R Project: Mozambique Study

Kaci Pickett

October 22, 2018

## Background

Malaria is a parasitic illness transmitted by mosquitos in tropical and subtropical areas. Infection can be moderate to severe and presents as flu like symptoms such as fever, chills, nausea and vomiting, and body aches. Severe cases can lead to death and include more extreme symptoms such as organ failure, anemia due to hemolysis and neurologic defects especially in children. Typically malaria is curable if diagnosed and treated quickly and correctly, but in areas with high incidence prevention of infection is more effective.  
In order to effectively prevent malaria, insecticide can be sprayed and bed nets can be distributed, but the efficacy of these measures decays over time. Due to this issue there is a need to understand the life cycle of malaria and what affects levels of incidence to be able to predict the best time to apply insecticide measures.

The life cycle of malaria parasites is split between a human host and the Anopheles mosquito. The parasite starts its life in a human liver as a sporozoite delivered by a mosquito, invades the bloodstream where it replicates and matures causing clinical manifestations of disease. The gametocytes are then ingested by a mosquito during a blood meal where the male and female gametocytes merge to form an oocyst that releases sporozoite to start the cycle over again. For this cycle to happen, an ideal environment must be available for mosquitos to breed, survive long enough to have a blood meal containing sporozoites, and incubate the parasite during the complete growth cycle. This means that adequate rainfall for water collection must happen to allow a 9 to 12-day mosquito development. The mosquito must then survive long enough to have a blood meal and incubate the parasite for 9 to 21 days to maturation. This requires temperatures of 25°C or higher as well as adequate humidity and rain. Once a human is infected with the parasite the incubation from liver to blood stream when symptoms present ranges from 7 to 30 days.

The complex life cycle shown above illustrates that the issue of determining what weather trends are associated with observed incidence rates takes diligence. From ambient weather to human symptoms the time gap can be anywhere from 2 weeks to 16 weeks or possibly longer. Understanding what this relationship is, and how weather patterns act in certain areas will allow better knowledge of when and where to implement preventative measures.

## Introduction

In this study information about weather trends and Malaria was collected from 142 districts in Mozambique. The goal of the analysis was a preliminary review of the relationships between weather trends and malaria incidence in the population under 5 years of age with a long-term goal of eventually implementing a prevention plan in the most effective way. Data was collected at the district level each week from the years of 2010 to 2016, with partial information available for 2017. Weather variables of interest included weekly data on rain totals, average temperatures, relative humidity, and surface barometric pressure. Population of the entire district was given along with the proportion of the population under 5 years of age, and the number of cases of malaria reported each week. Each district was also categorized into 11 provinces and 4 regions to allow for spatial pattern analysis. The main goal was to determine if malaria incidence was variable over time or space and how this variability was affected by weather patterns.

## Methods

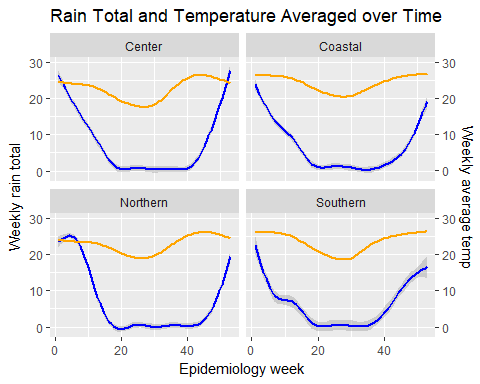
Malaria incidence was represented by cases per 1000 population in children under 5. This was calculated using malaria cases in the under 5 population divided by the product of total population of the district and proportion of population under 5. This was calculated at the district level at each week and as an average over the weeks for each district. Weather variables including total weekly rain total and average weekly temperature were tested for optimal lag time compared to malaria incidence using a cross correlation function and inspection of behavior of different lags versus incidence over time. Relationships between and within weather variables, incidence and time were explored graphically to determine trends.

