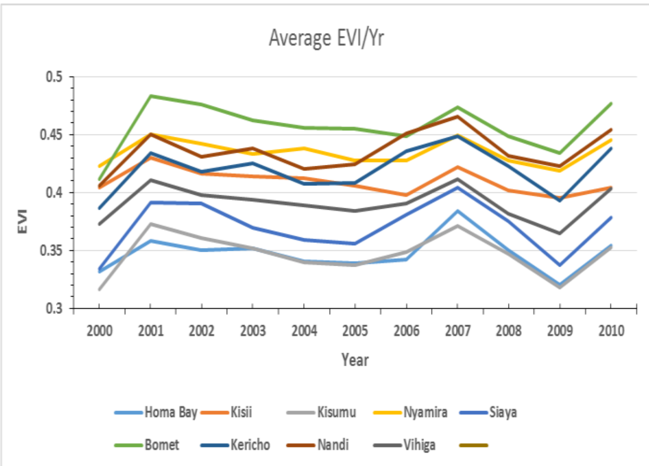

Figure 6: Average EVI change over time

Malaria Risk

Figure 7 shows maps of changing Plasmodium *falciparum* ($PfPR_{2\text{-}10}$) risk in the study area. The malaria risk ranges between 0% (light yellow) the lowest to 100% (dark red) the highest.

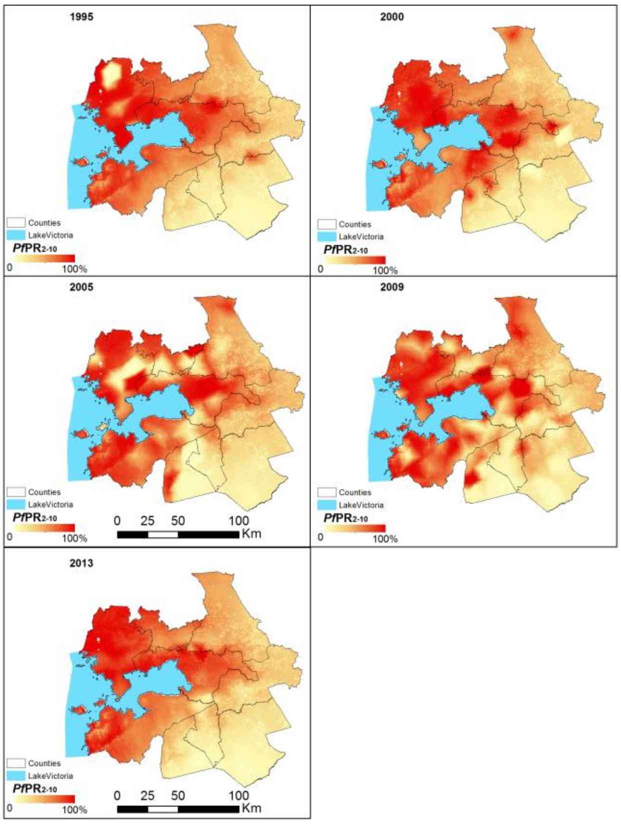

Figure 7: Maps of changing Malaria risk

In 1995 there was high malaria risk around the lake region and reduces as one moves away towards highland areas. The border between Kisumu and Nandi counties had a high risk malaria hotspot. Bomet County had low risk in most of the areas, however, there is increased risk towards the north where there was a hotspot at the border with Kericho County
In the year 2000, Siaya, Kisumu and Homabay counties had mid to high malaria risk. The highest risk was in the eastern part of Kisumu counties as well as the border between Siaya and Kisumu counties. High malaria risk hotspots were also experienced at the border between Kericho and Kisumu counties as well as northern part of Nandi County.

In 2005 high malaria risk hotspots occurred on the northern part of Kisumu County and at the border with Siaya County. The western region of Homabay County was characterized by medium to high risk in malaria transmission. Vihiga County experienced high malaria risk at the north eastern parts towards Nandi County. Kisii, Nandi and Kericho Counties experienced moderately low risk in malaria transmission with Bomet County experiencing the lowest malaria risk.

