Malaria and ENSO

El Nino Final Project - Sam Mayers

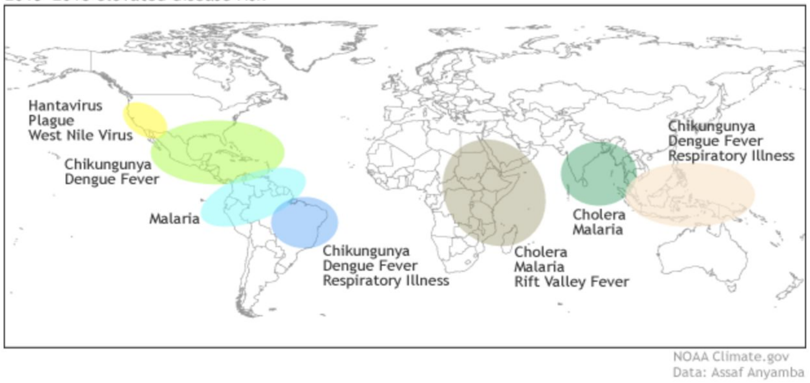
1. Introduction

Malaria infects between 350-500 million people per year, with over 1,000,000 deaths, most of whom are children in Sub-Saharan Africa (1). The disease causes flu-like symptoms, such as fevers, headaches, nausea, among others (2). Those with low or compromised immunity, including pregnant women and travellers, are also at higher risk of infection (2). Malaria is a risk for almost half of the world's population, with 91 countries and territories in Malaria-prevalent regions (2). It is estimated that Malaria has cumulative direct costs of over $12 billion per year, with even larger damages to economic growth in impacted regions (2). Malaria is well understood, and through intervention the numbers of infected have decreased in Malaria impacted regions. However, intervention costs are high in poorer countries, so understanding when Malaria will be most dangerous is essential for resource allocation.

Malaria is transmitted by the Anopheles mosquito that can develop a parasite called Plasmodium falciparum (1), after which this mosquito becomes a possible transmitter of the disease. These mosquitoes live in tropical climates, thriving in places with high rainfall, warm temperatures, and high humidity (1). Humidity and warm temperatures are necessary for these mosquitoes to live long enough for the parasite to develop, which takes at least nine days in warm climates (30o celsius), and longer in cooler temperatures (2). Rain is necessary to keep larval habitats wet, otherwise the young insects die if their habitat dries out (2). Since the adult female Anopheles mosquito rarely lives longer than 1-2 weeks, climate conditions must be ideal for the mosquitoes to survive long enough for the parasite to successfully develop and for Malaria to be a threat to humans in that location. Adult males live on average only a week in the wild, which is not long enough for the Malaria parasite to develop, thus only female Anopheles mosquitoes transmit Malaria. Therefore, since Malaria is dependent on climate factors, both for the survival of the mosquito and the development of the parasite, understanding the patterns of precipitation, temperature, and humidity in Malaria-impacted regions is essential to efficient prevention and control of the disease.

Previous work has studied the connections between many different climate-sensitive diseases and El Nino Southern Oscillation (ENSO), a climate pattern that causes anomalies across the globe. One study led by Dr. Assaf Anyamba in collaboration with a multitude of various governmental organizations, with scientists from NASA, USDA, DoD, WHO, and FAO, among others, found that during El Nino years, diseases were 2.5-28% more prevalent in ENSO impacted regions than during other years (3). ENSO causes changes in typical conditions, sometimes leading to more favorable environments for disease vectors. Since ENSO is predictable, understanding the relationship between ENSO and the change in climate conditions around the world can lead to accurate predictions about when conditions for climate-sensitive disease outbreak may be more favorable and thus, when these diseases are more threatening. In August 2015, Dr. Assaf Anyamba and his team issued early warnings for the possible outbreak of diseases, shown below (3,4).

