USING REMOTE SENSING AND MODELING TECHNIQUES TO INVESTIGATE MALARIA PREVALENCE IN LORETO, PERU

by

Aneela Mousam
A Thesis
Submitted to the
Graduate Faculty
Of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Master of Science
Civil and Infrastructure Engineering

Committee:

_____ Dr. Viviana Maggioni, Thesis Director

_____ Dr. Paul L. Delamater, Committee Member

_____ Dr. Mark H. Houck, Committee Member

_____ Dr. Liza W. Durant, Department Chair

_____ Dr. Kenneth S. Ball, Dean, Volgenau School of Engineering

Date: _____ Spring Semester 2016
George Mason University
Fairfax, VA

Using Remote Sensing and Modeling Techniques to Investigate Malaria Prevalence in Loreto, Peru

A Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at George Mason University

by

Aneela Mousam
Bachelor of Science
Virginia Tech, 2013

Director: Viviana Maggioni, Assistant Professor
Department of Civil and Infrastructure Engineering

Spring 2016
George Mason University
Fairfax, VA

DEDICATION

This is dedicated to my parents, Shahnaz Begum and Mousam Khan.

ACKNOWLEDGEMENTS

I would like to thank the many friends, family, and professors who have made this happen. I would also like to thank the Office of the Provost at George Mason University for funding a summer research fellowship, which gave me the opportunity to travel to Peru. I would also like to thank the Loreto Ministry of Health for providing the malaria dataset, NASA Precipitation Processing System (PPS) for the TRMM 3B42 data and the Global Modeling and Assimilation Office (GMAO) and the GES DISC for the dissemination of MERRA.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

ABSTRACT

USING REMOTE SENSING AND MODELING TECHNIQUES TO INVESTIGATE MALARIA PREVALENCE IN LORETO, PERU

Aneela Mousam, M.S.

George Mason University, 2016

Thesis Director: Dr. Viviana Maggioni

Peru is a country still working toward completely eliminating malaria. Between 2001 and 2010 significant progress was made towards reducing the number of malaria cases, but the country saw an increase between 2011 and 2012. This work attempts to uncover the associations among various climate and environmental variables and malaria prevalence in the Peruvian region of Loreto, which is located in the Amazon basin. A Multilevel Mixed-effects Poisson Regression model is employed to investigate the relationship between malaria prevalence and climate and environmental conditions during 2009-2013. The results indicate that increase in elevation (β=0.78; 95% confidence interval (CI) 0.75-0.81), soil moisture (β=0.0021; 95% CI 0.0019-0.0022), rainfall (β=0.59; 95% CI 0.56-0.61) and normalized difference vegetation index (β=2.13; 95% CI 1.83-2.43) are associated with higher malaria prevalence, while increase in temperature (β=-0.0043; 95% CI -0.0044, -0.0041) is associated with a lower malaria prevalence. The results from

this study are especially useful for healthcare workers in Loreto and have the potential of being integrated within malaria elimination plans.

INTRODUCTION

Malaria continues to be one of the most severe public health problems worldwide. According to the World Health Organization (WHO, 2015), 1.2 billion people are at a high risk of being infected with malaria and developing the disease and 214 million cases were reported in 2015. Malaria can be especially fatal in children under the age of 5, who account for 69% of all deaths. Globally, malaria transmission is ongoing in 96 countries and territories (WHO, 2015). In many of these places, malaria transmission does not necessary occur in all parts of the country because transmission can be impacted by various regional factors. In general, warmer regions close to the equator see a more intense transmission that occurs year-round (CDC, 2012). Figure 1 presents the triangle of Human Ecology for malaria (Mead and Emch, 2010). The triangle gives an overview of different risk factors associated with malaria and speakers to complexity of understanding this disease.

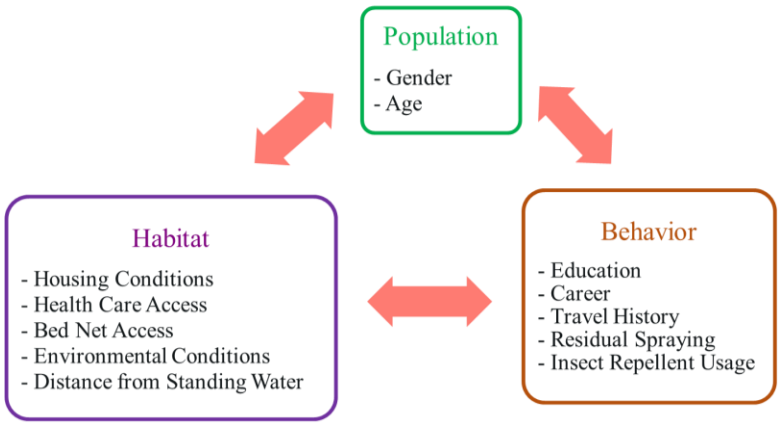

Figure 1. Triangle of Human Ecology for Malaria

Malaria is caused by a parasite that is transmitted to people through the bite of an infected female Anopheles mosquito. The four main parasite species responsible for causing the disease in humans include: *Plasmodium falciparum, Plasmodium vivax, Plasmodium malariae and Plasmodium ovale* (WHO, 2015). In humans, the parasite grows and multiplies in the liver cell producing merozoites. The merozoites exit the liver cell and re-enter the bloodstream and invade the red blood cells. In the blood cell, the parasite grows and destroys the cell and releases daughter parasites. These parasites then invade other red blood cells continuing the cycle. The parasites are then picked up by a mosquito, when they feed on an infected human and the mosquito can continue spreading the infection, when they bite another human. Therefore, the mosquito acts as the vector through which the parasite is spread from human to human (CDC, 2012). The symptoms of the disease usually start appearing 10-15 days after the bite. The most common symptoms include fever, headaches, chills and vomiting. If the symptoms are not treated within 24 hours the disease may lead to a more severe condition and possibly cause death (WHO, 2015).

WHO states that malaria has an especially heavy burden on the poorest and most marginalized communities, primarily affecting low and lower-middle income countries. The high risk associated with malaria is primarily due to lack of effective services for prevention, diagnosis and treatment. Therefore, it is critical to control and eliminate malaria in order to improve public health and reduce poverty (WHO, 2015). Due to the impact of this disease, many countries have taken initiatives in an attempt to eliminate it completely. In 2010, 79 countries were successful in eliminating malaria. But there are still approximately 99 endemic countries from which, 67 are controlling malaria and 32 are pursuing an elimination strategy (Feachem et al., 2010). The main approaches to elimination are killing the parasite with appropriate medication such as Primaquine and 8-aminoquinoline tafenoquine, vector control technologies like insecticide treated nets and indoor residual spraying, reduction in infection importation, and cross regional initiatives (Moonen et al., 2010; Ferguson et al., 2010).

Controlling and eliminating malaria can bring forth many benefits for a country. Moving from a high malaria burden to low can help in increasing productivity through increased human capital and increased productivity of factor such as land or capital (Modrek et al., 2012). Elimination can generate a boost in economic activities of a country by creating an environment that is open to foreign investment and tourism. Malaria elimination has also the potential to strengthen the components for health system, which in turn would improve public health (Sabot et al., 2010).

Although the benefits of controlling malaria are wide range, there are many challenges being faced by countries that are trying to pursue an elimination strategy

(Feachem et al., 2010). One ongoing threat to elimination is the resistance to chemical agents, such as DDT, which are used to target adult and larval stages of vectors (Moonen et al., 2010). There is also a presence of drug resistant parasites that are emerging such as artemisinin resistance in the Asia Pacific (Cotter et al., 2013). Even if a country has been successful in eliminating malaria, the occurrences of imported malaria cases can be a very critical threat to the achievement and maintenance of elimination. This threat is greatest for countries that are neighboring high endemic areas (Cotter et al., 2013; Moonen et al., 2010).

The majority of the cases occur in Africa and South-East Asia, but transmission continues in several parts of South America as well (WHO, 2015), as shown in Figure 2. In 2012, approximately 25% of the malaria burden in South America was experienced by only 12 municipalities in Peru, Brazil, and Venezuela (Zaitchik et al., 2012). Figure 3 below presents the most recent malaria distribution in South America.

Peru is progressing towards controlling malaria but has not been able to completely eliminate the disease, thus making it the country with the second highest number of malaria cases in South America (WHO, 2015; Bautista et al., 2006). In 2015, Peru has an estimated population of 30,973,148, of which 12,165,089 have at least some risk of contracting malaria (WHO, 2015).