In 2009, high malaria risk hotspots were observed in Siaya, Homabay and Kisii counties with the highest risk experienced in Kisumu County. Larger extents of Nandi, Vihiga and Kericho counties had moderate malaria risk. Bomet County had lower malaria risk throughout the county except in the hotspot that occurred at the border with Nyamira County.

In 2013 malaria risk distribution is high around Lake Victoria. Siaya, Vihiga, Kisumu and Homabay all had high risk of malaria infection, but less compared to other years. Nandi and Kericho counties had higher risk in areas bordering lake counties. A similar scenario is observed in Kisii County where the risk northwards to the border with Homabay County.

Population at Risk

Figure 8 shows the percentage of the population at risk over the years in the counties under study. The graph shows that malaria risk increased between 1995 and 2000 in Homabay, Kisii, Kisumu and Nyamira counties followed by decrease in 2005 and 2009 with an increase in 2013. Population at risk of malaria decreased between 1995 and 2000 in Siaya, Vihiga, Kericho, Bomet and Nandi. All counties under study except Nandi, Bomet and Kisii had increase in population at risk of malaria between 2009 and 2013. Homabay, Siaya, Kisumu and Vihiga had the highest population at risk with Bomet County having low malaria population at risk.

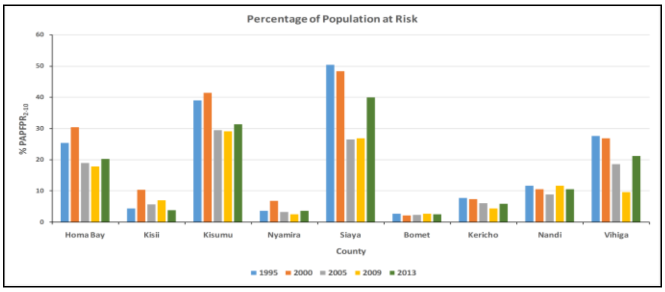

Figure 8: Graph of percentage of population at risk

Figure 9 represents the maps of population at risk of malaria infection. The map shows a general increase in population at risk of malaria from 1995 to 2000, followed by a decrease between 2000 and 2009 with an increase in 2013.

