

Assessing the Relationship between various Climatic Risk Factors & Mosquito Abundance in Recife, Brazil

Anwar Musah
Centre for Digital Public Health & Emergencies
University College London
United Kingdom
a.musah@ucl.ac.uk

Adrian Rubio-Solis
Centre for Engineering & Industrial Development – CIDESI
Mexico
adrian.rubio@cidesi.edu.mx

Georgiana Birjovanu
Centre for Digital Public Health & Emergencies
University College London
United Kingdom
georgiana.birjovanu.15@ucl.ac.uk

Wellington P. dos Santos
Department of Biomedical Engineering
Federal University of Pernambuco
Brazil
wellington.santos@ufpe.br

Tiago Massoni
Department Systems & Computing
Federal University of Campina Grande Brazil
massoni@dsc.ufcg.edu.br

Patty Kostkova
Centre for Digital Public Health & Emergencies
University College London
United Kingdom
p.kostkova@ucl.ac.uk

ABSTRACT

Recife is one of many regions in the North Eastern section of Brazil to be hit hard by the Zika virus outbreak in 2015. Zika infection is caused by the arboviruses transmitted by the *Aedes* mosquitoes which are abundant in Brazil. It is a known fact that increased abundance of the *Aedes* mosquito is typically associated with standing (or stagnant) water which serves as a reservoir or hotspot for breeding. Apart from the presence of standing (or stagnated) water in human dwellings, it is important to consider other intermediate factors that drive the mosquitoes to increase in population size. We consider the example of a set of climate conditions such as land surface temperature, humidity, precipitation, seasonality etc. These are typical risk factors that must be taken into consideration especially for the spatial and early warning prediction of breeding hotspots for the *Aedes* mosquito species responsible for transmitting Zika viruses in Brazil. In this work-in-progress study, we bring together historical records from a mosquito surveillance and control programme conducted by the Centre for Environmental Surveillance Agency in Recife and open source climate information. We use a population-based ecological study design to explore the regional-level spatial-temporal relationship between the prevalence of households detected to have increased mosquito larvae density and environmental risk factors such as temperature and precipitation.

CSS CONCEPTS

• Probability and statistics • Statistical paradigms • Regression analysis; • Information system applications • GIS

KEYWORDS

Vector control; surveillance; Zika virus; climate conditions; spatial-temporal models

ACM Reference Format

Anwar Musah, Georgiana Birjovanu, Adrian Rubio-Solis, Wellington P. dos Santos, Tiago Massoni, Patty Kostkova. 2019. Assessing the Relationship between various Climatic Risk Factors & Mosquito Abundance in Recife, Brazil. In the DPH'19: 2019 International Digital Public Health Conference Proceedings (2019), November 20-23, 2019, Marseille, France. 4 pages. <https://doi.org/10.1145/3357729.3357744>

1 INTRODUCTION

Neglected tropical diseases (NTDs) such as Zika, Chikungunya and Dengue are all vector borne arboviruses transmitted by the *Aedes* mosquitoes [1, 2]. These classes of NTDs are indiscriminate; they infect infants, children and adult causing severe morbidity and mortality if not treated. For decades, they remain endemic in the tropical (and sub-tropical) parts of Sub-Saharan Africa, Latin America and South East Asia claiming the lives of more than 700,000 people annually [3]. In the case of Zika virus, it has devastating health implications on newborn babies causing microencephaly which is a form of congenital birth defect where the head is malformed.

Zika Virus has become a global concern and a major health priority following the outbreak in Brazil in 2015. It is estimated that the outbreak has resulted in close to 1.5 million cases of infection with the virus of which majority are spatially concentrated in the districts that make-up the Northeastern region of Brazil [4]. At the moment, there is no effective vaccine or treatment for preventing Zika (or any of the arboviruses) in an infected human and thus the situation remains complex. Currently, the only solution to preventing Zika transmission from the *Aedes* mosquitoes to the human population is to

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

DPH'19, November 20-23, 2019, Marseille, France.

