**R Project**

**Biostatistics 6640 Python and R Data Science**

**Yuli Chen**

Introduction:

Malaria is a major global health problem responsible for 438,000 deaths in 2015 with 80% of cases and 90% of deaths occurring in sub-Saharan Africa.1 Malaria is caused by parasites of the *Plasmodium* family*,* with *Plasmodium falciparum* being responsible for most of the malaria cases in Africa and is spread to humans by 30 different species of mosquitos in the *Anopheles* genus, mainly *Anopheles gambiae* and *Anopheles arabiensis*. When left untreated, the parasite invades and destroys red blood cells, leading to severe anemia, long-term neurological consequences, and death in severe cases. Partial immunity to the disease is acquired via surviving through repeated infections but do not develop until adulthood.1 Children are especially prone to severe malarial disease due to lack of acquired immunity.1

Sub-Saharan Africa is a high transmission area where children under 5 years are especially prone to severe disease and death. The WHO reported an estimated 86% of malaria deaths occurred in this group in 2000.1 Severe anemia, hypoglycemia, and cerebral malaria are all common manifestations of severe malaria and more commonly seen in children than in adults.1 Since malaria has been established as a priority in the 2000 United Nations General Assembly and designated as one of the Millennium Development Goals, the global burden of malaria has decreased substantially. It has been estimated that 6.2 million malaria-related deaths have been averted, the incidence of malaria cases has decreased globally by 37%, malaria deaths has decreased by 60%, and malaria deaths among children under 5 years has decreased by 65% between 2000 and 2015.1 This success can be attributed to preventive measures through wide distribution of insecticide-treated mosquito nets and indoor residual spraying, two major and cost-effective interventions to fight malaria.

Despite progress in reducing malaria burden, it remains to be a major health concern in sub-Saharan Africa as a disproportionate proportion of malaria cases (80%) and deaths (91%) still occurring in this region, with children under five bearing the largest burden.1 Mozambique is among the 15 countries in the world with the highest disease burden and has experienced slowed decline in malaria incidence (37%) compared to other countries (54%).1 Additionally, not enough work has been done to assess the impact of weather on malaria incidence to better inform intervention implementation.

Climate is an important factor in malaria transmission affecting both the parasite and mosquito populations. Previous studies have established links between temperature and rainfall to malaria transmission. Temperature and rainfall are determining factors in mosquito reproduction and mortality.2 A study by Imbahale *et al* estimated that the mean temperature ideal for the development of the mosquito vectors is 25-27 °C.3 Intermittent rainfall with periods of sunshine provides the ideal breeding conditions necessary for mosquito vector proliferation while continuous rains results in flooding that clears out mosquito breeding sites.4  Research has also shown that an increase in the population of female *An. Gambiae,* a species of mosquito carrying the malaria vector, is correlated with rainfall during the previous week due to 7-14 day incubation period after exposure to an infected bit.5 This report aims to explore the lagged relationship between under-five malaria incidence and climate patterns in Mozambique in hopes that it will better inform future efforts to further reduce the impacts of malaria on children under 5, the most vulnerable group.

Methods:

The dataset used for this study was obtained in a CSV file from Dr. Colborn, the instructor for the Biostatistics 6640 course at the University of Colorado. All analysis was done using R statistical software version 3.5.1. The raw dataset contained a number of malaria cases, rainfall, and other climate data, and geographic data on a district level. All data were collected from January 2010 to July 2017. Data between January 2010 to December 2016 were included in this analysis due to incomplete data for 2017. Malaria cases were reported at the district level weekly. Spatial and geographic information were also reported including, province, region, latitude and longitude data. Under five malaria incidence per 1000 was created using census information including the proportion of the population under 5 years and total population. Climate data used for this report include total weekly rainfall measures, average temperature, a number of days temperature was above 35 and below 15 degrees Celsius. Lagged rainfall and average temperature data were created at 2, 4, and 8 weeks to demonstrate the lagged relationship between exposure of an infected bite and the onset of symptoms. A series of maps were created to explore geographic distributions of under 5 malaria incidence, total rainfall, and average temperature (Figure 1-3). Column scatter plots were compiled to show relationships between under 5 malaria incidence, total rainfall, average and extreme temperatures by region (Figure 4 and 5). To demonstrate the lagged relationship between under 5 malaria incidence and climate patterns, a series of line plots were created (Figure 6 and 7).