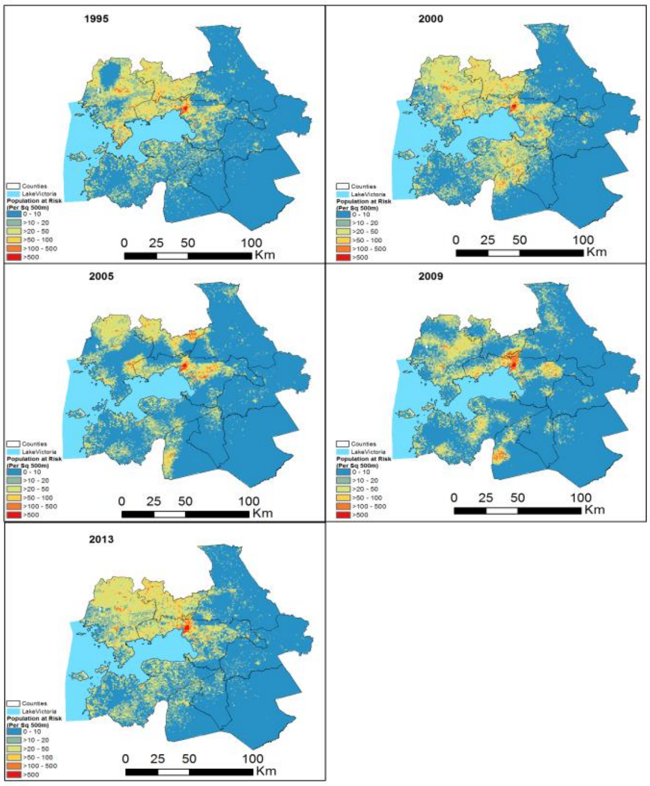

Figure 9: Maps of population at risk

Table 2: Percentage change in population at risk

County	1995 to 2000 % change in Population under risk	2000 to 2005 % change in Population under risk	2005 to 2009 % change in Population under risk	2009 to 2013 % change in Population under risk
Homa Bay	5.0	-11.6	1.1	2.4
Kisii	5.9	-4.6	-1.2	-3.1
Kisumu	2.5	-12.0	0.4	2.2
Nyamira	3.2	-3.6	0.8	1.2
Siaya	-2.1	-21.6	-0.3	13.1
Bomet	-0.6	0.1	-0.4	-0.3
Kericho	-0.4	-1.3	1.8	1.5
Nandi	-1.0	-1.8	-2.9	-1.2
Vihiga	-0.7	-8.4	8.9	11.4

Table 2 shows that population at risk of malaria increased from 1995 to 2000 in Homabay, Kisii, Kisumu and Siaya counties. Other counties experienced a decrease in malaria of ranging 0.7% to 2.1%. For the period from 2000 to 2005, all counties except Bomet experienced decrease in population at risk with the greatest decrease in Siaya County at 21.6%.

In the period 2005 to 2009, most counties experienced slight changes in population at risk with the highest decrease of 2.9% experienced in Nandi County. The highest increase during this period was in Kericho which experienced an increase of 1.8%.
For the period 2009 to 2013 there was a sharp increase in population at malaria risk in Vihiga and Siaya Counties which experienced an increase of 11.4% and 13.1% respectively. Homabay, Kisumu, Nyamira and Kericho experienced a slight increase at 1.2% to 2.4%. However a decrease in population at malaria risk was experienced in Kisii, Bomet and Nandi Counties at 3.1%, 0.3% and 1.2% respectively.

Table3 shows the total population and the estimated population at risk in each county per year of prediction.

Table 3: Population and population at risk per county

District	1995 Population	1995 Population at Risk	2000 Population	2000 Population at Risk	2005 Population	2005 Population at Risk	2009 Population	2009 Population at Risk	2013 Population	2013 Population at Risk
Homa Bay	604,968	153,842	652,034	198,132	732,033	137,402	799,144	141,415	883,109	177,155
Kisii	727,379	31,331	783,969	79,959	883,199	49,356	967,831	66,171	1,069,520	39,603
Kisumu	595,345	231,644	641,663	265,423	739,821	217,616	831,041	240,863	918,357	286,749
Nyamira	390,089	13,887	420,438	28,428	470,484	14,808	511,766	11,956	565,536	19,870
Siaya	523,064	262,991	563,759	271,576	632,451	167,928	689,860	185,039	762,343	304,127
Bomet	570,549	15,300	614,938	12,591	688,445	15,009	749,224	19,705	827,943	19,352
Kericho	372,546	28,610	401,531	29,383	452,877	27,166	496,904	21,039	549,112	31,303
Nandi	497,498	57,099	536,203	56,321	599,487	52,228	651,432	75,632	719,877	75,127
Vihiga	511,193	140,943	550,964	147,903	615,435	113,572	668,092	64,071	738,288	155,087
Total	4,792,631	935,647	5,165,499	1,089,716	5,814,232	795,084	6,365,294	825,891	7,034,085	1,108,372

7. Model Validation

In order to assess the validity of the estimated model a residual analysis was performed for the 10% of data which was used as validation data. The resulting correlation coefficient was 0.96 which shows the model ability to predict values at unobserved locations was very strong. The root mean square error was 0.07 indicating the model accuracy was good for assessing the average accuracy of the individual observations.

Conclusion

The results for this study provided a model based malaria risk maps for western Kenya at high resolution (0.5 X 0.5 km). Moreover, the study explored the underlying spatial and temporal processes that are the key causes of redistribution of malaria risk in the region. Additionally the study was able to determine the areas with high incidence of malaria risk and the number of people at risk.

From the results, warm and humid areas are experiencing high malaria transmission relative to colder and humid areas. Moreover areas with increased rainfall, temperature and high vegetation indices show increased malaria risk. This is in contrast with areas with moderately low average temperature of less than 20˚C which throughout the years experienced low malaria risk with low changes in population at risk over the period of study.