© 2019 Association for Computing Machinery.

ACM ISBN 978-1-4503-7208-4/19/11...\$15.00.

<https://doi.org/10.1145/3357729.3357744>

implement a comprehensive vector control program that integrates routine surveillance (i.e. by public health agents, fieldworkers and citizen-driven participatory reporting) and technology (e.g. mobile phone application, geographical information systems (GIS) etc.) to target mosquito populations and sources of standing (or stagnated) water that serve as hotspots for mosquitoes to breed [5 – 7].

There are several risk factors that have an impact on the presence and abundance of mosquitoes, specifically the *Aedes* species in an environment. These include poor sanitation practices in terms of disposing waste substances (i.e. from kitchen, bath and lavatory) in the environment, lack of proper drainage infrastructure in home and land-use (e.g. mechanics & car-washes with discarded tires, outdoor bathrooms etc.) [8, 9]. In a low-income setting - these are typical anthropogenic-related activities that contribute to water remaining stagnated for mosquito breeding in human dwellings. Consider another example such as the use of water containers in the form of vases, flower pots and water tanks which serves a beneficial purpose to humans but inadvertently act as breeding hotspot for the *Aedes* mosquitoes [8, 9]. So far, the best strategy to counter this unfortunate trend of the effect of anthropogenic-related factors is through engagement and education of the dangers these risk factors may pose. Other larger and uncontrollable risk factors such as local weather and climatic conditions such as reduced temperature and dryness [8, 9], as well as elevated levels of humidity and rainfall (or precipitation) vastly contribute to the increased surface areas for water to remain, and also making the environment favorable for mosquito breeding which increases population size and cycle [8, 9].

When predicting mosquito populations for vector control, it is imperative that contextual risk factors are taken into consideration – especially, if it is going to be utilized within an early warning system application for forecasting future outbreaks. This research, is in collaboration with the Brazilian health ministries and environmental agency, and is currently at the early stages of developing a vector control surveillance application for real-time geospatial triangulation of affected households in endemic study areas of Recife, Campina Grande and Olinda for predicting potential breeding hotspots and increased abundance of *Aedes* mosquitoes. But first, we begin by bringing a wide range of heterogenous and contemporary data sources together in an attempt to develop a surveillance and early warning system [6]. In this study, we have assembled a modest number of climatic variables which are spatially referenced and linked to historical vector surveillance data collected in Recife. We use a spatial-temporal Bayesian regression model to determine whether a set of weather-related conditions are associated with the increased prevalence of *Aedes* mosquito species across Recife, Brazil.

2 METHODOLOGY

2.1 Vector Control & Surveillance Data

The data for this research originated from a comprehensive vector control and entomological survey that took place from

January 2009 to December 2017 in Recife, Brazil. In brief, this longitudinal dataset was compiled over the course of eight years of follow-up at a 2-month interval (often with complete information for the following months: January, March, May, July, September and November) by the Brazilian's health authorities and environmental agency. The data collected during the survey was in two phases:

1. Collection of water samples from 7 different types of water containers inside and around households to measure the presence of mosquito larvae and pupal in a house;
2. Oviposition traps were set for the collection of larval eggs laid by adult mosquitoes;

The sample from traps were sent for laboratory analysis to count the number of eggs as well as determine the *Aedes* species. Laboratory analysis was carried out by the Laboratory of Immunopathology Keizo Asami (LIKA) in Recife. The fieldwork survey was conducted at a household-level; however, all survey information was released as aggregated data at a neighbourhood-level. The overall number of households that were surveyed during the surveillance period was 1,282,882. These residential houses were visited across 72 neighbourhoods in Recife (see figure 1).

