Spatial Modelling of Malaria Risk Factors in Ruhuha Sector in the East of Rwanda

Tuyishimire, J.1, Karema, F.2, Mugisha, J.1, Amer, S.3, Mens, P.2
1Center for Geographic Information Systems and Remote Systems-University of Rwanda (CGIS-UR) 2University of Twente 3KIT Amsterdamb
tuyishimirejoseph@gmail.com, f.karema2011@gmail.com, s.amer@utwente.nl

ABSTRACT

Malaria is a vector borne disease posing a severe health risk to the population of Sub Saharan Africa and particularly in the East African Valley. Over the past decades, billions of dollars have been invested to combat malaria, but it is still killing hundreds of thousands of people annually. The aim of this study is to spatially model malaria risk factors in a primarily rural area situated in south eastern Rwanda. The underlying data for the research were obtained through a combination of high-resolution ortho-photography and primary data collection. In addition to this, detailed malaria occurrence and socio-economic data were made available. Spatial clusters of malaria occurrence were subsequently determined using Getis and Ord spatial statistics. This cluster analysis showed that malaria distribution is characterized by zones with high malaria risk, so called hot spots, zones with moderate malaria risk known as not significant spots and zones of low malaria risk known as cold spots. Preliminary outcomes demonstrate that malaria prevalence varies from one household to another and from one administrative unit (village) to another. The relationship between malaria prevalence and malaria risk factors was assessed using a logistic regression model. Results clearly indicate that malaria infection increases with the proximity to irrigated farmland. It also increases with household size. It was also proven that lower housing quality (mud houses; unburnt brick walls, earth floor) is associated with higher risk of malaria infection.

Key words: Spatial Modelling, Malaria risk, Anopheles habitat, malaria causing factors, Ruhuha Sector, Rwanda

I. INTRODUCTION

Malaria is a mosquito-borne infectious disease caused by protists of the genus Plasmodium which are introduced into the circulatory system by the bite from an infected female Anopheles mosquito (Smith et al., 2013). Human malaria is one of the most widespread endemic vector borne diseases throughout the tropical and subtropical regions of the World especially around the Equator (Ahmad et al., 2011). Many studies about malaria were done in different parts of the World. Most of them investigated the ecology of the vector (Smith et al., 2013), efficiency of control measures (Karema et al., 2012) and spatial disease modelling (Hay et al., 2006).

In Rwanda, studies about malaria mainly addressed malaria control policies and programmes. Thus, the focus was on the contribution of anti malarial drugs, mosquito nets and residual sprays (Karema et al., 2006; NISR, 2012a; President's Malaria Initiative, 2013), the prevalence of malaria in pregnant HIV positive women (Ivan et al., 2012) and the resistance of Plasmodium falciparum to antimalarial drugs (Karema et al., 2012).

Most all the studies done in Rwanda have established a comprehensive malaria framework, but they did not take into consideration geospatial factors that might affect malaria infection. This paper takes an explicitly geographical perspective by incorporating environmental factors and performing a geospatial analysis to identify environmental risk factors associated with malaria infection in Ruhuha.

The use of High resolution orthophotographs helped to identify potential anopheles vector breeding sites, Getis and Ord hot spots analysis is a useful geostatistical technique that can help in malaria clusters identification.

The information gathered through the combination of GIS tools and spatial statistics analysis, could allow the spatial analysis of malaria underlying factors and to show their relationship with malaria prevalence at small scale within the rural administrative Sector of Ruhuha.

II. MALARIA CONCEPTUAL FRAMEWORK

Malaria conceptual framework used in this study consists of five components (Figure 1):

1. Environmental variables. These variables are categorised into three groups as follows: climatic (temperature, rainfall, humidity), topographic (altitude and slope), and ecological (hydrology, wetlands, vegetation and soil). Together these factors represent the main determinants for anopheles mosquito breeding sites (Machault et al., 2011; Nagasaki University, 2007).
2. Vector life cycle. The vector life cycle is characterized by breeding and mortality of the vector itself, which together determine vector density.
3. The parasite cycle. The parasite cycle starts when humans are bitten by infected female anopheles mosquitoes. Infection is detected when the parasite multiply quickly in non-immune people.
4. Demographic and socioeconomic characteristics. The level of vulnerability is associated with demographic variables (household size, age and gender), socioeconomic variables (income, education)
5. Malaria control. Vector control consists of two methods. Natural control, which deals with environmental management and land use planning, and artificial control (use of ITNs and IRS).