The results shows that counties which had increase in temperature resulted in increased risks of malaria as observed in Kericho County.

One of the limitations of this study was the uneven distribution of case based survey data which could have resulted in under prediction or over prediction in some areas. This is was mainly because many of the case based survey were concentrated in areas which have shown to experience high number of reported malaria cases over the years. Future studies should develop ways of accounting for uneven distribution of malaria survey data.

The malaria control and management strategies can be effectively formulated using the findings from this study. This will ensure that limited resources are utilized effectively and efficiently by focusing on the malaria redistribution areas as well as high risk areas instead of blanket application in entire region.

References

[1] Sachs, J. and P. Malaney "The Economic and Social Burden of Malaria," Nature 415 (2002)
[2] World Health Organization, World Malaria Report(WMR) 2013
[3] Cyril Caminade, Sari Kovatsc , Joacim Rocklovd , Adrian M. Tompkinse , Andrew P. Morseb , Felipe J. Colón-Gonzáleze , Hans Stenlundd , Pim Martensf , and Simon J. Lloydc, Impact of climate change on global malaria distribution, 2014
[4] President's Malaria Initiative, Malaria operation plan(MOP,2013) FY 2013
[5] Najera JA (1989). Malaria and the Work of WHO. Bulletin of the World Health Organization 67(3): 229-243
[6] P. Ceccato, S.j. Connor, I.Jeane, M.C. Thomson, Application of Geographical Information Systems and Remote sensing technologies for assessing and monitoring malaria risk, 2005.
[7] Yaw A. Afrane, Tom J. Little, Bernard W. Lawson, Andrew K. Githeko, and Guiyun Yan, Deforestation and Vectorial Capacity of Anopheles gambiae Giles Mosquitoes in Malaria Transmission, Kenya, 2008
[8] Andrew K Githeko, John M Ayisi, Peter K Odada, Francis K Atieli,, Bryson A Ndenga, John I Githure and Guiyun Yan, Topography and malaria transmission heterogeneity in western Kenya highlands: prospects for focal vector control, 2006
[9] Yaw A. Afrane, Andrew K. Githeko and Guiyun Yan, Malaria Transmission in the African Highlands in a Changing Climate Situation: Perspective from Kenyan Highlands, 2011Lawrence N Kazembe, Immo Kleinschmidt, Timothy H Holtz and Brian L Sharp, Spatial analysis and mapping of malaria risk in Malawi using point-referenced prevalence of infection data, 2006
[10] Lawrence N Kazembe, Immo Kleinschmidt, Timothy H Holtz and Brian L Sharp, Spatial analysis and mapping of malaria risk in Malawi using point-referenced prevalence of infection data, 2006
[11] Kenya National Bureau of Satistics, 2009 Kenya Population and Housing Census report.
[12] Havard Rue and Sara Martino, Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations, J. R. Statist. Soc. B (2009) 71, Part2, pp.319–392

Author Profile

Gilbert K. Sang is a GIS officer at Cedia Africa (Kenya). He holds B.Sc. (Honors) degree in Geomatic Engineering and Geospatial Information Systems from JKUAT (2011) and is about to complete his M.Sc. degree in GIS and Remote Sensing at the same University. He has a keen interest in application of GIS in Health, security, Agriculture, Environmental management, Urban and regional planning and Community Development

Dr Hunja Waithaka is a Senior lecturer in the Department of Geomatic Engineering and Geospatial Information Systems of Jomo Kenyatta University of Agriculture and Technology (Kenya). He holds a BSc. (Eng) Surveying & Photogrammetry, University of Nairobi (1993) and a M. Sc. (Earth & Planetary Sciences), Hokkaido University and Ph.D. (Earth & Planetary Sciences), Hokkaido University (2004). His area of specialization include Geo-informatics, Surveying, Geodesy, GNSS, Geodynamics, GIS, Remote Sensing and Land Management