2.2 Climatic & Weather Data

This study used open source meteorological data from the Climate Forecast System Reanalysis (CFSR) database (<https://globalweather.tamu.edu/#pubs>). The CFSR data is available globally at a 38-km resolution – it contains daily and monthly weather measurements for the variables; land surface temperature, rainfall (or precipitation), wind, humidity and solar radiation from January 1979 through to July 2014 [10, 11]. For the neighbourhoods in Recife, we calculated month-specific estimates for the weather variables and matched them accordingly to the neighbourhood-level vector control and surveillance data as described in section 2.1. We first downloaded the entire 38-km gridded data for the whole of the North East region of Brazil. The grids were resampled for weather estimates to be interpolated on to a new grid template with higher resolution of 100m. A series of geostatistical models (i.e. ordinary Kriging with a fitted Gaussian theoretical model) were implemented to spatially interpolate the estimates for each weather variable over the 100m grid template to generate month-specific raster data for the weather variables. Geoprocessing of the monthly raster included cropping the grids to the size of the study area and the aggregating the pixel estimates within the neighbourhoods to derive neighbourhood-level averages for each weather variable. Note: We have generated for each of the variables land surface temperature and precipitation only. The remaining analysis for humidity, wind and solar radiation is a work-in-progress.

2.3 Study Area

The routine surveillance data described in section 2.1. is limited to Recife City which is one of eight major municipalities

that form the region of Recife. The geographical area lies between latitudes -8.17 & -7.90 (South) and longitudes -35.05 & -34.85 (West). As shown in figure 1, the area has been delineated into six major health zones for which, each health zone is subdivided into 72 lower health neighbourhoods. This research uses ecological study design for the analysis of spatiotemporal data since we are using aggregated surveillance and climate data that is month-specific, as well as the neighbourhoods being spatial areal units. The reader should bear in mind that the analysis was limited to periods between January 2010 and July 2014, inclusive, as there were no open source weather data available from August 2014 and onwards. In addition, surveillance data from August 2014 were incomplete.



Figure 1: The study area of Recife. It consists of 6 health units, wherein, the health units are further sub-divided into 72 neighbourhoods.

2.4 Determination of Prevalence Rates

The primary outcome measures for our descriptive analysis are the prevalence rates of households detected to have *Aedes* mosquito species. The study reports the overall, as well as provides a breakdown of the overall prevalence rates by health neighbourhoods in Recife observed from January 2010 to July 2014. The prevalence rate is simply calculated as the number of houses detected to have the *Aedes* mosquito that were caught in traps divided by total number of households visited during the survey period. All prevalence rates per capita were reported as 1,000 per households in Recife.

2.5 Model formation of a Spatiotemporal Bayesian Regression Model

Bayesian spatiotemporal regression model was used to assess the relationship between the neighbourhood-level prevalence of *Aedes* species in households and temperature and precipitation only. As shown in figure 1 the region of Recife comprises a set of $k = 1, \dots, 72$ non-overlapping neighbourhoods $S = \{S_1, \dots, S_{72}\}$, where vector surveillance and weather data are available for each unit at certain month t . The prevalence for the

neighbourhoods is denoted as a vector Y_{kt} which corresponds to the prevalence observed in the k^{th} neighbourhood at month t . $O_t = (O_{1t}, \dots, O_{kt})$ is a vector of offsets for the months at t . Finally, $X_{kt} = (x_{kt1}, \dots, x_{ktp})$ is a vector with p known covariates for neighbourhood k and month t . In this case, the covariates in this analysis correspond to temperature and precipitation. Briefly, the mathematical formula for modelling thus relationships is given as follows:

$$\text{Binomial: } Y_{kt} \sim \text{Binomial}(n_{kt}, q_{kt})$$

$$\log[q_{kt}/(1 - q_{kt})] = X_{kt}\beta + O_{kt} + \psi_{kt}.$$

In the binomial model n_{kt} denotes the number of houses visited in neighbourhood k and in the month t , and q_{kt} is the probability of detecting *Aedes* mosquito in households that were visited in neighbourhood k and in the month t . The regression coefficients $\beta = (\beta_1, \dots, \beta_p)$ are exponentiated in order to report odd ratios (OR) with 95% credible intervals (95% CrI). ORs are statistically significant if the 95% CrIs exclude the null value of 1 between the lower and upper limits. All statistical analysis was conducted using the software RStudio version 1.2.1335 with the ‘CARBayesST’ package [12].