III. MATERIALS AND METHODS

1. Study site

This study was carried in Ruhuha Sector which is one of the 13 administrative sectors of Bugesera district in the Eastern Province of Rwanda. Ruhuha Sector consists of 5
Ruhuha Sector is characterized by a succession of low plateaux, hills, dry valleys and swamps. The latter are mostly used for agricultural purposes or occupied by Papyrus Species and Cyperus species. (Especially near Cyohoha Lake). The micro-environment of these swampy areas represent a suitable habitat for anopheles mosquitos (Bugesera District, 2013). Because of the environmental conditions in Ruhuha Sector and the Eastern Province of Rwanda in general, malaria represents a serious health threat (NISR, 2012a). Malaria is the second cause of morbidity (23.3%) in the area, after the sharp infections of the superior or lower respiratory ways which represent 27.3% (Bugesera District, 2013). Its effect could have become higher unless control measures had been taken.

2. Data
   i. Data collection
This study is based on secondary data consisting of malaria infection, demographic and economic data that were provided by the Rwanda Biomedical Center/Malaria Research program (RBC/MRC) and on spatial data that were provided by the National Institute of Statistics of Rwanda (NISR). Primary data included identified anopheles mosquito breeding sites and land use. Table 1 shows all the collected data.

   • Secondary data
Secondary data were obtained from RBC/MRC in collaboration with the Royal Tropical Institute (KIT) through a research entitled “Empowering the Community towards Malaria Elimination”. Another section of secondary data consisted of spatial information obtained from the NISR. These data consisted of demographic household size, age and gender), economic (animal ownership and household material), Malaria control measures (artificial and natural approaches and treatment), Malaria infection and spatial data (sector, cells and village boundaries and water sources).

The demographic, economic and malaria control measures data were directly recorded in tablets (Samsung Galaxy tabs). In addition, location (GPS coordinates) was recorded for each household. The collected data were then uploaded to the RBC / MRC server where they were combined from different villages and cleaned for further uses. 3443 households were surveyed.

Spatial data consisted of the shapefiles of the study area (sector, cells and village boundaries and water sources). They were all obtained from the National Institute of Statistics of Rwanda (NISR).

   • Primary data
Primary data consisted of potential malaria breeding sites (irrigated farmland and water reservoir), land use and local people’s view about malaria causing factors and transmission.
With the Guidance of administrative cells CHWs, the coordinates and the elevation of each breeding site were marked using a hand-held Geographic Positioning System (GPS), GARMIN etrex 12 CHANNEL (Garmin Ltd, 2003) and the land use type was described. The captured data were later imported via map source 6.16.3 Software and processed in ArcGis 10.2 package.
3. Data analysis

- Data integration

The data from RBC consisted of laboratory results that were stored in laboratory notebook and in RBC/MRC general register and demographic, economic and malaria control measures that were stored on the online RBC/MRC server. A link code was created so that data from the three sources could be integrated. The same code was used to link data from the field, RBC and NISR.

- Steps for data analysis

  a. Potential anopheles mosquito habitat maps

The land use was delineated from an orthophotograph of the study area. An accuracy assessment of the produced land use map was carried out to check its reliability for further analyses (Phillips et al., 2006). Potential anopheles mosquito habitat map was created through the integration of the information extracted from the land use map and field observations (see figure 10).

  b. Malaria Prevalence and clustering

Malaria prevalence (ratio of infected people over the total population) was analyzed at village level. Getis and Ord (G*) index (Cromley & MCLafferty, 2012) was used to test the significance of malaria clusters within the surveyed households.

  c. Malaria causing factors

Malaria causing factors in the study area were determined from the literature and interviews with the local Community Health Works (CHWs) and the RBC/MRC staff in charge of malaria control. They were categorized in three groups: environmental, demographic and economic. The identified malaria causing factors were either nominal or ratio. The latter consisted of the altitude and household size the distance to Anopheles breeding sites. To know the distance from each of the observed households to the nearest vector breeding, distance maps were created.