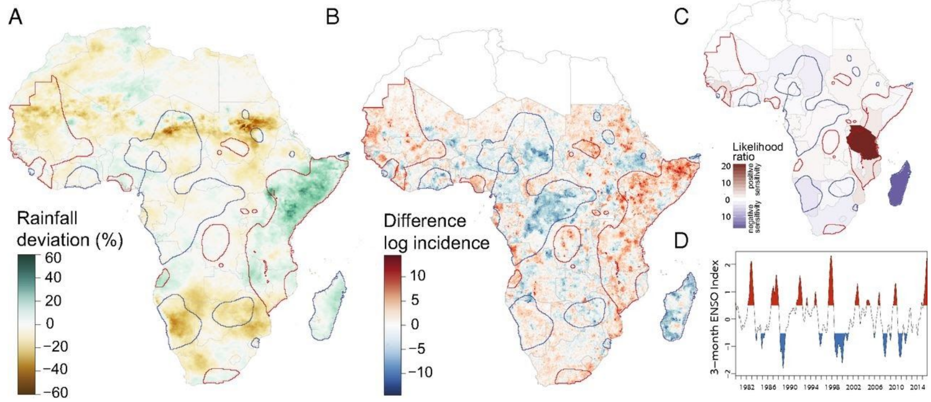
2015–2016 elevated disease risk



Guidance provided to stakeholders by Dr. Anyamba's team on August 28, 2015, indicating the elevated risk for specific disease outbreaks throughout the globe due to the forecast of El Niño conditions through spring of 2016. El Niño conditions did, in fact, strengthen through the fall of 2015 and persist until the spring of 2016. Climate.gov modified figure from A. Anyamba.

The El Nino event of 2015-2016 ended up being particularly strong, and despite the early warnings, unfortunately many of these diseases ended up having higher rates during this time period than normally (4). However, certain prevention methods for Rift Valley Fever were confirmed to have prevented an outbreak. Livestock, usually intensifying the outbreak by infecting mosquitoes that then could infect humans, were vaccinated in response to the warnings in preparation for an outbreak, since no human vaccine for Rift Valley Fever exists. No outbreak occurred, even with the heightened number of infected mosquitoes observed through field studies (4). The success of this prevented outbreak demonstrates how effective intervention can be, and in regions where intervention is particularly expensive, knowing when intervention is especially crucial can save the country's economy and many peoples' lives. The value in predicting disease outbreak is well recognized; the U.S. Global Change Research Program in June 2019 issued a report called "Predicting Climate-Sensitive Infectious Diseases to Protect Public Health and Strengthen National Security" (5).

Malaria specifically has been studied in relation to ENSO. One study found that in Venezuela, Malaria mortality and morbidity following El Nino years were on average 36.5% higher (95% confidence interval, 3.7%-69.3%; P=.004) than following non El Nino years (6). ENSO's impact on Africa has also been studied. One study looked at cholera epidemics in Africa following El Nino events. This study found that during El Nino years, cholera in East Africa increases by 50,000 cases (95% confidence interval), while in South Africa, cholera decreases by 32,000 cases (95% confidence interval) (7). Below is a figure showing multiple maps of Africa, with anomalies during El Nino years of rainfall deviation, difference log incidence, and likelihood ratio shown (7).



From map A above, there is heightened rainfall during El Nino in East Africa, and decreased rainfall during El Nino in South Africa. The study concluded that El Nino affects the distribution of cholera in Africa (7).

For this project, I will be analyzing climate patterns in East Africa where Malaria is endemic, specifically Tanzania. I will look at climate-variables that directly impact Malaria's development and spread, such as precipitation, temperature, and humidity. In addition, recent studies have found that observing only mean monthly temperature overpredicts Malaria in warm temperatures and underpredicts Malaria in cooler temperatures (8,9). Both studies determined that temperature fluctuation impacts the parasite development rate in mosquitoes, leading to relatively faster development in cooler temperatures and relative slower development in warmer temperatures (8,9). Because of these findings, I will also analyze daily temperatures in order to look at temperature variation in connection to ENSO. My goal is to determine if there exists a relationship between these climate variables in Tanzania and ENSO in order to understand the underlying relationship between Malaria in Tanzania and ENSO.

2. Data

The datasets I use provide precipitation, temperature, humidity, and daily temperature data in Tanzania. All datasets are modified to begin in 1980 and ignore any earlier data, since the Nino 3.4 datasets used begin in 1980. The precipitation dataset has values for 491 dates beginning on January 16th, 1980 for a 22 by 28 map of Tanzania. The temperature dataset has the same number of dates beginning on the same date, also for a 22 by 28 map of Tanzania. The humidity dataset provides the same number dates beginning on the same date for a 7 by 8 map of Tanzania, and the daily temperature dataset has values for 14958 dates, beginning on January 1st, 1980, for a 7 by 8 map of Tanzania. I also use a very minimal dataset for Malaria in Tanzania.

The graphs below show the map of the last value (by time) in the datasets. The last figure shows the Malaria numbers (per 1000) in Tanzania from 2000 to 2018. The climate datasets were provided by Professor Tippett, and I acquired the dataset for Malaria in Tanzania from the World Health Organization website (10).

3. Methods

For each climate variable, precipitation, temperature, and humidity, I use the same methods.