*Table 1. Descriptive Statistics by Year*

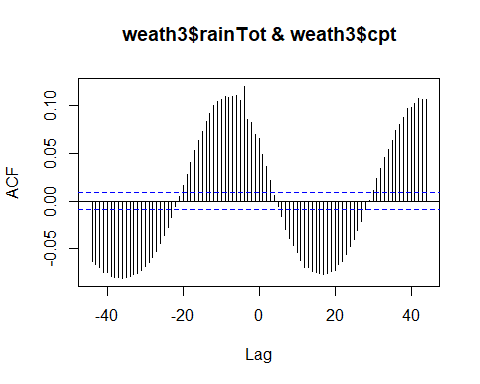
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Year | Cases/1,000 U5 | Median Rainfall | Average Temp (C) | Relative Humidity |
| 2010 | 9.84 | 0.01 | 22.91 | 77.20 |
| 2011 | 9.05 | 0.00 | 22.89 | 75.08 |
| 2012 | 8.87 | 0.00 | 23.01 | 74.37 |
| 2013 | 10.46 | 0.01 | 22.93 | 74.17 |
| 2014 | 14.23 | 0.00 | 23.07 | 74.55 |
| 2015 | 14.20 | 0.00 | 23.80 | 67.78 |
| 2016 | 15.64 | 0.75 | 23.92 | 67.78 |

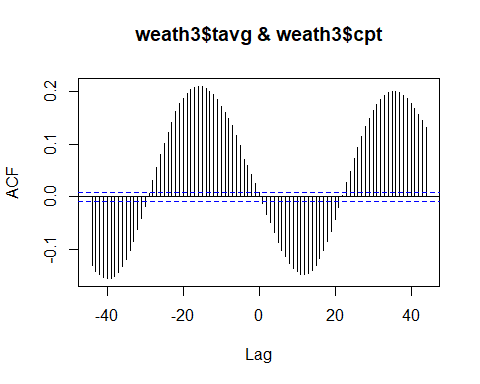
## Results

A plot of the smoothed behavior of weekly average temperature revealed that rainfall and temperature share a similar trend but are not completely associated so the lag time for rainfall compared to incidence may not be identical to that of temperature. Figure 1 shows that the first few weeks and the last few weeks of the year tend to be the rainy season and midyear is both drier and cooler, but temperature tends to increase before rainfall does.

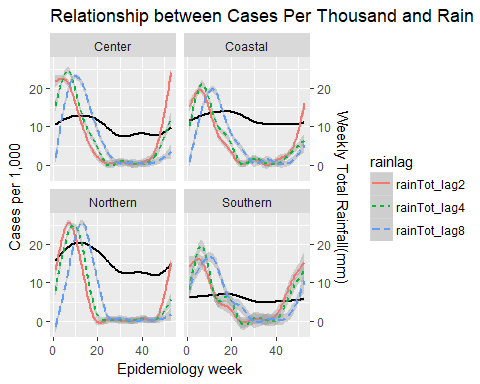
*Figure 1. Smoothed Plot of Temperature and Rain over time* 

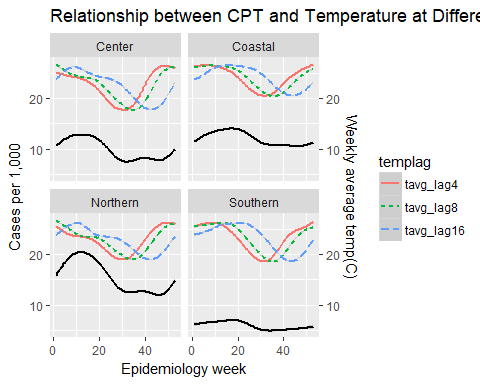
When observing the cross-correlation functions for both rainfall and temperature compared to Cases Per Thousand (CPT) it appears that the optimal lag time for rainfall occurs around 4 to 8 weeks before incidence and the temperature lag that is most correlated is between 8 and 16 weeks (Figures 2-3).