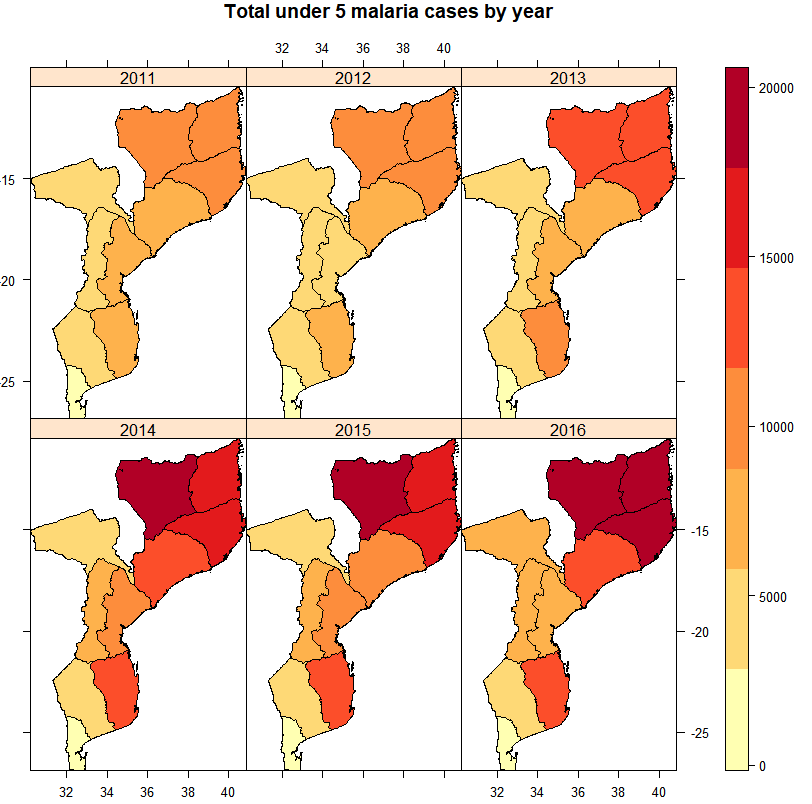
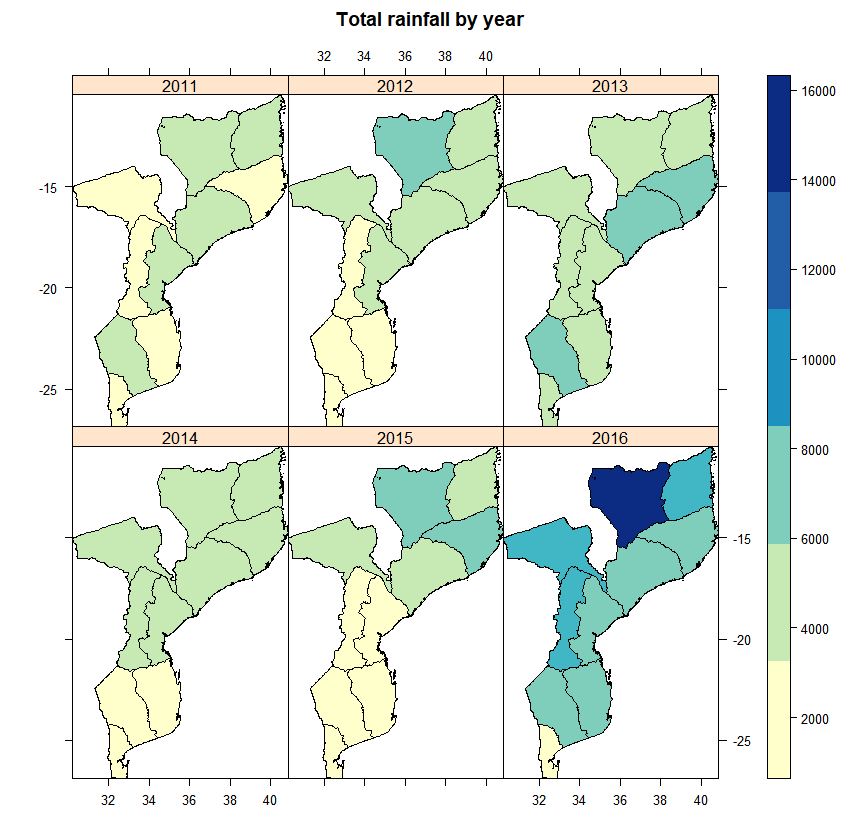
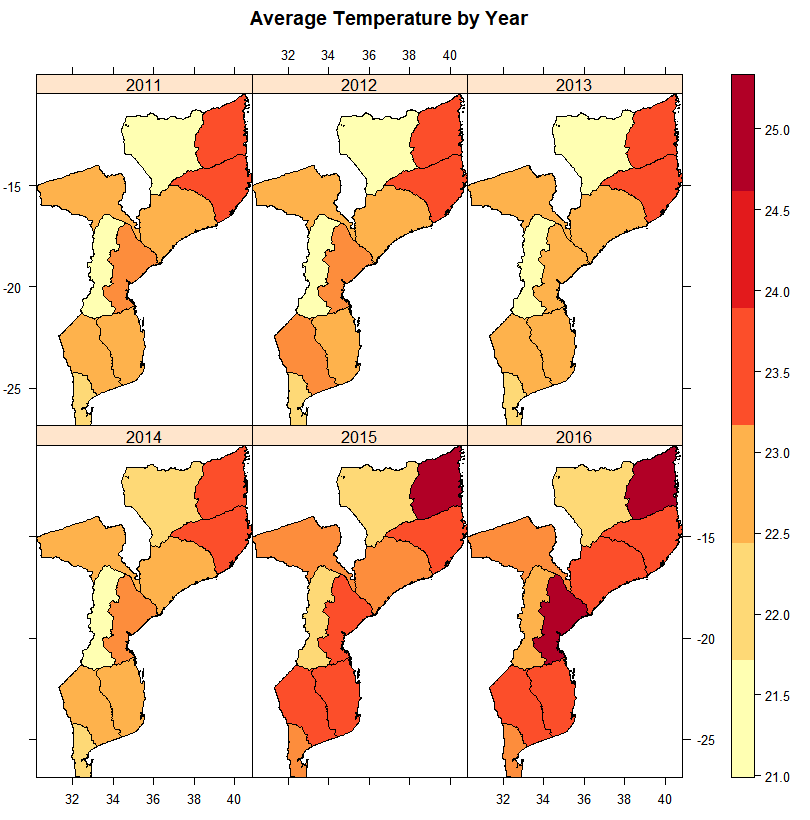
Results:

**Geographic Distribution of Malaria Incidence, Rainfall, and Average Temperature**

To explore the geographic relationship between malaria incidence among children under 5 and climate patterns, a series of maps were created in Figure 1-3. All maps show Mozambique divided into its ten provinces with malaria incidence (Figure 1), total rainfall (Figure 2), and average temperature (Figure 3) plotted by year.

Based on this map series, we observe that malaria incidence among children under 5 years is not equally distributed in Mozambique. Northern and coastal regions of the country experience higher incidence. This pattern is consistent during the extent of the years studied, 2011-2016. Incidence has also increased from 2011 to 2016 (Figure 1). Similarly, rainfall is also geographically clustered in Mozambique. Total rainfall is greater in the northern and coastal regions of Mozambique and lowest in southern and eastern regions of the country, except for 2016 (Figure 2).

Similar to the geographic distribution of under 5 malaria incidence, rainfall is also geographically clustered in Mozambique. Total rainfall is greater in the northern and coastal regions of Mozambique and lowest in southern and eastern regions of the country, with the exception of 2016 (Figure 2).



**Figure 1(Top Left):**

Choropleth map of Mozambique displaying malaria incidence among children under 5 years (total cases per 1000) by year and province. Light yellow represents lowest incidences and dark red represents highest incidence.

**Figure 2(Top Right):**

Map of Mozambique displaying total rainfall in each province by year. Light yellow represents low rainfall and dark blue represents regions with highest rainfall.

**Figure 3(Bottom Left):**

Map displaying average temperature of each province by year. Light yellow represents lowest average temperature and dark red represents regions with highest average temperatures.

Overall, the maps show a general trend of higher malaria incidence among children under 5, rainfall, and the average temperature in the northern coastal provinces of Mozambique. Malaria incidence, rainfall, and average temperature have all increased between 2011 and 2016.

To further examine malaria and climate patterns, a series of column scatter plots were produced by year and depicted by region (Figure 4 and 5). Figure 4 shows under 5 malaria incidence, total rainfall, and average temperature (y-axis) by year (x-axis). Regions were color-coded. There is an upward trend with malaria incidence. Northern and coastal regions tend to have higher incidence while central and southern regions generally have lower malaria incidence. Total rainfall seems to be constant, however, the range of rainfall is greater in 2016 with greater number of districts having relatively high total rainfall compared to 2010. Average temperature increases between 2010 and 2016. The central and northern regions of Mozambique experiences lower average temperature while coastal regions tend to have the highest average temperature for all years.