3 RESULTS

The overall prevalence rate of households with the *Aedes* species to have been detected during the surveillance periods (i.e. January 2010 to July 2014) was estimated as 33.34 per 1,000 households in Recife. Figure 2 illustrates the overall spatial distribution of the prevalence rates across the 72 neighbourhoods in Recife. Areas with prevalence rates exceeding 60.0 per 1,000 households are neighbourhoods known to contain densely populated urban slum (or Favelas) in Recife.

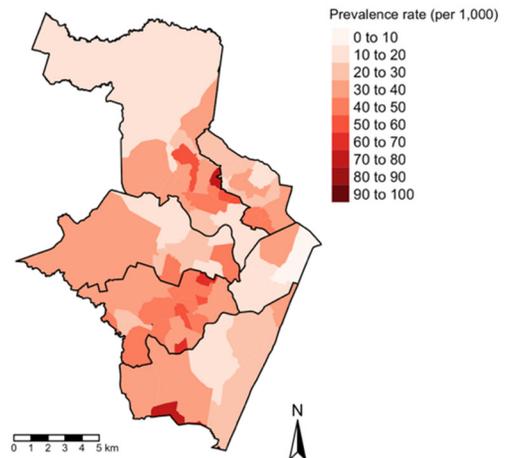


Figure 2: Neighbourhood-level prevalence rates (per 1,000 households) of houses with the *Aedes* mosquito larval in Recife (2010 to 2014)

When determining the relationship between the monthly neighbourhood-level prevalence with land surface temperature

and precipitation using the Bayesian spatiotemporal model, we found that for land surface temperature, the odds of increased prevalence for neighbourhoods reduced by 18.0% (OR: 0.82, 95% CrI: 0.70 to 1.02) for every unit increase in the levels for temperature. However, the association between these two variables are non-significant due to the null-value of 1 lies between the lower and upper limits of the CrI. For precipitation, we found that a per unit increase in the levels of precipitation in neighbourhoods significantly increases the odds for households to be detected with the *Aedes* mosquito larvae by 12.0% (OR: 1.12, 95% CrI: 1.01 to 1.21).

4 DISCUSSION

In this paper, we bring together historical entomological surveillance records which was uniquely linked to monthly weather data from an open source meteorological online database. We used a conditional autoregressive model within a Bayesian framework to assess the association between a modest number of weather-related factors that impact on mosquito abundance in the region of Recife. More importantly, and not surprisingly, is our finding for precipitation which significantly contributes to the increased *Aedes* mosquito prevalence in Recife. The elevated levels of precipitation are an important environmental risk factor to consider since it inadvertently contributes for the *Aedes* mosquito species to have an increased habitat availability for breeding. This study benefits from the fact that it uses a rich historical surveillance database linked with representative weather to make predictions for neighbourhoods in Brazil using a robust statistical model – which takes into consideration the spatial and spatiotemporal aspects that may also have potential impact on the prevalence rates across Recife.

In brief, the authors acknowledge that there are substantial limitations with the current findings – these include: the fact that this study uses aggregated data at neighbourhood-level instead of resorting to data measured at a household-level; this makes the results prone to ecological fallacy. A type of bias whereby results interpreted at an area level are assumed to be generalizable and thus representative even at a household-level. Secondly, there are several risk factors that have not been accounted for in our Bayesian model. Examples include demographic factors such as population density and levels of socioeconomic status [8, 9], as well as other environmental and climatic variables such as vegetation, humidity and aridness [8, 9], as well as the percentage of land-use in accordance to land-use classification in each neighbourhood. These factors affect not only the breeding habits, but also the feeding behavior of the *Aedes* mosquitoes. Therefore, the limited information in the models will certainly yield some residual confounding.