  d. Relationship between malaria infection and malaria causing factors

The potential relevant variables for malaria prediction were screened through the process of collinearity analysis (Correlation coefficients and VIFs calculation).

  e. Regression model

The logistic regression model was chosen for this study because it has the capacity predicting categorical variables by nominal and ratio variables at the same time. The stepwise logistic regression model was used for testing the relationship between malaria infection (Presence of parasites in the blood=1, absence=0) and its causing factors and control measures. In fact, the stepwise regression helps to detect the model improvement if new predictors are added and is better for logistic regression with dummy variables (Field, 2009). Odds ratio (OR and Wald were used to test the significance of the predicted coefficients (Lubetzky-Vilnay et al., 2013) Getis and Ord (G*) index (Cromley & MCLafferty, 2012) was used to test the significance of malaria clusters within the study area?

f. Malaria causing factors

<table>
<thead>
<tr>
<th>Data</th>
<th>Specification</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orthophoto</td>
<td>25 cm resolution acquired in August 2008 and 2009</td>
<td>RNRA/DLM</td>
</tr>
<tr>
<td>Study area shapefiles</td>
<td>Rwanda baseline maps 2010</td>
<td>NISR</td>
</tr>
<tr>
<td>Household XY coordinates</td>
<td>GPS coordinates</td>
<td>RBC/MRC</td>
</tr>
<tr>
<td>Malaria infection data</td>
<td>Number of people with Plasmodium parasites per household</td>
<td>RBC/MRC</td>
</tr>
<tr>
<td>Number of people</td>
<td>Household level</td>
<td>RBC/MRC</td>
</tr>
<tr>
<td>Population density</td>
<td>Cell and village levels</td>
<td>NISR</td>
</tr>
<tr>
<td>Age and sex of the population of study</td>
<td>Presence of animals in the household</td>
<td>RBC/MRC</td>
</tr>
<tr>
<td>Animal ownership</td>
<td>The material that makes the roof, walls and the floor</td>
<td>RBC/MRC</td>
</tr>
<tr>
<td>House material</td>
<td></td>
<td>RBC/MRC</td>
</tr>
<tr>
<td>Use of control measures</td>
<td>Number of ITNs per Household, sprays, cutting bushes, clearing stagnant water and others</td>
<td>RBC/MRC</td>
</tr>
<tr>
<td>Anopheles micro-habitat in Ruhuha</td>
<td>GPS coordinates and elevation of Anopheles breeding sites</td>
<td>Field work</td>
</tr>
<tr>
<td>Land use of the study area</td>
<td>Training point for accuracy assessment</td>
<td>RBC/MRC</td>
</tr>
</tbody>
</table>

Table 1. Collected data
Malaria causing factors in the study area were determined from the literature and interviews with the local Community Health Works (CHWs) and the RBC/MRC staff in charge of malaria control. They were categorized in three groups: environmental, demographic and economic. The identified malaria causing factors were either nominal or ratio. Ratio factors consisted of the altitude and household size the distance to Anopheles breeding sites. To know the distance from each of the observed households to the nearest vector breeding, distance maps were created.

g. Relationship between malaria infection and malaria causing factors

The potential relevant variables for malaria prediction were screened through the process of collinearity analysis (Correlation coefficients and VIFs calculation).

h. Regression model

The logistic regression model was chosen for this study because it has the capacity predicting categorical variables by nominal and ratio variables at the same time. The stepwise logistic regression model was used for testing the relationship between malaria infection (presence of parasites in the blood=1, absence=0) and its causing factors and control measures. In fact, the stepwise regression helps to detect the model improvement if new predictors are added and is better for logistic regression with dummy variables (Field, 2009). Odds ratio (OR and Wald were used to test the significance of the predicted coefficients (Lubetzky-Vilnai et al., 2013)