**Figure 4.** Column scatter plots showing malaria incidence among children under 5 years (top panel), total rainfall (center panel), and average temperature (bottom panel) plotted by year and classified by region. Each point represents a single district.

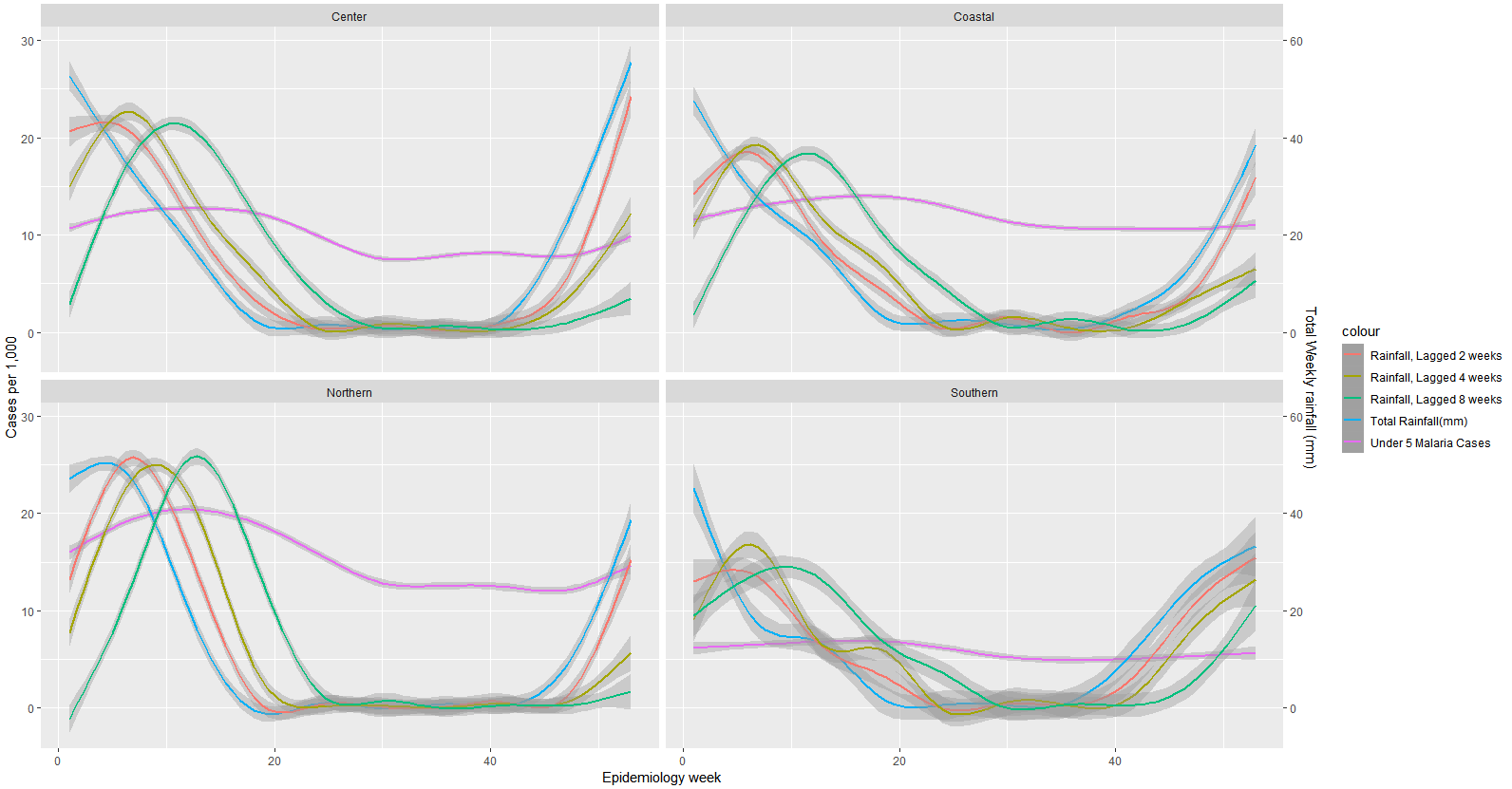
Studies have suggested a role of extreme temperatures on mosquito vector survivorship. Figure 5 shows column scatter plots of malaria incidence, average temperature, high temperature (above 35 degrees Celsius), and low temperature (below 15 degrees Celsius) by year and region. Malaria incidence and average temperature trends were described above. Figure 5 shows the number of days with temperatures above 35 degrees Celsius was relatively constant between 2010 and 2016. However, there are clusters of central and coastal districts that experienced a greater number of high temperatures. The number of days with temperatures below 15 degrees Celsius was also relatively constant between 2010 and 2016. Northern and central districts tend to have greater number days with low temperatures.