5 CONCLUSIONS

In conclusion, this study used a robust modelling approach taking into consideration the spatiotemporal data structure of the historical mosquito surveillance records and open source climate data to establish how risk factors such as precipitation

and land surface temperature are related to neighbourhood-level abundance of *Aedes* mosquitoes. Further work is needed in order to together more information from reliable sources to increase the predictive strength of causal patterns and decrease residual confounding. While the approach is robust, these must be interpreted with caution. For future studies, and alongside usage with the current historical surveillance data – our intention is the integration of mobile phone and m-gamification applications for data collection and real-time surveillance of mosquitoes for vector control.

REFERENCES

- [1] Anthony S Fauci and David M Morens. (2016). Zika in the Americas & another arbovirus threat. *New England Journal of Medicine*, 374(7), 601-604.
- [2] Farakh Javed, Khanzdi N Manzoor, Mubashar Ali, Irshad U Haq, Abid A Khan, Assad Zaib and Sobia Manzoor. (2018). Zika virus: what we need to know? *Journal of Basic Microbiology*, 58(1), 3-16.
- [3] Enny S Paixao, Maria G Teixeira and Laura C Rodrigues. (2018). Zika, Chikungunya & Dengue: the causes and threats of new and re-emerging arboviral diseases. *BMJ Global Health*, 3, e000530
- [4] Rooseclis Brasil Martines. (2016). Notes from the field: evidence of Zika virus infection in brain and placental tissues from two congenitally infected newborns and two fetal Brazil. *MMWR Morbidity and Mortality report* 65.
- [5] Juan D Beltran, Andrei Boscor, Wellington P. dos Santos, Tiago Massoni and Patty Kostkova. (2018). ZIKA: A new system to empower health workers and local communities to improve surveillance protocols by E-learning and to forecast Zika virus in real-time in Brazil. In *Proceedings of the 2018 International Conference in Digital Health*. ACM, 90-94.
- [6] Patty Kostkova. (2013). A Road map to integrated digital public health surveillance: the vision and the challenges. *ACM Proceedings of the 22nd International Conference on World Wide Web 2013*. 1, 687-694
- [7] De Quincey E and Patty Kostkova. (2009). Early warning and outbreak detection using social networking websites: The potential of twitter. *International Conference on Electronic Healthcare 2009*. 1, 21-24.
- [8] Thomas C McHale, Claudia M Romero-Vivas, Claudio Fronterre, Pedro A Padilla, Naomi R Waterlow, Chad D Nix, Andrew K Falconar. and Jorge Cano (2019). Spatial Heterogeneity in the Distribution of Chikungunya and Zika virus cases incidences during their 2014 to 2016 epidemics in Barranquilla, Colombia. *International Journal of Environmental Research and Public Health*, 16(10), 1759.
- [9] Claudia M Romero-Vivas and Andrew K Falconar. (2005). Investigation of relationships *Aedes Aegypti* egg, larvae, pupae and adult density indices where their main breeding sites were located indoors. *Journal of the American Mosquito Control Association*, 21(1), 15-21.
- [10] Yihun Taddele Dile and Raghavan Srinivasan. (2014). Evaluation of CFSR climate data for hydrologic prediction in data-scarce watersheds: an application in the Blue Nile River Basin. *Journal of the American Water Resources Association (JAWRA)*, 1-16. DOI: 10.1111/jawr.12182
- [11] Daniel R Fuka, M Todd Walter, Charlotte MacAlister, Arthur T Degaetano, Tammo S Steenhuis, Zachary M Easton. (2013). Using the climate forecast system reanalysis as weather input data for watershed models. *Hydrological Processes*, 28(22), 5613-5623
- [12] Duncan Lee, Alastair Rusworth and Gary Napier. (2018). Spatial-temporal areal unit modeling in R with conditional autoregressive priors using the CARBayesST package. *Journal of Statistical Software*, 84(9), 1-39.