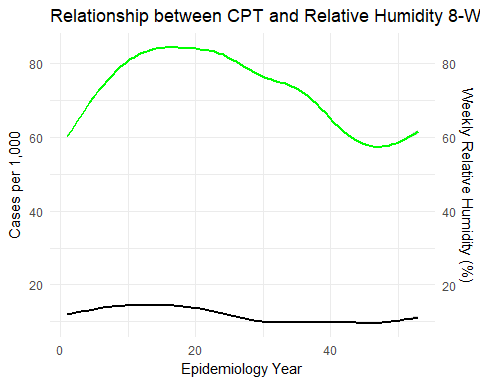
*Figure 2. Cross Correlation of Cases per Thousand and Weekly Rain Total* 

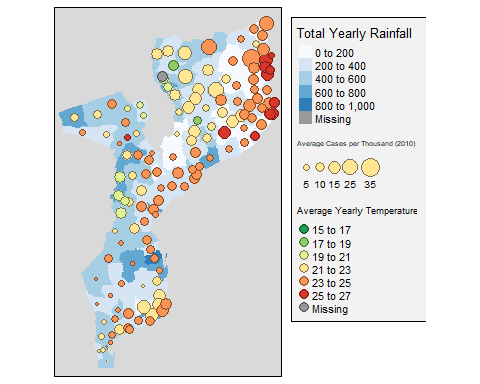
*Figure 3. Cross Correlation of Cases per Thousand and Weekly Average Temperature* 

This can also be seen when plotting smoothed curves of a few of the most clinically relevant lags overlaid with CPT. The lag of 8 weeks appeared to have the most similar peaks and decreases to CPT across the regions for both rainfall and temperature(Figure 4-5). This relationship can be seen with other weather variables as well as Figure 6 Relative Humidity as a percentage has a similar 8 week lagged relationship with CPT. All further mapping and plotting was done with a lag of 8 weeks for both rain and temperature when associations with CPT were considered.

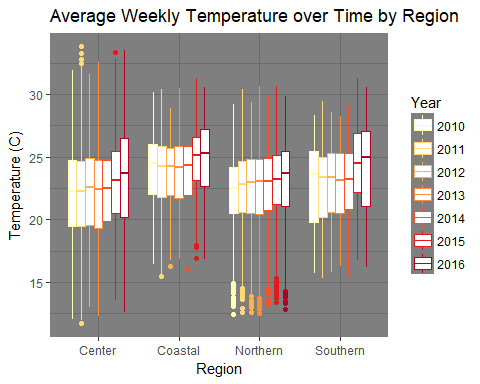
*Figure 4. Smoothed curves of CPT and Different Lags for Weekly Rain Total* 

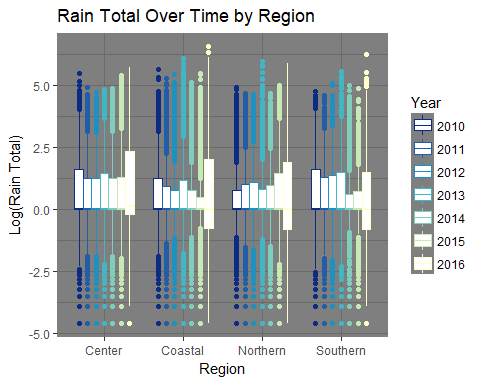
*Figure 5. Smoothed curves of CPT and Different Lags for Weekly Average Temperature* 

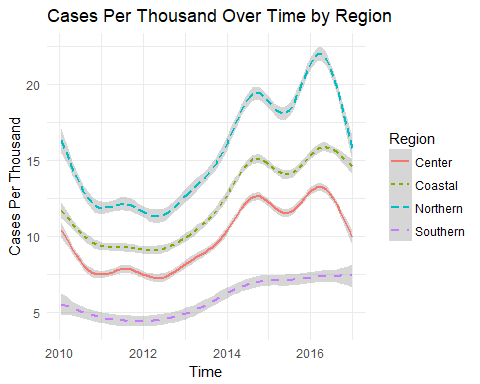
*Figure 6. Smoothed curves of CPT and Different Lags for Weekly Rain Total*  A map depicting average CPT in 2010 shows that Malaria incidence is different across regions with the Coastal and Northern regions having the highest amounts of clustering (Figure 7). This is very interrelated with large amounts of rainfall and temperatures ranging from 21 to 27. It appears as if the temperatures that are warmer than 21 but lower than 25 create the best environment for incidence. Rainfall seems to be necessary in small amounts as there seems to be less instances in areas with a lot of rainfall. The trend in 2010 is noticeable so it would be of interest to see how this trend changes over the 7 years of data.