**Figure 5.** Column scatter plots showing malaria incidence among children under 5 years (top panel), average temperature (2nd panel), number of days with temperatures above 35 degrees Celsius (3rd panel), and number of days with temperatures below 15 degrees Celsius (bottom panel) plotted by year and classified by region. Each point represents a single district.

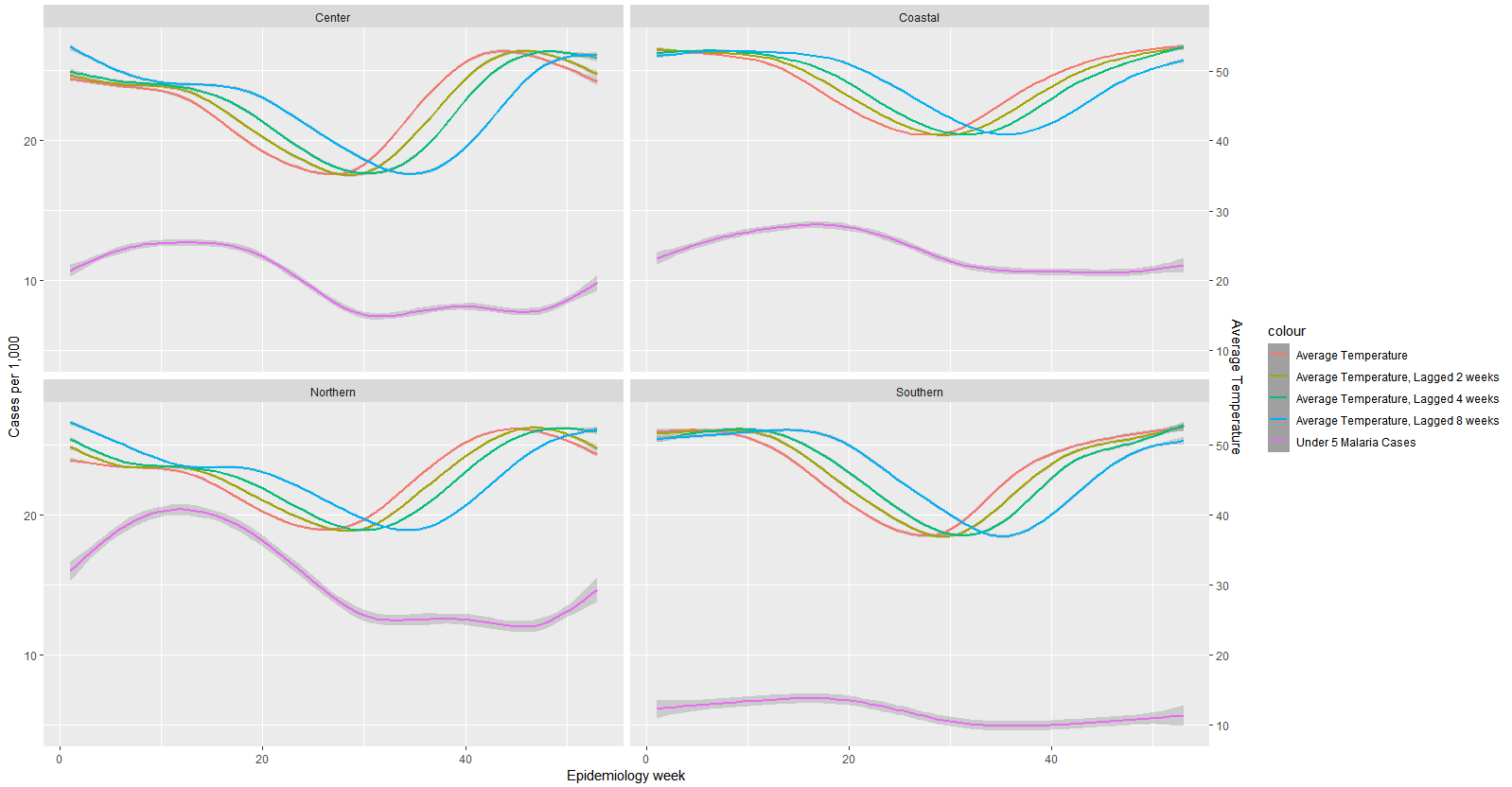
**Exploring Lagged Relationship between Malaria Incidence, Rainfall, and Average Temperature**