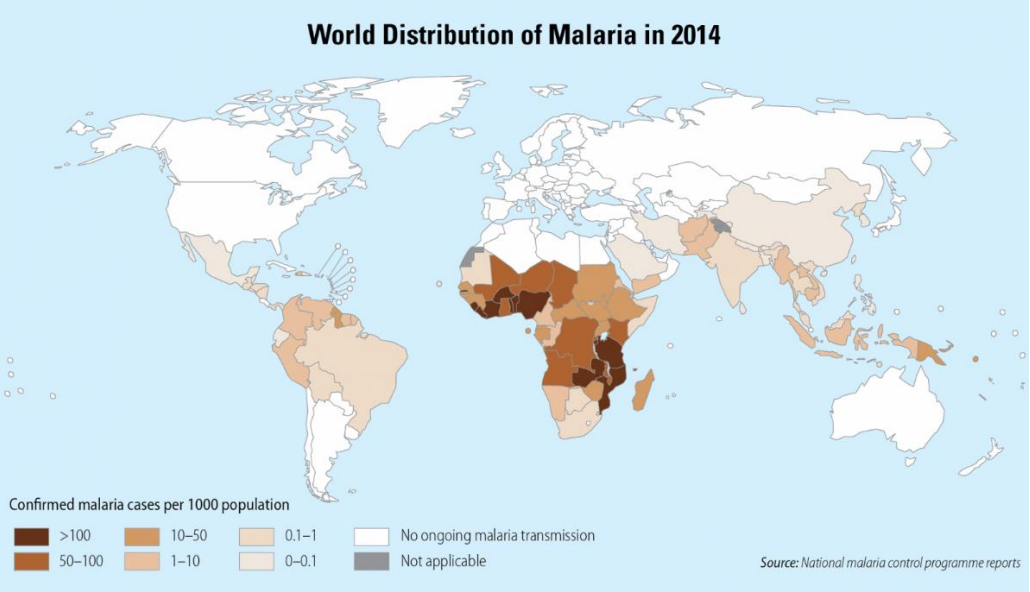

Figure 2. World Distribution of Malaria in 2014

During the 1990s, there has been a 7-fold increase in malaria incidence in Peru, rising from 13 per 10,000 inhabitants in 1990 to a peak of 88 per 10,000 in 1996 (Roper et al., 2000). Specifically, over 60% of all malaria cases occurred in the Loreto Department of Peru (Zaitchik et al., 2012). Therefore, the Department of Loreto has been the major focus of the malaria control. In 1990, there were only 641 cases in Loreto, but reached 121,268 cases by 1997 (Roper et al., 2000). Peru saw an overall decline in malaria cases from 2001-2010. However the number of cases increased from 2011-2013, especially in Loreto.

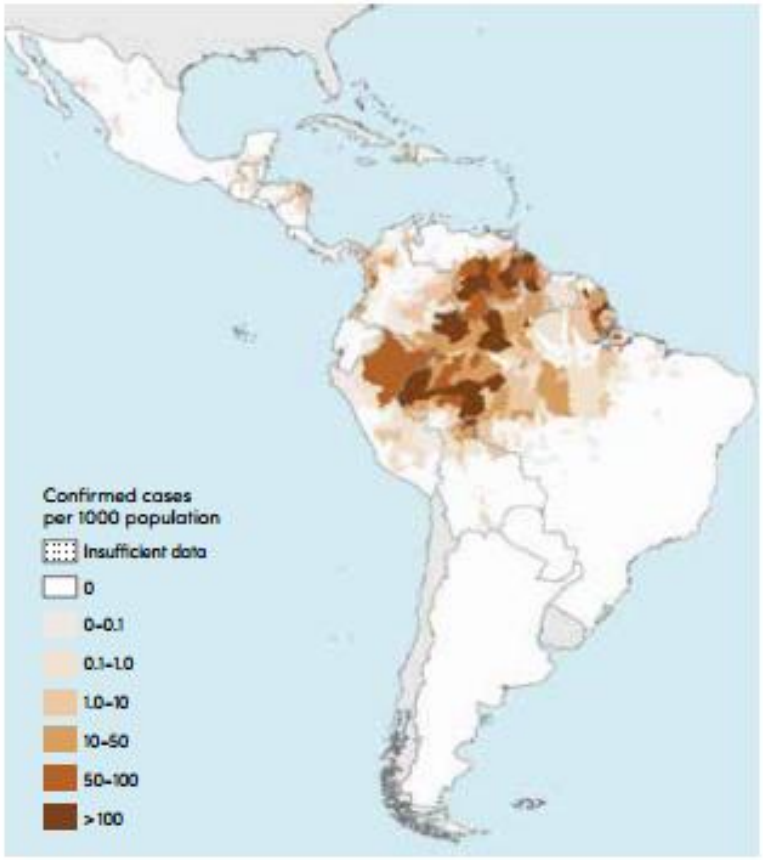

Figure 3. Malaria Distribution in South America in 2014

Peru has implemented various initiatives in an effort to control malaria. The Peruvian Malaria Program provides free antimalarial drugs under a Directly Observed Therapy (DOT) protocol (Chuquiyauri et al., 2012). In addition, regional efforts to improve malaria surveillance, early detection, prompt treatment, and vector management have been employed since 2000 (Herrera et al., 2012). Despite these efforts and increased funding for malaria control in the region, there are still gaps in understanding how different factors impact malaria transmission and elimination (Herrera et al., 2012). This brings into question the role of climate and environmental factors.

Past studies clearly indicate that global climate variability already has and will continue to have an impact on malaria transmission, which will ultimately impact countries that are pursuing a malaria elimination strategy. Specifically, Patz et al. (2005) and Parham and Michael (2010) state that climatic variations and extreme weather events have a profound impact on infectious agents and their associated vector organisms. These two studies also show that vectors such as mosquitoes are devoid of thermostatic mechanisms, so their reproduction and survival rates are strongly impacted by fluctuations in temperature. Parham and Michaels (2010) show that environmental variables such as temperature, humidity, rainfall, and wind speed can affect the incidence of malaria by impacting the changes in the duration of the parasite's life cycle and parasite behavior.

Githeko and Ndegwa (2001) focus on the East African Highlands and argue that the underlying cause of the malaria epidemic is due to the changing climatic conditions in this normally cool area. An increase in temperature has also been shown to accelerate the rate of mosquito larval development and the frequency of bites on humans, as well as impact the time it takes for the malaria parasite to mature into the mosquito stage. Increases in rainfall can create additional habitats for mosquitoes to breed, thus increasing vector populations. Githeko and Ndegwa, (2001) conclude that in the past decade there has been an increase in the anomalies of mean monthly temperatures, which has a strong relationship with the number of malaria cases.

Few past studies have attempted to identify the relationship between climate variables and malaria risk in Peru. Jones et al. (2004) proposes that environmental factors are responsible for changes in the mosquito population over time. This study focused on

Loreto, where a higher overall mosquito population is observed from October 1996 through March 1997, which corresponds to the rainy season (Jones et al., 2004). Aramburú Guarda et al. (2004) finds a positive correlation between malaria transmission periods and rainfall and higher temperatures near the Amazon River. Additionally, Aramburú Guarda et al. (1999) show that the two precipitation peaks in 1997 occurred three months and one month before the malaria cases reached their highest levels in Loreto. These studies take place in the Loreto region but fail to take into account a long time series of the climate and environmental data, which is critical to observe temporal trends.

This work goes one step further, by investigating how remote sensing and modeling products can be used to analyze trends in malaria prevalence by expanding on these past studies to include a longer time series, a larger study area (i.e., the whole Loreto Department), and by looking at a more complete set of environmental variables. Field observations in the region are limited, as the Loreto department comprises nearly one-fourth of the land mass of Peru and has a low population density (Aramburú Guarda et al., 1999), making it difficult to conduct field data collections of climate and environmental factors. Thus, remote sensing and modeling data are extremely valuable to obtain the necessary information of the current environmental and climate conditions of the region and to investigate the impacts of the factors on the malaria transmission.

The yearly malaria prevalence rates from 315 health centers located in Loreto and environmental and climate variables such as temperature, humidity, soil moisture, vegetation index, and elevation are analyzed. All of the variables are entered into a Multivariate Poisson Regression Model to study the dependence of malaria prevalence on

environmental conditions and identify which regions of the department are suitable for malaria transmission based on the noted factors. Results from this study can be applied for surveillance purposes to target regions that are at a higher risk for elimination strategies.

STUDY AREA AND DATASETS

Study Area

Loreto is one of the 25 departments in Peru, located in the Northeast region of the country. Loreto comprises one fourth of Peru's land area and has a total area of approximately 348,177 km^2 (Griffing et al., 2013; Vittor et al., 2006). The region lies in the amazon rainforest basin and has ecological characteristics of the amazon lowlands (Aramburú Guarda et al., 1999). Loreto has high and low jungle and almost all of the area is covered by thick vegetation (National Geographic, 2015; CIA, 2015). The region has two distinct seasons: wet and dry. The rainy season is between November and May, although precipitation occurs year around. Loreto's annual average temperature is 28°C and the region has a persistent, high relative humidity of more than 87% year around (Aramburú Guarda et al., 1999).

The Loreto region only has one major paved road with small unpaved roads connecting villages. However, in practice, most of the movement in this region happens along the river networks (Kvist & Nebel, 2001; Abizaid, 2005; GOREL, 2006). As a result, majority of the population resides in close approximation of the river (Figure 4).