First, I graph the climate variable. The first graph of each climate variable's section begins with the average, by month, of that climate variable in Tanzania. Each climate variable's dataset starts in 1980 and goes through 2020. Next, I plot the climate variable on average by month on a line graph, instead of by maps. On this graph, it can be seen how the variable changes through an average year. Lastly, I plot the climate variable on average by year for the 42 years in the dataset on a line graph.

Next, I correlate the climate variable in Tanzania with Nino 3.4 by month. I show the average correlation on a map between Nino 3.4 and the entire climate variable dataset, not grouped by any time period. Then I show the correlation between Nino 3.4 and the climate variable anomaly, with the mean of each month removed for that month (instead of the mean of the whole dataset removed). Then I plot the maps of the correlation between Nino 3.4 and the climate variable for every month in Tanzania. For each of these maps, I find which points are significant, and display another map for each month showing which areas had a significant correlation. I calculate what percent of the points have a significant correlation for every point on every month's map.

To determine if there is a significant difference between the climate variable during El Nino years and La Nina years, I use a T-Test for difference in means. I map the difference in means between the climate variable during El Nino years and La Nino years by month to visualize the potential differences. A p-value is calculated for each month, and a visual of where these values fall by significance is provided. Then, I find the p-values for each coordinate on each map by month to account for possible differences across the country.

I use KNN to predict the climate variable for 2020 by month from the values of Nino 3.4. I show the 10 closest years to the 2020 value of Nino 3.4 (k=10) and the corresponding values of the climate variables in those years to make a conjecture about the accuracy of the following KNN analysis. Then, I plot a 2020 forecast of the climate variable for each month. I show the actual map of the variable for 2020, and then subtract the actual from the forecast and map the difference in order to visualize how close the predictions were to the actual values. The "actual" value for December 2020 is shown as the prediction value, since the actual value is not available yet. Thus, on the maps showing the difference between the forecast and the actual, for December there is no difference, since the difference is unknown. For precipitation and temperature, I repeat this KNN analysis for predicting the 2015 precipitation and temperature values.

I also use linear regression to predict the November 2020 climate variable's map based on Nino 3.4.

For the daily temperature dataset, I test to see if there is a significant difference in the monthly variance of daily temperature for each month between El Nino and La Nina years.

Lastly, I correlate the Malaria in Tanzania data with many different versions (different months and month groupings) of the Nino 3.4 dataset.

4. Analysis

Available in the lab notebook.

5. Results

For all three climate variables, precipitation, temperature, and humidity, there was a significant correlation between the climate variable and Nino 3.4 when broken down by month. This result provides evidence that Nino 3.4 is associated with precipitation, temperature, and pressure in Tanzania.

For precipitation, there was no significant difference between El Nino and La Nina years for the precipitation average across the whole map for the whole year. Because some regions of the map have more precipitation during El Nino and some regions have less during El Nino, this result is expected since the changes cancel each other out. The only month that had a significant difference in means between El Nino years and La Nina years was March, which had significantly less precipitation during El Nino ($p = .0022$). When the p-values were broken down on the map, other months had specific coordinates that showed a significant difference in precipitation.

For temperature, there was a significant difference between El Nino and La Nina years for the precipitation average across the whole map for the whole year ($p=.0011$). The difference between temperature during El Nino and La Nina years showed that for the whole year, every month had on average higher temperatures during El Nino, so the significant difference in temperatures logically follows. When broken down by month, all months except October and November showed significant differences between El Nino years and La Nina years, with significantly higher temperatures observed during El Nino (January ($p = .0241$), February ($p = .0018$), March ($p = 1.0419\text{e-}6$), April ($p = .0013$), May ($p = .0418$), June ($p = .0197$), July ($p = .0025$), August ($p = .0065$), September ($p = .0021$), and December ($p = .0003$)). When the p-values were broken down on the map, other months also had specific coordinates that showed a significant difference in temperature.

For humidity, there was no significant difference between El Nino and La Nina years for the humidity average across the whole map for the whole year. Because some regions of the map have more humidity during El Nino and some regions have less during El Nino, this result is expected since the changes cancel each other out. The only months that had a significant difference in means between El Nino years and La Nina years was April and December, which had significantly more humidity during El Nino ($p = .0121$ and $p = .0118$, respectively). When the p-values were broken down on the map, other months had specific coordinates that showed a significant difference in humidity.

For all three datasets, the KNN predictions from the Nino 3.4 values for the monthly precipitation, temperature, and humidity were all terrible for both 2020 and 2015. This result was surprising given that a strong correlation was observed between the climate variables and Nino 3.4 earlier in the analysis. The linear regression predictions were also terrible for all the datasets.