Research has shown an incubation period of 7 to 14 days between exposure to an infected mosquito bite and the onset of malaria symptoms. Here, we explore the length of lag on malaria incidence patterns. Figure 6 shows incidence, total rainfall and lagged total rainfall (by 2, 4, and 8 weeks) by epidemiology week and region. Highest incidence in all four regions is most closely correlated with a lag a lag of 8 weeks.



**Figure 6.** Malaria incidence among children under 5 plotted (pink line) with total rainfall and lagged rainfall at 2, 4, and 8 weeks.

Figure 7 shows incidence, average temperature, and lagged average temperature (by 2, 4, and 8 weeks) by week and region. High and low average temperatures seem to be correlated with lower incidence while moderate temperatures are correlated with higher incidence. However, visual observation of this does not seem to identify an obvious lab that is most associated with malaria incidence.



**Figure 6.** Malaria incidence among children under 5 plotted (pink line) with average temperature and lagged average temperature at 2, 4, and 8 weeks.

Discussion:

This report aims to explore the relationship between malaria incidence among children under five and climate patterns measured by total rainfall and average temperature. Descriptive maps and scatter plots suggest a geographic distribution of high malaria cases, higher rainfall, and higher temperature in northern and coastal regions of Mozambique. Despite a global decrease in the burden of malaria, initial results from this report suggest that malaria incidence under 5 has consistently increased since 2010 in Mozambique. This is consistent with the WHO report of children under 5, particularly in sub-Saharan countries being a vulnerable population.

Upon visual inspection of lagged rainfall and temperature patterns suggests a longer than expected lag time between rainfall and spike in malaria incidence. Rainfall creates ideal breeding environments for mosquitos. It is generally thought that the incubation period is 7-14 days, this is the time between an infected bite and onset of malarial symptoms. However, the lag time in rainfall and incidence may be lengthened due to the amount of time required for the development of mosquito vectors.

This report is an exploratory and descriptive analysis showing relationships between climate patterns and malaria incidence. Further research on the role of climate change on malaria incidence, especially among the most vulnerable population is needed to fully understand slowed progress in reducing malaria burden among children under 5 in sub-Saharan Africa. This report does not include any statistical analysis to test the significance of differences in climate patterns, malaria incidence, and geographic distribution. Additional analysis is needed to describe whether these patterns are significant.

Note: I collaborated with Anjin Singh and Ana Babinec on this project

References:

1. World Health Organization | UNICEF. (2015). Achieving the malaria MDG target: Reversing the incidence of malaria 2000-2015. Accessed from <http://apps.who.int/iris/bitstream/handle/10665/184521/9789241509442_eng.pdf;jsessionid=EB3E8DC8AA1B9934B0B18EDFA8C45004?sequence=1>
2. Lindsay, S.W.; Birley, M.H. Climate change and malaria transmission. *Ann. Trop. Med. Parasitol.* 1996, *90*, 573–588.
3. Imbahale, S.S.; Paaijmans, K.P.; Mukabana, W.R.; van Lammeren, R.; Githeko, A.K.; Takken, W. A longitudinal study on anopheles mosquito larval abundance in distinct geographical and environmental settings in Western Kenya. *Malar J.* 2011, *10*.
4. Martens, P.; Kovats, R.S.; Nijhof, S.; de Vries, P.; Livermore, M.T.J.; Bradley, D.J.; Cox, J.; McMichael, A.J. Climate change and future populations at risk of malaria. *Glob. Environ. Chang.* 1999, *9*, S89–S107.
5. Koenraadt, C.J.; Githeko, A.K.; Takken, W. The effects of rainfall and evapotranspiration on the temporal dynamics of *Anopheles gambiae sensu stricto*. and *Anopheles arabiensis* in a Kenyan village. *Acta Trop.* 2004, *90*, 141–153.