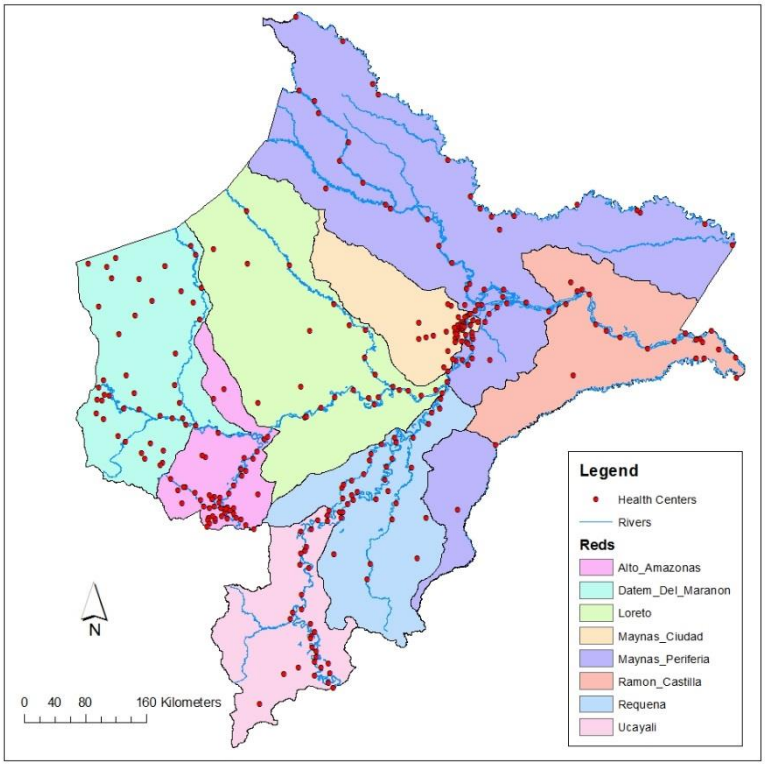

Figure 4. Health Center locations in the red and micro red network

There are 315 health centers in Loreto. The health centers are nested in 8 "red" networks, which are based on the boundaries of the provinces in the region. Figure 4, displays a map of the health center locations, the rivers, and the boundaries of the reds. The figure shows the distribution of the health centers within each of the networks. Some reds have several health center clustered in the same area, while in other regions the health centers are more dispersed.

Malaria Data

The malaria data used for this study are obtained from the Loreto Ministry of Health for the 2009 - 2013 time period. The dataset includes yearly case counts for both *Plasmodium Falciparum* and *Plasmodium Vivax* cases at each of the 315 health centers. The two species

differ morphologically, immunologically, in their geographical distribution, and in their relapse and drug response (Tuteja, 2007).

Data on the estimated population at each health center are used to determine the prevalence rate, as defined in equation 1 (SJSU, 2012). The population estimates per health center are obtained from the Loreto's Ministry of Health for the year 2009. The population estimates for 2010-2013 are calculated based on linear interpolation from the 2009 population assigned to each health center and the total population estimates of the department at each year.

Equation 1

$$\text{Prevalence} = \left(\left(\frac{number\ of\ reported\ cases\ during\ a\ time\ period}{population\ during\ time\ period} \right) * 1000 \right)$$

The total weekly malaria case count for the whole department is extracted for 2009 - 2013 from the Loreto Ministry of Health's weekly health bulletins. The results of the weekly malaria counts are presented in Figure 5.

Climate and Environmental Variables

Several satellite and modeling products are used to study climate and environmental variables at each health center for 2009 – 2013, including outputs from the NASA MERRA (Modern-Era Retrospective analysis for Research And Applications) model, precipitation data from the Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA), vegetation products from the moderate-resolution imaging spectroradiometer **(**MODIS) instrument, and elevation data from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Elevation Model (GDEM).

NASA MERRA Model

The MERRA model provides historical time series of the hydrological cycle variables, such as precipitation, temperature, humidity, and surface pressure from 1979 to present (Reichle et al., 2011; Reinecker et al., 2011). MERRA incorporates information from remote sensing observations of the atmosphere from many modern satellites and provides estimates of surface meteorological data such as precipitation, radiation, air temperature, and humidity as well as land surface variables such as soil moisture and runoff. Data are available at hourly steps and at 1/2° × 2/3° spatial resolution in latitude and longitude. The climate variables obtained from MERRA include specific humidity at 2 m above the displacement height (QV2M), temperature at 2 m above the displacement height (T2M), and soil moisture content in the top soil layer (SFMC).

Precipitation Data

Precipitation data is obtained from the TRMM TMPA dataset (NASA, 2015). TMPA provides precipitation estimates by merging information from multiple satellite sensors and ground-based gauges (Huffman et al., 2007; Huffman et al., 2010). Specifically, this product combines the rainfall estimates of passive of several passive microwave sensors (PMW) that are onboard Low Earth Orbit Satellites and sensors that are on board platforms of the Defense Meteorological satellite products (DMSP) and NOAA (Mantas et al., 2015). TMPA estimates are produced by using the PMW rain rates for each sensor through the Goddard Profiling algorithm (Kummerow et al., 2011; Gopalan et al., 2010). TRMM data is available from latitude band 50°N–S from 1998-2014 for 3 hourly steps and 0.25° x 0.25° resolution (Huffman et al., 2007). The TMPA rainfall products are available in two

versions: a real-time version (TMPA 3B42RT) and gauge-adjusted post real-time research version (TMPA 3B42). The main difference between the two products is the use of rain gauge data for bias adjustment and the date of release (Melesse, 2011). For this study, the TMPA 3B42V7 product is used, since it is reported to have better performance in terms of bias with respect to the real-time version (Habib et al., 2009; Maggioni et al., 2016).

Vegetation Data

Satellite based vegetation indices are obtained from the global 16-day composite of MODIS vegetation indices that provide spatial and temporal comparisons of vegetation conditions. The indices include the MODIS normalized difference vegetation index (NDVI) and the Enhanced Vegetation Index (EVI), which is more responsive to canopy structural variations. For this study, data is obtained from the MOD13C1 Product at a 0.05° resolution (Herring, 2000). NDVI is defined as follows

Equation 2

$$\mathrm{NDVI} = \frac{\mathrm{N}-\mathrm{R}}{\mathrm{N}+\mathrm{R}}$$

where N and R are the reflectance in the near-infrared (NIR) and red bands, respectively. EVI is defined as

Equation 3

$$\mathrm{EVI} = \mathrm{G}\left(\frac{\mathrm{N}-\mathrm{R}}{\mathrm{N}+\mathrm{C_1R}-\mathrm{C_2B}+\mathrm{L}}\right)$$

where N, R, and B are atmosphere-corrected surface reflectance in near-infrared, red, and blue bands. G is a gain factor, C_1, and C_2 are coefficients of the aerosol resistance term. The coefficients used in the MODIS EVI algorithm are, L=1, C_1=6, C_2=7.5, and G=2.5 (Jiang et al., 2008).

Elevation Data

The ASTER GDEM model provides elevation data globally at a 30 meter resolution. The ASTER instrument is launched onboard NASA's Terra Spacecraft and has the capability of using near infrared spectral band and nadir-viewing and backward-viewing telescopes to acquire stereo image data (Japan Space Systems, 2011; NASA, 2016). To produce the ASTER DEMs, the ASTER archive are processed through an automated method which includes, cloud masking, stacking all cloud-screened DEMs, removing bad values, and averaging selected data to create final pixel values. The ASTER GDEM is available for land surface regions between 83^{o} N-S in $1° \times 1°$ tiles (Japan Space Systems, 2011).