IV. RESULTS AND DISCUSSION

1. Malaria prevalence

Compared to the country statistics (1.2%) as described by NISR (2012b), the study area was characterized by a high malaria prevalence (2.7%). A large number of households were characterized by one infected individual. Malaria is differently distributed in the study area with the highest prevalence in Gikundamvura Cell (See Figure 2). According to local CHWs, the higher malaria prevalence in Gikundamvura cell could be justified by the fact that a large proportion of people in Gikundamvura are far from Ruhuha Health Center and have houses of low quality. Malaria prevalence at village level is heterogeneously distributed as well (Figure 2). Most of the villages in the study area are characterized by a high malaria prevalence compared to the national statistics. Gikundamvura, Rukurazo, Kanombe, and Kazabagarura of Gikundamvura Cell; Nyaburiba of Gatanga Cell; and Kamweru of Kindama Cell, Rwanzunga of Bihari Cell and Mubano of Ruhuha Cell are characterized by higher malaria prevalence (> 5%). Compared to the WHO (2008) report suggesting that areas characterized by more than one malaria case per thousand people are at high risk, all the villages of the study area are at high risk of malaria.

2. Malaria clustering

The biggest number of hot spots was observed in Gikundamvura cell. Ruhuha cell is the cell with less malaria cases (Figure 3). The probability of getting malaria is higher (Zone characterized by a high prevalence) near the place where the observed prevalence was higher too. This distribution map is realistic given the fact that humans are at the same time hosts and reservoirs of Plasmodium Sp. In accordance with Tobler (1970), malaria distribution in the study area is also justified by the principle of spatial autocorrelation by which neighboring people share the same characteristics such as malaria causing factors and
prevention measures and are therefore under the same risk, unless different intervention measures are relied upon.

3. Land use of the study area

The land use of the study area was classified in six classes: Open agriculture consisting mainly of perennial crops, closed agriculture made of banana and coffee plantations, forest made mostly of eucalyptus plantations, lake, settlement and irrigated farmland where the predominant crop was rice (Figure 5).

The overall accuracy and the kappa (K*) value were calculated. In accordance with Phillips et al. (2006), the accuracy assessment proved a strong agreement (K*=90%, overall accuracy=91%) and the produced land use map was therefore reliable for future analyses.

4. Potential Breeding Sites

Potential anopheles mosquito habitats in the study area consist of irrigated farmlands, water reservoirs, closed agriculture, forest plantation and the lake (See Figure 4). Anopheles mosquito depends on blood meal and accordingly, its habitat combines the breeding and resting sites.

Irrigated farmlands and water reservoirs are places where ecological conditions are favourable for adult mosquito breeding and larva survival. This was confirmed by Wielgosz et al. (2012) in for East Africa and Yamamoto et al. (2010) in Burkina Faso. In accordance with the findings of Stoler et al. (2009) in Accra and Munga et al. (2006) in the Western Kenya, irrigated farmlands and water reservoirs were considered as breeding sites and habitats at the same time because mosquitoes could stay there before their migration for blood meal search. The intermediate habitats (forest plantations, closed agriculture and vegetation on the lake shore) may serve as resting sites during the mosquito flights.

5. Relationship between malaria infection and underlying factors

The relationship between malaria and its underlying factors was analyzed through a stepwise logistic regression where the Wald statistics and the OR proved their significance (See Table 3). The logistic regression model, showed a low predictive power (R²=13%). This is explained by the fact that the chosen variables are not the only malaria underlying factors in the study area.

- Malaria infection towards demographic variables
  Malaria infection increases with the household size (Wald=77.3 and OR>1). The increase of malaria with household size is realistic in the study area because the household size varied between 1 and 14 people with a mean of 4 individuals. In general, more people in the study area live in low income and large families with bad quality houses and are therefore prone to malaria. In conformity with the results of this study, the high household size was also suggested among malaria underlying causes in the fringe zones of Kigali City by Bizimana et al. (2009). The association between malaria infection and the household size was also found by WHO (2008) in the annual malaria report while in their review on the persistence of malaria problem, Stratton et al. (2008) mentioned poverty and increasing population as associated problems which are among the main causes of malaria in developing countries.

- Malaria infection towards economic variables
  House materials are the only economic malaria predictors (see Table 3). Compared to houses with medium quality walls (reference category during the process of dummy coding), the infection is lower in houses made of high quality walls (Wald=5.7, OR<1) and higher in houses with low quality walls (Wald=39.9, OR>1). Malaria infection is lower in houses made of cement floor (Wald=31.3, OR<1) compared to those with earth floor (reference category).