For the daily temperature dataset, there was no significant difference in monthly variance of temperature. Since the studies that found temperature fluctuation is associated with Malaria development for daily fluctuation (8,9), and this dataset used only had daily temperature (and thus monthly fluctuation), the lack of significant difference does not have much meaning for understanding Malaria in Tanzania.

Lastly, the Malaria in Tanzania dataset had a weak correlation with Nino 3.4 directly. However, the decrease in Malaria (as shown in the Data section), is likely due to factors such as human intervention. With an increase in prevention and control methods, it follows that rates of Malaria decrease. Understanding the underlying factors that make Malaria more threatening will hopefully help reduce these numbers even further in the future.

6. Conclusions

While it is clear from the analysis in this project that precipitation, temperature, and humidity in Tanzania are all associated with Nino 3.4, the predictions in this project yielded poor forecasts and should not be trusted to predict these climate variables in the future. Previous work supports the relationship found between precipitation, temperature, and humidity in Tanzania and Nino 3.4. Significant differences in precipitation, temperature, and humidity in Tanzania were found in this project between El Nino years and La Nina years, confirming the association observed from the project's correlations. However, future work should focus on building more accurate forecasting models.

For temperature, which had the strongest correlation with Nino 3.4 and the most significant p-values for the difference in means, KNN analysis could be improved by removing the increasing temperature trend. Future work could also look at the impact of climate change on Malaria at-risk regions, with some regions becoming warm enough for Plasmodium falciparum to develop in mosquitoes, and some regions becoming too warm and no longer having the risk.

In addition, since I did not have access to a daily temperature fluctuation dataset, future work could study the relationship between daily fluctuation and Nino 3.4, since daily temperature fluctuation impacts Malaria development within mosquitoes. Future work could also confirm the direct relationship between Malaria in Tanzania and Tanzania's precipitation, temperature, and humidity if access to a more thorough dataset of Malaria in Tanzania was available. This data analysis would be important in substantiating the biology behind the development of Malaria in the Anopheles mosquito, especially historically before intervention was as common. The direct relationship between ENSO and Malaria in Tanzania could also be studied much better than in this project if such a dataset could be procured.

Overall, a relationship between Malaria's climate dependent factors in Tanzania and ENSO was found, suggesting that conditions for a Malaria outbreak may be more likely during El Nino.

7. References

1. Climate Change and Vector-Borne Disease. (n.d.). Retrieved December 16, 2020, from https://scied.ucar.edu/longcontent/climate-change-and-vector-borne-disease
2. CDC - Malaria - About Malaria. (2020, March 12). Retrieved December 16, 2020, from https://www.cdc.gov/malaria/about/
3. Anyamba, A., Chretien, J., Britch, S., Soebiyanto, R., Small, J., Jepsen, R., . . . Linthicum, K. (2019, February 13). Global Disease Outbreaks Associated with the 2015–2016 El Niño Event. Retrieved December 16, 2020, from https://www.nature.com/articles/s41598-018-38034-z
4. ENSO and your health: How the 2015-16 El Niño led to early warnings for global disease outbreaks: NOAA Climate.gov. (2019, May 30). Retrieved December 16, 2020, from https://www.climate.gov/news-features/blogs/enso/enso-and-your-health-how-2015-16-el-ni%C3%B1o-led-early-warnings-global-disease
5. Predicting Climate-Sensitive Infectious Diseases to Protect Public Health and Strengthen National Security. (2019, June). Retrieved from https://www.globalchange.gov/sites/globalchange/files/EID%20science%20plan.pdf
6. Dye, C., & Bouma, M. J. (1997). Cycles of malaria associated with El Niño in Venezuela. Retrieved December 16, 2020, from https://europepmc.org/article/med/9388155
7. Moore, S., Azman, A., Zaitchik, B., Mintz, E., Brunkard, J., Legros, D., . . . Lessler, J. (2017, April 25). El Niño and the shifting geography of cholera in Africa. Retrieved December 16, 2020, from https://www.pnas.org/content/114/17/4436
8. Blanford, J., Blanford, S., Crane, R., Mann, M., Paaijmans, K., Schreiber, K., & Thomas, M. (2013). Implications of temperature variation for malaria parasite development across Africa. Retrieved December 16, 2020, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3575117/
9. Paaijmans, K., Read, A., & Thomas, M. (2009, August 18). Understanding the link between malaria risk and climate. Retrieved December 16, 2020, from https://www.pnas.org/content/106/33/13844
10. hUnited Republic of Tanzania. (n.d.). Retrieved December 16, 2020, from https://www.who.int/data/gho/data/countries/country-details/GHO/united-republic-of-tanzania?countryProfileId=40f06adc-047c-435c-b576-ebd4da789bb2