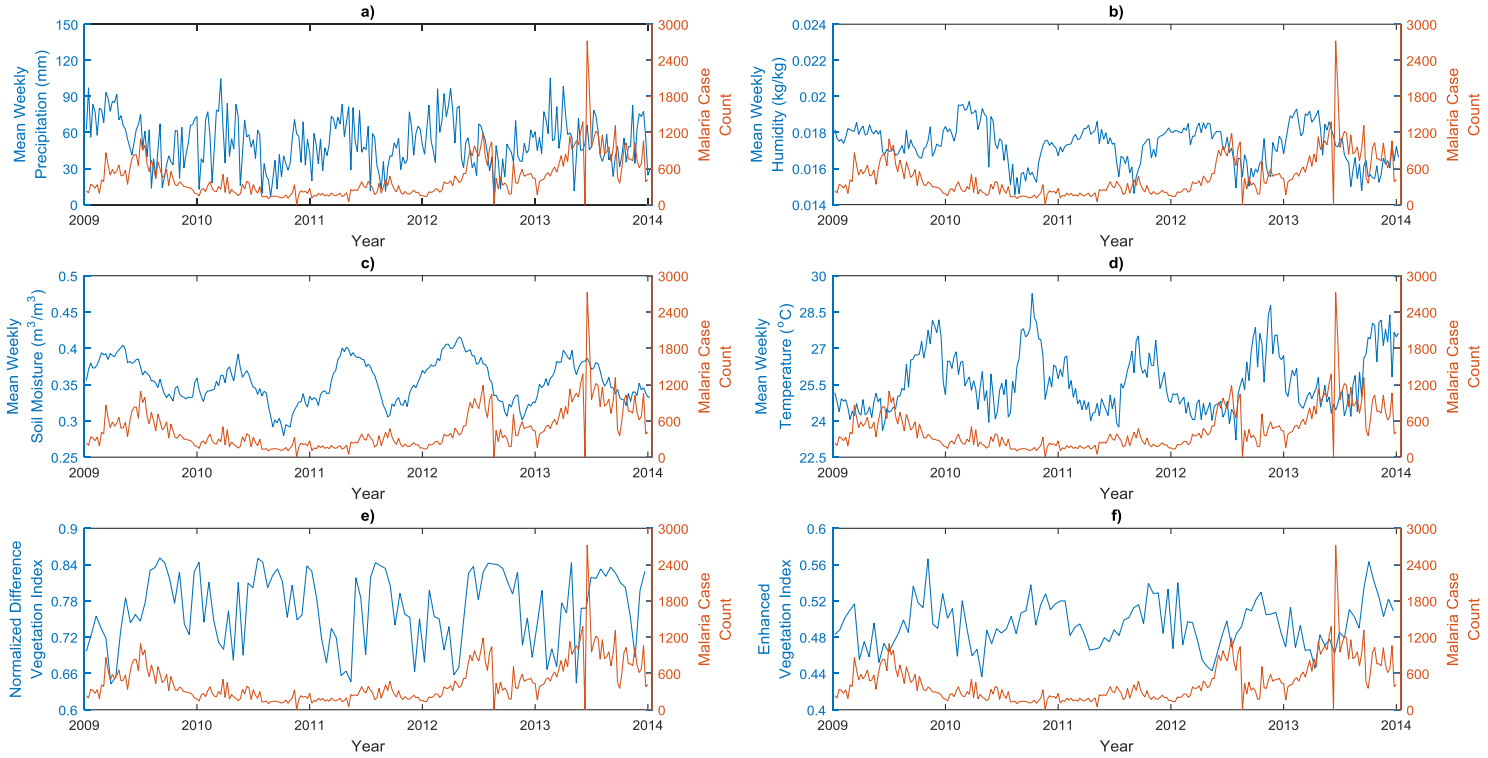

Figure 5. Weekly time series plots for weekly malaria cases and (a) mean weekly precipitation, (b) mean weekly humidity, (c) mean weekly soil moisture, (d) mean weekly temperature, (e) mean weekly normalized difference vegetation index (NDVI), and (f) mean weekly enhanced vegetation index (EVI) for the whole Loreto Department

Time Series

The time series of the weekly average precipitation, temperature, humidity, NDVI, EVI and number of malaria cases from 2009 to 2013 for the entire area are shown in Figure 4. The time series for the climate and environmental factors indicate that there is a strong seasonality. In Figure 5, plot a, plot b, and plot d, the values begin increasing in the last few months of the year and the peak occurs in the beginning of the year, which lines up with wet season (November-May) in Loreto. For temperature (plot d), NDVI (plot e), and EVI (plot f), the peak occurs in the second half of each year between May to December. The peak malaria cases occur approximately in the middle of each year. The maximum weekly precipitation, temperature, and humidity average for the entire time period occur between 2010 and 2011, which align with the time when malaria cases begin to increase.

METHODOLOGY

Mixed Effects Poisson Regression Mode

For this study, a mixed-effects Poisson regression model is used to study the relationship between malaria prevalence and climate and environmental variables. This type of model is a Poisson regression that contains both fixed effects and random effects. There are many advantages of using a mixed-effect model for this type of analysis. First, mixed-effect models can be applied to continuous and non-normally distributed outcomes (e.g., Poisson distribution). Second, this family of models is robust when handling missing data, time-invariant and time varying covariates (Gibbons et al., 2010). Mixed-effects models also allow for modeling the correlation that might exist in grouped data and, therefore, the nesting of the groups can be treated as random effects within the model (Buckley et al., 2003; Ren et al., 2015).

A Poisson regression model is generally utilized when the response variable is a discrete number (n = 0, 1, 2, …,N). In comparison to ordinary regression models, this technique has the constraint that: 1) predicted values are non-negative numbers, and 2) the mean and the variance of the errors are equal to each other. Moreover, the Poisson regression model assumes that the probability distribution of the response follows a Poisson distribution (Gardner et al., 1995; Long, 1997):

Equation 4

$$\Pr(y_{ij} = y|x_{ij}, u_j) = \frac{(\exp(-\mu_{ij}))\,\mu_{ij}^{y}}{y!}$$

where $\mu_{ij} = \exp(x_{ij}\beta + u_j)$, $j = 1, 2, \ldots, m$ clusters (reds), with cluster j consisting of $i = 1, 2, \ldots, n_j$ (health centers). The responses are counts y_{ij} (malaria prevalence). The row vector x_{ij} corresponds to the covariates for the fixed effects, with the regression coefficients (fixed effects) β. The random effects are represented by u_j (STATA, 2013). In this model, the fixed effects represent the climate and environmental variables and the random effect accounts for the correlation that might exist in the red networks.

Poisson regression assumes that the logarithm of its expected value can be modeled by a linear combination of unknown parameters (Ahmed, 2014). The equation written as a generalized linear function is:

Equation 5

$$\log(\mu_{ij}) = \beta_0 + \beta 1 x_{ij} \ldots + \beta_n x_{ij} + u_j$$

The model outputs coefficient values for the constant (β_0), the fixed effects (β_n) and the random effects (u_j). In Poisson regression, the response is conceptualized as a rate. Therefore, a positive coefficient indicates higher rate and a negative coefficient indicates a lower rate (Schofer, 2007).

The coefficients can be exponentiated to determine the incidence rate ratio (IRR), which is computed as:

Equation 6

$$IRR = exp(\beta_n)$$

The IRR can be converted to a percentage by using the following formula:

Equation 7

$$IRR(\%) = IRR - 1 * 100$$

The percentage can help to indicate the magnitude of the change of the fixed effects on the response (Schofer, 2007).

Data Analysis

In this study, the mixed-effects model is implemented using a Multilevel Mixed-effects Poisson Regression in STATA 13. The model includes malaria prevalence at each health center. The centers are clustered by red networks to take into consideration the shared, clustered-level random effects. The climate and environmental data are gridded, therefore it is necessary to determine in which grid each of the health center is located. To determine the conditions at each health center data, the nearest neighbor value of the climate and environmental data is used.

All the climate and environmental variables are transformed in an attempt to better characterize the conditions throughout the year. For example, rather than simply considering the average yearly temperature, the number of days with temperature above a certain degree is considered. Due to the fact that malaria data is available on a yearly scale, numerous transformations for the variables in an attempt to test a broad range of potential representations of the variables. The full list of variable transformations is presented in Appendix D.

Each of the climate and environmental variable transformation is examined in a univariate analysis to identify the transformation that will be included in the multivariate mode. Thus, the Akaike Information Criterion (AIC), which provides a measure of the relative quality of a model for a set of data, is computed. The AIC is defined as:

Equation 8

$$\mathrm{AIC} = -2\,\mathrm{lnL} + 2\mathrm{k}$$

where lnL is the maximized log-likelihood of the model, k is the number of parameters estimated (STATA, 2013).

AIC criterion is used to estimate the quality of each model relative to other models. The AIC values give insight into the fitness of the model and help to select the best variable transformation. Given two models, the one with the smaller AIC indicates a better-fitting model (STATA, 2013).

The AIC for each of the various transformations of all variables is calculated in order to test what is the best transformation for each variable to use for the multivariate model. The variable transformations with the lowest AIC value are then adopted in the multivariate analysis. The forward selection is used to determine the variables for the multivariate model. In this approach, each variable is added to the model one at a time starting with the one that is the most significant (i.e., lowest AIC) and assessing what effect adding a variable has on the AIC value. The variable that results in the lowest AIC for the model is included.

Pearson's correlation coefficients are also calculated for each of the variable to avoid multicollinearity. Multicollinearity can inflate the variance of one of the estimated regression coefficients and produce untrustworthy model results. A coefficient value greater than 0.5 or less than -0.5 corresponds to a high degree of correlation between the two variables, so one of the two variables would be eliminated. The multivariate analysis is applied first to the total malaria prevalence and then to the two types of malaria (*P. Falciparum* and *P. Vivax*) separately.

RESULTS

Sample Characteristics

The total number of malaria cases for Loreto decreased from 2009 to 2010, then began increasing afterward, reaching 43,737 cases in 2013. The yearly number of malaria cases, population, and malaria prevalence for the entire department are presented in Table 1 for 2009 to 2013.

Table 1. Year malaria count and population for the whole department

	2009	2010	2011	2012	2013
Malaria Count	23,486	11,445	11,779	25,148	43,737
Population	766,169	766,578	775,321	784,113	792,935
Malaria Prevalence	30.7	14.9	15.2	32.1	55.2

Table 2 shows the mean yearly malaria prevalence for the health centers in each of the eight red networks, as well as the number of health centers in each. In 2013, Maynas Ciudad, Datem Del Maranon and Maynas Periferia showed the highest malaria prevalence, whereas Ucayali recorded the lowest average prevalence among all the reds in 2013. As shown in Figure 4, Maynad Ciudad and Maynas Periferia are located closer to the equator compared to the other reds. Datem Del Maranon is located in the western region of the department near the border with Ecuador. Ucayali is located at the southernmost part of the department and the furthest away from the Equator compared to the other reds. The distance from the equator is critical as areas closer to it experience hot climate with little

seasonal variations (National Geographic Society, 2011). Therefore, the environmental conditions are more favorable for malaria transmission.

Another point to note in Table 2 is that the standard deviation of the prevalence is higher than the mean, which indicates high variability in malaria prevalence among the health centers in the each of the red networks.