Malaria infection was lower in houses with high quality walls (burnt bricks and cement blocks). Few houses of good quality made especially of burnt bricks or cement...
blocks for walls and tiles or iron sheets for roofs, have tight doors and windows so that mosquito could not enter easily except when doors or windows are open. Conversely, most of houses in the study area are made of medium quality (unburnt bricks) walls which in most cases have windows and doors that are not tight enough to block the mosquito entrance. In addition, a large proportion of houses with low quality (mud) walls were also proved positively associated with malaria infection. The quality of mud wall houses is the worst and they favour an easy contact between the mosquito and humans.

The relationship between poor housing quality and malaria infection that was found in this study is in confirmation with Stratton et al. (2008) in their review on the persistent problem of malaria. The same relationship is in contradiction with the results of Yamamoto et al. (2010) which suggested a lack of relationship between malaria prevalence and house material in the semi-urban areas of Burkina Faso.

The results of this study showed that:

- Malaria infection is influenced by those land use types that serve as anopheles mosquito habitats (breeding and resting sites).
- Malaria prevalence (number of infected people/the whole population) is higher in the study area compared to the country statistics and is differently distributed in the surveyed villages and cells. Malaria prevalence is higher in less populated areas which are characterized by a big number of clusters.
- Malaria decreases with the distance to irrigated farmland. Houses with earth floors are characterized by a higher malaria infection compared to those with cement floors. Malaria infection is higher in low quality (mud) walls and lower in high quality (burnt bricks and cement blocks) walls.

The results of this study need to be complemented by anopheles entomological data so that malaria prevalence could be integrated with the vector spatial pattern. This integration will help decision makers to deal with all aspects of malaria transmission in the study area.

### Table 2. Relationship between malaria prevalence and causing factors

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>S.E.</th>
<th>Wald</th>
<th>df</th>
<th>P-value</th>
<th>OR</th>
<th>95% C.I. for OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dist_marsh</td>
<td>-1.005</td>
<td>.235</td>
<td>18.238</td>
<td>1</td>
<td>.000</td>
<td>.366</td>
<td>.231 - .580</td>
</tr>
<tr>
<td>Dist_HH</td>
<td>.463</td>
<td>.110</td>
<td>17.825</td>
<td>1</td>
<td>.000</td>
<td>1.589</td>
<td>1.282 - 1.970</td>
</tr>
<tr>
<td>H_size</td>
<td>.253</td>
<td>.029</td>
<td>77.368</td>
<td>1</td>
<td>.000</td>
<td>1.288</td>
<td>1.217 - 1.363</td>
</tr>
<tr>
<td>High</td>
<td>-.487</td>
<td>.618</td>
<td>5.788</td>
<td>1</td>
<td>.016</td>
<td>.226</td>
<td>.067 - .759</td>
</tr>
<tr>
<td>Low</td>
<td>.782</td>
<td>.124</td>
<td>39.968</td>
<td>1</td>
<td>.000</td>
<td>2.186</td>
<td>1.716 - 2.786</td>
</tr>
<tr>
<td>Cement</td>
<td>-.942</td>
<td>.200</td>
<td>22.243</td>
<td>1</td>
<td>.000</td>
<td>.390</td>
<td>.264 - .577</td>
</tr>
<tr>
<td>Dist_HC</td>
<td>-1.013</td>
<td>.174</td>
<td>33.918</td>
<td>1</td>
<td>.000</td>
<td>.363</td>
<td>.258 - .511</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.591</td>
<td>.339</td>
<td>58.499</td>
<td>1</td>
<td>.000</td>
<td>.075</td>
<td></td>
</tr>
</tbody>
</table>

$R^2 = 0.13$ (Nagelkerke). Model $\chi^2 = 207$, $p < 0.001$

### V. CONCLUSIONS

The results of this study showed that:

- Malaria infection is influenced by those land use types that serve as anopheles mosquito habitats (breeding and resting sites).
- Malaria prevalence (number of infected people/the whole population) is higher in the study area compared to the country statistics and is differently distributed in the surveyed villages and cells.

### I. References