*Figure 7. Map of Weather and Incidence Patterns by District in 2010* 

Temperature trends tended to have similar ranges across regions with slightly warmer temperatures noticeable in the Coastal and Center regions. These trends are not highly variable over the years but a slight increase in the median temperatures can be seen (Figure 8). A graph of rainfall with a log transform on all values other than 0 was also performed and showed now real difference in the average weekly rainfall in different districts and showed only slight increases in rainfall in the Northern region over time (Figure 9).

*Figure 8. Boxplot of Temperature by Year* 

\*Figure 9. Boxplot of Weekly Rainfall by Year  Over the last 7 years counts of Malaria were relatively constant up until 2014 when increasing cases were seen and it appears as if there were outbreaks in mid-2014 and early 2017(Figure 10). This trend was consistent across region but was most noticeable in the Northern and Coastal regions where the highest incidence was seen.

*Figure 10. Cases Per Thousand Over Time by Region* 

## Conclusions

In general these graphs gave good insight into the fact that it is possible to predict a relationship between Malaria incidence and weather patterns. Warm areas with moderate amounts of average rainfall tended to breed to most mosquitos that transmit Malaria. It appears that seasonal patterns should be considered and the times early and late in the year with warm wet periods should be the time when preventetive measures should be planned 8 to 16 weeks after these weather patterns.

A consideration that may need to be made is if the relationship between the CPT incidence rate and weather patterns changes over the years since the graphs that were performed today were mainly performed for values averaged across years. Supplemental materials show smoothed plots of the relationship between total weekly rainfall and average temperature across the years to show there may be some differences in the trend over the years(Appendix A).

The map comparing the clustering of cases compared to rainfall and temperature showed that extreme temperatures or large amounts of rain tended decrease incidence. This showed that the cultivation of the Malaria requires enough warmth and moisture to allow both the parasite and the mosquito to thrive but large amounts of rain or extreme temperatures threaten this process. Careful consideration needs to be taken in how to classify rainfall. Since seasonal patterns are present using weekly averages for rain instead of total rainfall may mask some of the rain effects areas with spurts of dry time and others of downpours will appear similar to areas where there is consistent small amounts of rainfall. In future investigation, it may be interesting to look at average rainfall excluding days with no rain in order to determine the effect of frequency versus quantity of rain on incidence.

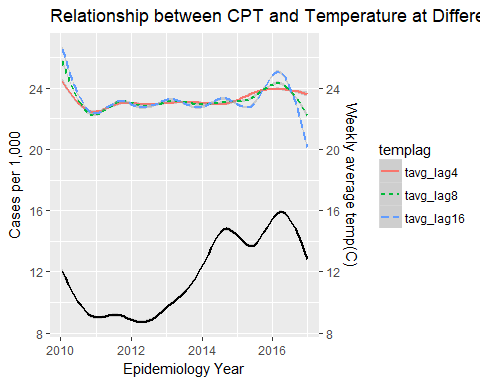
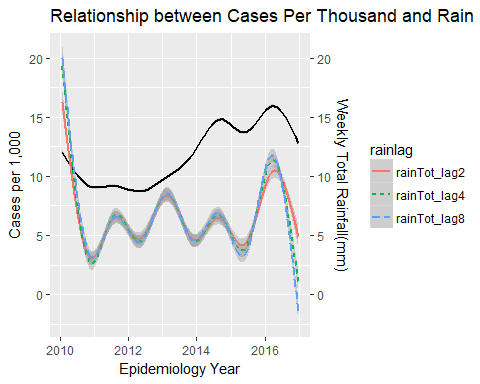
## References

<https://www.cdc.gov/malaria/about/biology/index.html>

## GitHub Location

## https://github.com/picketka/Bios6640\_projects/tree/master/R\_project

## Suplemental Materials/ Appendix A



\*code and plots were developed primarily with Emma Jones and Melissa Lowe