Table 2. Mean and standard deviation of yearly prevalence per red network

	No. of Health Centers	2009		2010		2011		2012		2013	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Alto Amazonas	52	26.5	41.8	18.3	28.4	18.7	29.6	18.8	45.2	43.0	99.06
Datem	45	59.2	107.6	32.9	63.0	33.5	64.7	65.1	143.7	143.3	319.8
Del Marañon	27	52.4	78.1	21.3	28.6	21.6	29.2	43.6	85.4	80.4	140.6
Loreto	45	64.3	141.4	36.5	81.1	36.5	82.7	171.6	525.5	201.3	407.1
Maynas Ciudad	57	115.9	254.9	51.7	95.2	52.3	96.36	65.9	138.6	131.2	299.2
Maynas periferia	21	58.8	52.2	17.9	16.1	18.0	16.4	60.1	89.1	120.2	208.23
Ramon castilla	34	23.1	98.5	12.9	53.1	12.9	53.4	28.2	121.3	42.9	152.1
Requena	34	0.0	0	0.03	.17	0.03	.171	0.09	.287	0.8	4.454
Ucayali											

Univariate Analysis

The AIC is used to determine which variable transformation to use in the multivariate analysis (Appendix A). First, the model is created for the total malaria prevalence including both the *P. Vivax* and *P. Falciparum*. According to the univariate analysis, the variable transformations with the lowest AIC value, which indicate a better model fit, are the P2 (cumulative precipitation during the wet season (in m)), E2 (the elevation above 100 meters), H4 (the number of days with humidity above 0.018 $kg_{vapor} \times kg_{air}^{-1}$), SM4 (the number of days with soil moisture above $0.400 m^3/m^3$), T9(the number of days with temperature above 25°C), and NDVI (the yearly average normalized difference vegetation index).

Next, a univariate analysis is conducted for the *P. Falciparum* and the *P. Vivax* prevalence independently to investigate if the two types of malaria are associated with different climate and environmental conditions. For *P. Falciparum* (Appendix B), the transformed variables with the lowest AIC value are the same as the one for the total malaria prevalence with exception of precipitation, where P7 (the number of days with rainfall over 15 mm) had the lowest AIC number rather verses the cumulative precipitation during the wet season. For the *P. Vivax* (Appendix C), the variables transformations with the lowest AIC values are the same as the ones for the total malaria prevalence.

Multivariate Analysis

The forward selection is used to determine the variables for the multivariate model. Additionally, the correlation coefficient is calculated for each of the variables to avoid multicollinearity, similarly to the univariate analysis. Table 3 indicates that the correlation

between the variables for each multivariate model is low (i.e., between -0.075 and 0.44). Five climate and environmental factors are included in the multivariate mixed effects model based on the univariate model for the total malaria prevalence (Table 4).

Table 3. Correlation matrix of variables used in the multivariate analysis for the total and P. Vivax prevalence model

	T9	E2	SM4	P2	NDVI
T9	1				
E2	-0.075	1			
SM4	-0.26	-0.38	1		
P2	-0.13	0.0034	0.12	1	
NDVI	0.044	0.44	-0.32	-0.024	1

Based on the results presented in Table 4, the incidence rate ratio (IRR) can be calculated for the fixed-effects variables. The IRR for T9 is 0.996 which corresponds to the 0.40% reduction in the malaria prevalence when the number of days in the year with temperature above 25°C increases by one day. Aramburú Guarda et al. (1999) also show a negative correlation with temperature and malaria cases though the study area is only limited to the Iquitos region. Due to the fact that mean temperature in the rainy season is still within optimal range for mosquito development, the temperature correlation might not have as much impact (Aramburú Guarda et al., 1999).

The IRR value for E2 corresponds to a 118% increase in malaria with an increase in the number of health center having an elevation above 100m. This corroborates what presented by previous studies that showed that mosquito densities and parasite prevalence decrease with increasing altitude (Drakeley et al., 2005; Attenborough et al., 1997; Bødke et al., 2003; Akhwale et al., 2004).

Table 4. Multivariate analysis for total malaria prevalence

Coefficient	Estimate	95% CI	IRR – Fixed Effects
β_0	-472.34	-482.66, -462.03	
β_{T9}	-0.0043	-0.0044, -0.0041	0.996
β_{E2}	0.78	0.75, 0.81	2.18
β_{SM4}	0.0021	0.0019, 0.0022	1.002
B_{P2}	0.59	0.56, 0.61	1.797
β_{NDVI}	2.13	1.83, 2.43	8.448
β_{year}	0.24	0.23, 0.24	1.265
Random effects	3.29	1.23, 8.83	

In terms of the SM4 IRR, a 0.2% increase in malaria prevalence is observed when there is an increase in the number of days in the year with soil moisture above 0.400 m^3/m^3 by one day. This is in line with a previous study conducted over Kenya, which concluded that soil moisture better predicts the biting rates compared to rainfall (Patz, et al., 1998).

Finally, the increase in P2 and the NDVI are also associated with an increased in malaria prevalence. Since NDVI, and vegetation in general, are linked to temperature and precipitation, this result confirms that vegetation and precipitation are fundamental factor because they provide breeding sites for mosquitos (Cui et al., 2009; Hao et al., 2011).

Similar analysis is conducted for the two types of malaria present in Peru separately, i.e., *P. Falciparum* and *P. Vivax*. The results of the correlation coefficients and multivariate analysis for the *P. Falciparum* prevalence are shown in Table 5 and Table 6.

Table 5. Correlation matrix of variables used in the multivariate analysis for the P. Falciparum model

	T9	NDVI	E2	SM1	P7
T9	1				
NDVI	0.044	1			
E2	-0.075	0.44	1		
SM1	-0.4	-0.43	-0.36	1	
P7	-0.12	0.0005	-0.017	0.12	1

Based on the coefficient estimates, the calculated IRR value for the *P. Falciparum* prevalence indicates that there is an increase in malaria prevalence associated with an increase in E2, P7, and NDVI. E7 and SM1 (yearly average soil moisture) are negatively associated with malaria prevalence.

Table 6. Multivariate analysis for the P. Falciparum prevalence

Coefficient	Estimate	95% CI	IRR – Fixed Effects
β_0	-612.94	-639.12, -586.77	
β_{T9}	-0.0097	-0.010, -0.0094	0.99
β_{E2}	0.69	0.59, 0.76	1.97
β_{SM1}	-1.79	-2.37, -1.23	0.166
B_{P7}	0.0084	0.0069, 0.0098	1.008
β_{NDVI}	9.14	8.43, 9.84	9286.17
β_{year}	0.3	0.29, 0.32	1.354
Random effects	3.8	1.37, 10.52	

The results of the multivariate analysis and the correlation coefficients for the *P. Vivax* prevalence are shown in Table 7 and Table 3, respectively. Based on the coefficient estimates, the IRR value for the *P. Vivax* prevalence indicates that there is an increase in malaria prevalence associated with an increase in E2, SM4, P2 and NDVI. The number of days with temperature above 25°C is associated with a decrease in the malaria prevalence. These results are the same as what is observed for the total malaria prevalence multivariate

model, which demonstrates that it is sufficient to use total cases instead of the two separate values.

Table 7. Multivariate analysis for the P. Vivax prevalence

Coefficient	Estimate	95% CI	IRR – Fixed Effects
β_0	-442.64	-453.78, -431.49	
β_{T9}	-0.003	-0.0031, -0.0028	0.997
β_{E2}	0.81	0.78, 0.85	2.254
β_{SM4}	0.003	0.0029, 0.0031	1.003
β_{P2}	0.63	0.60, 0.65	1.871
β_{NDVI}	0.5	0.16, 0 .83	1.642
β_{year}	0.22	0.22, 0.23	1.247
Random effects	3.11	1.16, 8.37	

However, the *P. Falciparum* and *P. Vivax* multivariate models have a difference in which variable transformation is included in the model. For the *P. Falciparum* cases, number of days with rainfall over 15 mm is included rather than cumulative rainfall in the wet season. Moreover, for the *P. Falciparum* model, average yearly soil moisture is included rather than the number of days with soil moisture above 0.400 m^3/m^3.

The main difference between the two malaria types consists in the severity of the disease. *P. Falciparum* can cause more severe effects because it multiplies more rapidly in the blood, whereas *P. vivax* has dormant liver stages and can relapse several months after the infecting mosquito bite (CDC, 2015). In Peru, the majority of reported cases are *P. vivax,* but *P. Falciparium* transmission continues to occur. For this study, a multivariate model is created for the two malaria types and the only difference is for the *P. Falciparum* model, where a different transformation is used for precipitation and soil moisture.

Ultimately, it is concluded that since the climate and environmental factors have a more impact on the vector instead of the parasite, there is not a need to separately test the two malaria types. Specifically, higher temperature increases the number of blood meals, and number of eggs laid which increases the number of mosquitos in a given area. Rainfall creates the sites the mosquitoes need to breed so increase in rainfall can increase the number of mosquitos. The rainfall also relates to the vegetation index and the soil moisture which affects the mosquito breeding sites (Open Learn Works, 2016).

CONCLUSION

In this study, we analyze malaria prevalence data from Health Center in the Loreto Department of Peru, located in the Amazon basin, to assess the association between malaria and various environmental and climate factors. This region has a low population density and a wide land area, making it difficult to directly collect high-resolution environmental data. Remote sensing and modeling techniques are particularly useful in remote areas like Loreto, providing temporally and spatially continuous information regarding the climate and environmental conditions. Results from this study indicate the climate and environmental conditions that increase malaria prevalence in Loreto include high NDVI (which corresponds to abundant and healthy vegetation), high precipitation amount in the wet season, elevation above 100m, and high soil moisture.

Malaria transmission is extremely complex. It is impacted by many factors, including gender, age, housing conditions, weather conditions, distance to standing water and occupation and more (Ayele et al., 2012). This work solely focuses on understanding the role of climate and environmental factors, which is particularly useful, since this data is free and generally easy to access in any area of the world. Thus, similar analysis could be conducted in any other region where vector-borne disease data are available. Further studies would benefit by also incorporating other factors such as population age, occupation, access to healthcare, education, and distance from standing water. Future work

should also include information regarding the population distribution data to better characterize the environmental and climate conditions the population is exposed to.

Because malaria prevalence is only available on a yearly scale, a lag analysis could not be performed. In a lag analysis, malaria cases are compared with the environmental or climate conditions that occurred in a certain past period, thus incorporating time for mosquitoes to breed and for malaria symptoms to show.

Despite the limitations, results from this study are especially useful for healthcare workers in malaria endemic areas and have the potential of being integrated within a malaria elimination plan or of being used to create a risk map. This work contributes a further understanding of climate variability's effect on malaria prevalence and has the potential to inform future efforts to develop a global framework for predicting and monitoring the spread of malaria, especially in relation to climate variables and their change over time.

APPENDICES

Appendix A: Univariate analysis for total cases

Parameter	Incidence rate ratio	(95% CI)	AIC
P_1	1.554854	1.532202, 1.577842	219412.8
P_2	2.09352	2.049979, 2.137986	218222
P_3	1.497786	1.452433, 1.544554	222214.4
P_4	.9929961	.9927128, .9932795	220440.6
P_5	1.012988	1.012498, 1.013478	220110.1
P_6	1.014243	1.01371, 1.014777	220088.8
P_7	1.017246	1.016641 , 1.017852	219707.2
P_8	1.062489	1.059469, 1.065518	221169.8
T_1	.6835416	.6749167, .6922767	219411.7
T_2	1.047394	1.034728, 1.060214	222804
T_3	.6881428	.6815615, .6947877	216885.1
T_4	.9299462	.9267455, .9331581	221102.7
T_5	.9008346	.896607, .9050821	220892.3
T_6	.9229218	.912395, .9335701	222675.9
T_7	.8212745	.8128405, .829796	221505.7
T_8	.9983729	.9979627, .9987834	222799.4
T_9	.9940594	.9939227, .994196	215851.4
T_{10}	1.983345	1.815193, 2.167074	222701.3
T_{11}	1		222857.9
T_{12}	.8128269	.806065, .8196456	220360
T_{13}	.8870561	.883481, .8906456	219256.7
H_1	4.18e-63	1.11e-67 , 1.58e-58	222143.3
H_2	1.5e-113	8.3e-119 , 2.5e-108	221067.9
H_3	6.85e-25	3.48e-28 , 1.35e-21	222652.3
H_4	.99591	.9957509, .996069	220335.4
H_5	.997846	.9976571, .9980348	222362.8
H_6	.9961877	.9959124, .9964632	222121.2
SM_1	885.1863	706.5091, 1109.051	219376.1
SM_2	279.6238	221.4949, 353.008	220638.8
SM_3	803.9682	656.6941 , 984.2708	218650.1
SM_4	1.003947	1.003837, 1.004057	218015.7
SM_5	.9981079	.9980063, .9982095	221520.7
EVI	1.246832	.8858714 , 1.754871	222858.3

NDVI	1485.89	1148.576, 1922.267	219755.7
E_1	1.000073	.9999382, 1.000208	222858.8
E_2	2.725138	2.652726, 2.799526	216387.8
E_3	1.011736	.9918508, 1.03202	222858.6
E_4	.5680049	.5521997, .5842626	221168.3

Appendix B: Univariate analysis for P. Falciparum cases

Parameter	Incidence rate ratio	(95% CI)	AIC
P_1	1.271951	1.2262, 1.319409	50210.91
P_2	1.505163	1.426906, 1.587712	50153.46
P_3	1.2508	1.158648, 1.350282	50343.34
P_4	1.001588	1.000933, 1.002244	50353.25
P_5	1.006574	1.00538, 1.00777	50258.29
P_6	1.009415	1.008102, 1.010729	50176.48
P_7	1.011509	1.010025, 1.012997	50143.23
P_8	1.052478	1.045016, 1.059992	50182.13
T_1	.5160248	.5020682, .5303694	48190.18
T_2	.7899309	.7698824, .8105015	50059.08
T_3	.5466947	.5348352, .5588172	47360.56
T_4	.9829511	.9750028, .9909642	50358.36
T_5	.9451331	.9346797, .9557034	50274.97
T_6	.7939135	.7731474, .8152375	50101.46
T_7	.7994203	.7795686, .8197776	50082.11
T_8	.9893658	.9884632 , .9902693	49837.17
T_9	.9896344	.9893101, .9899587	46616.65
T_{10}	2.724164	2.343268, 3.166974	50273.01
T_{11}	1		50373.68
T_{12}	.8528462	.8360345, .8699959	50119.94
T_{13}	.9697974	.9609156, .9787613	50332.45
H_1	3.7e-140	6.1e-151, 2.3e-129	49720.31
H_2	2.4e-210	1.0e-222, 5.4e-198	49243.7
H_3	5.66e-70	1.03e-77, 3.11e-62	50062.14
H_4	.9912787	.9909074, .9916503	48306.18
H_5	.9983938	.9979688, .9988189	50321.11
H_6	.9947558	.994148, .9953641	50088.89
SM_1	132.1643	79.19442, 220.5636	50029.02
SM_2	10.27684	5.97496, 17.67601	50305.45
SM_3	409.2887	258.3627, 648.38	49719.67
SM_4	1.000366	1.000084, 1.000648	50369.24
SM_5	.9978821	.9976551, .9981092	50038.23
EVI	65.62482	28.95229, 148.7487	50274.8
NDVI	1217627	648592.3, 2285897	48441.68
E_1	1.000636	1.00035, 1.000922	50357.63
E_2	3.657058	3.381051, 3.955597	48988.77
E_3	1.19057	1.137488, 1.246129	50319.66
E_4	.4564128	.426369, .4885737	49789.44

Appendix C: Univariate analysis for P. Vivax cases

Parameter	Incidence rate ratio	(95% CI)	AIC
P_1	1.615027	1.589359, 1.64111	184574.5
P_2	2.22569	2.175368, 2.277175	183398.1
P_3	1.548607	1.497473, 1.601487	187353.1
P_4	.9911521	.9908377 , .9914666	184823.6
P_5	1.014236	1.013699, 1.014773	185230.1
P_6	1.015167	1.014584, 1.015751	185358.8
P_7	1.018349	1.017687, 1.019012	185006.4
P_8	1.064313	1.061011, 1.067624	186487.5
T_1	.7345217	.7241225, .7450701	186183.2
T_2	1.126282	1.110849, 1.141928	187698
T_3	.7264618	.7187523, .734254	184456
T_4	.9189709	.915473, .9224822	186029
T_5	.8915243	.8869055, .8961671	186025.9
T_6	.95312	.9410716, .9653228	187932.2
T_7	.8258267	.8165289, .8352303	186922
T_8	1.000653	1.000193, 1.001112	187978.5
T_9	.9949669	.9948159, .995118	183856.1
T_{10}	1.772129	1.587929, 1.977695	187910.2
T_{11}	1		187984.3
T_{12}	.8042731	.7969028, .8117116	185704.4
T_{13}	.8697518	.8658423, .873679	184008
H_1	1.22e-45	1.05e-50 , 1.42e-40	187683.7
H_2	3.08e-91	4.82e-97, 1.97e-85	187047.8
H_3	1.65e-14	3.71e-18 , 7.33e-11	187931.3
H_4	.9969136	.9967375, .9970898	186813.3
H_5	.9976999	.9974889, .9979108	187531.9
H_6	.9964989	.99619, .9968079	187491.3
SM_1	1404.564	1092.356, 1806.005	184779.4
SM_2	601.7187	464.5387, 779.4085	185656.4
SM_3	954.1796	761.7007, 1195.297	184410.6
SM_4	1.004635	1.004515, 1.004755	182365.6
SM_5	.998161	.9980473, .9982746	186975.8
EVI	.5330091	.3658141, .7766204	187975.5
NDVI	396.5746	298.9542, 526.0721	186259.2
E_1	.9999342	.9997814, 1.000087	187985.6
E_2	2.607151	2.5334, 2.68305	182838.7
E_3	.9764676	.955138, .9982736	187981.8
E_4	.5975844	.5793285, .6164155	186838.3

Appendix D: Acronyms

	Precipitation
P_1	Cumulative in the whole year (m)
P_2	Cumulative in the wet season (m)
P_3	Cumulative in the dry season (m)
P_4	Number of days without rainfall
P_5	Number of days with rainfall over 10 mm
P_6	Number of days with rainfall over 12 mm
P_7	Number of days with rainfall over 15 mm
P_8	Number of months with rainfall over 60 mm
	Temperature
T_1	Average daily temperature in the whole year (°C)
T_2	Average daily temperature in the wet season (°C)
T_3	Average daily temperature in the dry season (°C)
T_4	Mean maximal temperature in the whole year (°C)
T_5	Mean maximal temperature in the wet season (°C)
T_6	Mean minimal temperature in the whole year (°C)
T_7	Mean minimal temperature in the dry season (°C)
T_8	Number of days with temperature above 30 °C
T_9	Number of days with temperature above 25 °C
T_{10}	Number of days with temperature under 20 °C
T_{11}	Number of days with temperature under 15 °C
T_{12}	Average temperature at night (6pm/6am) in the wet season (°C)
T_{13}	Average temperature at night (6pm/6am) in the dry season (°C)
	Humidity
H_1	Average daily humidity in the whole year (kg vapor × $kgair^{-1}$)
H_2	Average daily humidity in the wet season (kg vapor × $kgair^{-1}$)
H_3	Average daily humidity in the dry season (kg vapor × $kgair^{-1}$)
H_4	Number of days with humidity above 0.018 kg vapor × $kgair^{-1}$
H_5	Number of days with humidity under 0.016 kg vapor × $kgair^{-1}$
H_6	Number of days with humidity under 0.014 kg vapor × $kgair^{-1}$
	Soil Moisture
SM1	Average daily soil moisture in the whole year (m^3*m^{-3})
SM_2	Average daily soil moisture in the wet season (m^3*m^{-3})
SM_3	Average daily soil moisture in the dry season (m^3*m^{-3})
SM_4	Number of days with soil moisture above $0.400m^3*m^{-3}$
SM_5	Number of days with soil moisture under $0.300m^3*m^{-3}$
	Vegetation Index
EVI	Enhanced vegetation index
NDVI	Normalized Difference Vegetation Index
	Elevation
E_1	Elevation (m)
E_2	Elevation above 100 m
E_3	Elevation above 150 m

E_4	Elevation above 200 m

REFERENCES

Abizaid, C. (2005). Geographical field note. An anthropogenic meander cutoff along the Ucayali river, Peruvian Amazon. The Geographical Review 95: 1, 122-135.

Ahmed, S. E. (2014). *Penalty, Shrinkage and Pretest Strategies*. Cham: Springer International Publishing. Retrieved from http://link.springer.com/10.1007/978-3-319-03149-1

Akhwale, W. S., Lum, J. K., Kaneko, A., Eto, H., Obonyo, C., Björkman, A., & Kobayakawa, T. (2004). Anemia and malaria at different altitudes in the western highlands of Kenya. *Acta Tropica*, *91*(2), 167–175. http://doi.org/10.1016/j.actatropica.2004.02.010

Aramburú Guarda, J., Ramal Asayag, C., & Witzig, R. (1999). Malaria reemergence in the Peruvian Amazon region. *Emerging Infectious Diseases*, *5*(2), 209–215. http://doi.org/10.3201/eid0502.990204

Attenborough, R. D., Burkot, T. R., & Gardner, D. S. (1997). Altitude and the risk of bites from mosquitoes infected with malaria and filariasis among the Mianmin people of Papua New Guinea. *Transactions of The Royal Society of Tropical Medicine and Hygiene*, *91*(1), 8–10. http://doi.org/10.1016/S0035-9203(97)90373-4

Ayele, D. G., Zewotir, T. T., & Mwambi, H. G. (2012). Prevalence and risk factors of malaria in Ethiopia. *Malaria Journal*, *11*, 195. http://doi.org/10.1186/1475-2875-11-195

Bautista, C. T., Chan, A. S. T., Ryan, J. R., Calampa, C., Roper, M. H., Hightower, A. W., & Magill, A. J. (2006). Epidemiology and Spatial Analysis of Malaria in the Northern Peruvian Amazon. *The American Journal of Tropical Medicine and Hygiene*, *75*(6), 1216–1222.

Bødker, R., Akida, J., Shayo, D., Kisinza, W., Msangeni, H. A., Pedersen, E. M., &

Lindsay, S. falciW. (2003). Relationship Between Altitude and Intensity of Malaria Transmission in the Usambara Mountains, Tanzania. *Journal of Medical Entomology*, *40*(5), 706–717. http://doi.org/10.1603/0022-2585-40.5.706

Centers for Disease Control and Prevention (CDC). (2012). CDC - Malaria - About Malaria - Where Malaria Occurs. Retrieved March 16, 2015, from http://www.cdc.gov/malaria/about/distribution.html

Centers for Disease Control and Prevention CDC. (2015). CDC - Malaria - About Malaria - Biology - Malaria Parasites. Retrieved April 15, 2016, from http://www.cdc.gov/malaria/about/biology/parasites.html

Chuquiyauri, R., Paredes, M., Peñataro, P., Torres, S., Marin, S., Tenorio, A., Brouwer, K. C., Abeles, S., Llanos-Cuentas, A., Gilman R. H., Kosek M.,& Vinetz, J. M. (2012). Socio-Demographics and the Development of Malaria Elimination Strategies in the Low Transmission Setting. *Acta Tropica*, *121*(3), 292–302. http://doi.org/10.1016/j.actatropica.2011.11.003

CIA. (2015). The world fact book: Peru. Retrieved December 11, 2015, from https://www.cia.gov/library/publications/the-world-factbook/geos/pe.html

Cui, X., Parker, D. J., & Morse, A. P. (2009). The Drying Out of Soil Moisture following Rainfall in a Numerical Weather Prediction Model and Implications for Malaria Prediction in West Africa. *Weather and Forecasting*, *24*(6), 1549–1557. http://doi.org/10.1175/2009WAF2222240.1

Cotter, C., Sturrock, H. J. W., Hsiang, M. S., Liu, J., Phillips, A. A., Hwang, J., … Feachem, R. G. A. (2013). The changing epidemiology of malaria elimination: new strategies for new challenges. Lancet (London, England), 382(9895), 900–911. http://doi.org/10.1016/S0140-6736(13)60310-4

Drakeley, C. J., Carneiro, I., Reyburn, H., Malima, R., Lusingu, J. P. A., Cox, J., … Riley, E. M. (2005). Altitude-Dependent and -Independent Variations in Plasmodium falciparum Prevalence in Northeastern Tanzania. *Journal of Infectious Diseases*, *191*(10), 1589–1598. http://doi.org/10.1086/429669

Feachem, R. G., Phillips, A. A., Hwang, J., Cotter, C., Wielgosz, B., Greenwood, B. M.,

Sabot, O., Rodriguez M. H., Abeyasinghe, R. R., Ghebreyesus, T. A., & Snow, R. W. (2010). Shrinking the malaria map: progress and prospects. The Lancet, 376(9752), 1566–1578. http://doi.org/10.1016/S0140-6736(10)61270-6

Ferguson, H. M., Dornhaus, A., Beeche, A., Borgemeister, C., Gottlieb, M., Mulla, M. S., Gimnig, J. E., Fish, D., Killeen, G. F., & Killeen, G. F. (2010). Ecology: A Prerequisite for Malaria Elimination and Eradication. *PLoS Med*, *7*(8), e1000303. http://doi.org/10.1371/journal.pmed.1000303

Gardner, W., Mulvey, E. P., & Shaw, E. C. (1995). Regression analyses of counts and rates: Poisson, overdispersed Poisson, and negative binomial models. *Psychological Bulletin*, *118*(3), 392–404.

Gibbons RD, Hedeker D, DuToit S (2010) Advances in Analysis of Longitudinal Data. In: NolenHoeksema S, Cannon TD, Widiger T, editors. Annual Review of Clinical Psychology, Vol 6 Palo Alto: Annual Reviews; pp. 79–107

Griffing, S. M., Gamboa, D., & Udhayakumar, V. (2013). The history of 20th century malaria control in Peru. *Malaria Journal*, *12*, 303. http://doi.org/10.1186/1475-2875-12-303

Githeko, A.K. and Ndegwa, W. (2001). Predicting malaria epidemics in the Kenyan highlands using climate data: A tool for decision makers. *Global Change and Human Health* **2**: 54-63

Gopalan, K., Wang, N.-Y., Ferraro, R., & Liu, C. (2010). Status of the TRMM 2A12 Land Precipitation Algorithm. *Journal of Atmospheric and Oceanic Technology*, *27*(8), 1343–1354. http://doi.org/10.1175/2010JTECHA1454.1

GOREL. (2006). Mapa de la hidrografía principal y secundaria, departamento de Loreto. Gobierno regional de Loreto, Iquitos.

Habib, E., Henschke, A., & Adler, R. F. (2009). Evaluation of TMPA satellite-based research and real-time rainfall estimates during six tropical-related heavy rainfall events over Louisiana, USA. *Atmospheric Research*, *94*(3), 373–388. http://doi.org/10.1016/j.atmosres.2009.06.015

Hao, F., Zhang, X., Ouyang, W., Skidmore, A. K., & Toxopeus, A. G. (2011). Vegetation NDVI Linked to Temperature and Precipitation in the Upper Catchments of Yellow River. *Environmental Modeling & Assessment*, *17*(4), 389–398. http://doi.org/10.1007/s10666-011-9297-8

Herring, J. W. (2000). Measuring Vegetation (NDVI & EVI). Retrieved April 2, 2016, from http://earthobservatory.nasa.gov/Features/MeasuringVegetation/measuring_vegetation_2.php

Herrera, S., Quiñones, M. L., Quintero, J. P., Corredor, V., Fuller, D. O., Mateus, J. C., Calzada, J.E., Gutierrez, J. B., Llanos, A., Soto, E., Menendez, C., Wu, T., Alondo, P., Carrasquilla, G., Galinski M., Beier, J. C., & Arevalo-Herrera, M. (2012). Prospects for malaria elimination in non-Amazonian regions of Latin America. *Acta Tropica*, *121*(3), 315–323. http://doi.org/10.1016/j.actatropica.2011.06.018

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., … Stocker, E. F. (2007). The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. *Journal of Hydrometeorology*, *8*(1), 38–55. http://doi.org/10.1175/JHM560.1

Huffman, G. J., Adler, R. F., Bolvin, D. T., & Nelkin, E. J. (2010). The TRMM Multi-Satellite Precipitation Analysis (TMPA). In M. Gebremichael & F. Hossain (Eds.), *Satellite Rainfall Applications for Surface Hydrology* (pp. 3–22). Springer Netherlands. Retrieved from http://link.springer.com/chapter/10.1007/978-90-481-2915-7_1

Japan Space Systems. (2011). Data specification. Retrieved April 2, 2016, from http://www.jspacesystems.or.jp/ersdac/GDEM/E/4.html

Jiang, Z., Huete, A. R., Didan, K., & Miura, T. (2008). Development of a two-band enhanced vegetation index without a blue band. *Remote Sensing of Environment*, *112*(10), 3833–3845. http://doi.org/10.1016/j.rse.2008.06.006

Jones, J. W., Turell, M. J., Sardelis, M. R., Watts, D. M., Coleman, R. E., Fernandez, R., Carbajal, F., Pecor, J. E., Calampa, C., & Klein, T. A. (2004). Seasonal distribution,

biology, and human attraction patterns of culicine mosquitoes (Diptera: Culicidae) in a forest near Puerto Almendras, Iquitos, Peru. *Journal of Medical Entomology*, *41*(3), 349–360.

Kummerow, C. D., Ringerud, S., Crook, J., Randel, D., & Berg, W. (2010). An Observationally Generated A Priori Database for Microwave Rainfall Retrievals. *Journal of Atmospheric and Oceanic Technology*, *28*(2), 113–130. http://doi.org/10.1175/2010JTECHA1468.1

Kvist, L.P. & G. Nebel (2001). A review of Peruvian flood plain forests: ecosystems, inhabitants and resource use. Forest Ecology and Management 150: 1-2, 3-26.

Long, J.S. (1997). Regression Models for Categorical and Limited Dependent Variables, Chapter 8. Thousand Oaks, CA: Sage Publications.

Maggioni, V., Meyers, P. C., & Robinson, M. D. (2016). A Review of Merged High-Resolution Satellite Precipitation Product Accuracy during the Tropical Rainfall Measuring Mission (TRMM) Era. *Journal of Hydrometeorology*, *17*(4), 1101–1117. http://doi.org/10.1175/JHM-D-15-0190.1

Mantas, V. M., Liu, Z., Caro, C., & Pereira, A. J. S. C. (2015). Validation of TRMM multi-satellite precipitation analysis (TMPA) products in the Peruvian Andes. *Atmospheric Research*, *163*, 132–145. http://doi.org/10.1016/j.atmosres.2014.11.012

Meade, M. S., & Emch, M. (2010). *Medical Geography* (Third). New York: Guilford Publications.

Melesse, A. M. (2011). *Nile River Basin: Hydrology, Climate and Water Use*. Springer Science & Business Media.

Modrek, S., Liu, J., Gosling, R., & Feachem, R. G. (2012). The economic benefits of malaria elimination: do they include increases in tourism? Malaria Journal, 11(1), 244. http://doi.org/10.1186/1475-2875-11-244

Moonen, B., Cohen, J. M., Snow, R. W., Slutsker, L., Drakeley, C., Smith, D. L., Abeyasinghe, R. R., Rodriguez, M. H., Maharaj, R., Tanner, M., & Targett, G. (2010). Operational strategies to achieve and maintain malaria elimination. The Lancet, 376(9752), 1592–1603. http://doi.org/10.1016/S0140-6736(10)61269-X

NASA. (2015). TRMM Data Downloads | Precipitation Measurement Missions. Retrieved April 9, 2016, from http://pmm.nasa.gov/data-access/downloads/trmm

NASA. (2016). Reverb | ECHO. Retrieved April 9, 2016, from http://reverb.echo.nasa.gov/reverb/#utf8=%E2%9C%93&spatial_map=satellite&spatial_type=rectangle

National Geographic. (2015). Peru Factors. Retrieved December 11, 2015 from http://travel.nationalgeographic.com/travel/countries/peru-facts/

National Geographic Society, & Society, N. G. (2011). Equator. Retrieved April 6, 2016, from http://education.nationalgeographic.org/encyclopedia/equator/

Open Learn Works. (2016). Communicable Diseases Module: 6. Factors that Affect Malaria Transmission: View as single page. Retrieved April 16, 2016, from http://www.open.edu/openlearnworks/mod/oucontent/view.php?id=89&printable=1

Parham, P. E., & Michael, E. (2010). Modeling the Effects of Weather and Climate Change on Malaria Transmission. *Environmental Health Perspectives*, *118*(5), 620–626. http://doi.org/10.1289/ehp.0901256

Patz, J. A., Campbell-Lendrum, D., Holloway, T., & Foley, J. A. (2005). Impact of regional climate change on human health. *Nature*, *438*(7066), 310–317. http://doi.org/10.1038/nature04188

Patz, J. A., Strzepek, K., Lele, S., Hedden, M., Greene, S., Noden, B., … Beier, J. C. (1998). Predicting key malaria transmission factors, biting and entomological inoculation rates, using modelled soil moisture in Kenya. *Tropical Medicine & International Health: TM & IH*, *3*(10), 818–827.

Ren, Z., Wang, D., Hwang, J., Bennett, A., Sturrock, H. J. W., Ma, A., … Wang, J. (2015). Spatial-Temporal Variation and Primary Ecological Drivers of Anopheles sinensis Human Biting Rates in Malaria Epidemic-Prone Regions of China. *PLoS ONE*, *10*(1). http://doi.org/10.1371/journal.pone.0116932

Reichle, R. H., Koster, R. D., De Lannoy, G. J. M., Forman, B. A., Liu, Q., Mahanama, S. P. P., & Touré, A. (2011). Assessment and Enhancement of MERRA Land Surface

Hydrology Estimates. *Journal of Climate*, *24*(24), 6322–6338. http://doi.org/10.1175/JCLI-D-10-05033.1

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Lawrence, T., Kim, G., Bloom, S., Chen, J., Collins, D., Conaty, A., da Silva, A., Gu, Wei., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C. R., Reichle, R., Robertson, F. R., Ruddick, A. G., Sienkiewicz, M., & Woollen, J. (2011). MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *Journal of Climate*, *24*(14), 3624–3648. http://doi.org/10.1175/JCLI-D-11-00015.1

Sabot, O., Cohen, J. M., Hsiang, M. S., Kahn, J. G., Basu, S., Tang, L., Zheng, B., Gao, Q., Zou, L., Tatarsky, A., Aboobakar, S., Usas, J., Barrett, S., Cohen, J. L., Jamison, D. T., & Feachem, R. G. (2010). Costs and financial feasibility of malaria elimination. *The Lancet*, *376*(9752), 1604–1615. http://doi.org/10.1016/S0140-6736(10)61355-4

Schofer, E. (2007). Count Models 1 [PowerPoint Slides]. Retrieved from http://www.socsci.uci.edu/~schofer/2007soc8811/pub/Class%2012%20Count%20Models%201%201.0.ppt.

STATA. (2013). STATA MULTILEVEL MIXED-EFFECTS REFERENCE MANUAL RELEASE 13. Retrieved from https://www.stata.com/manuals13/me.pdf

SJSU. (2012). Calculating Prevalences and Incidences. Retrieved from: http://www.sjsu.edu/faculty/gerstman/hs161/pre-inc-key-points.PDF

Tuteja, R. (2007). Malaria − an overview. *FEBS Journal*, *274*(18), 4670–4679. http://doi.org/10.1111/j.1742-4658.2007.05997.x

Vittor, A. Y., Gilman, R. H., Tielsch, J., Glass, G., Shields, T., Lozano, W. S., … Patz, J. A. (2006). The Effect of Deforestation on the Human-Biting Rate of Anopheles Darlingi, the Primary Vector of Falciparum Malaria in the Peruvian Amazon. *The American Journal of Tropical Medicine and Hygiene*, *74*(1), 3–11.

World Health Organization. (2015). WHO | World Malaria Report 2015. Retrieved March 16, 2016, from

http://apps.who.int/iris/bitstream/10665/200018/1/9789241565158_eng.pdf?ua=1

Zaitchik, B., Feingold, B., Spangler, K., & Kosek, M. (2012). Towards Spatially Explicit Malaria Risk Models for the Peruvian Amazon. Retrieved March 16, 2015, from http://www.isprs.org/proceedings/XXXVIII/8-C23/pdf/Zaitchik.pdf

BIOGRAPHY

Aneela Mousam graduated from Yorktown High School, Arlington, Virginia, in 2009. She received her Bachelor of Science from Virginia Tech in 2013.