Discussing The Behavior of West Nile Virus Related to Climate Conditions and Various Modelling Techniques to Represent the Disease

Samarth Halyal, Alexander Golubow, Anthony Lam
Department of Engineering and Computer Science, California State University Fullerton USA

Abstract**-- During the years 1962 – 1963 Camargue, South France saw an outbreak of mosquito-borne disease. Later in 1996, an epidemic occurred over the entire Europe during which countries like Romania identified about 400 cases. This disease which was causing fever and other related symptoms was named West Nile Virus (WNV). The virus is now found throughout the world across the continental United States, Europe, Africa, The Middle East, and west Asia. In this paper, we will be discussing the spread of WNV and its future trajectory based on the temperature and climate of regions. We will show how intervention strategies will change the graph of the model.**

NOMENCLATURE

WNV – West Nile Virus
S – Susceptible population
I – Infectious Population
R – Recovered/Removed
E – Exposed
CDC – Center for disease control and prevention
CSV – Comma Separated Values
M – mosquitos
B – birds
H – humans
λ – recruitment (birth) rate
β – the probability of transmission
δ – the natural mortality rate
μ – the disease-induced mortality rate
γ – the rate of progression from exposure to infectious
ν – recovery rate

1. Background

West Nile Virus (WNV) is a virus that causes West Nile Fever. It is a single-stranded RNA virus. Cases of West Nile Virus normally occur during the seasons from summer through winter when mosquito growth is high. Most people do not develop any symptoms only a fraction of people develop some symptoms like fever.

a) **The pattern of transmission**: Is primarily spread to humans through mosquito bites. The primary host of the virus is birds. The virus is then typically transmitted to other animals from mosquitos that feed on the infected birds. The virus primarily remains in a bird-mosquito-bird cycle; however, it is occasionally spread to humans and other animals when an infected mosquito bites them. The virus is not spread through coughing, sneezing, or touching an infected individual. Handling live or dead infected specimens, or eating infected specimens, will not result in an infection for humans. However, birds can spread the virus to other birds through regular contact.

b) **Vaccination:** There is no vaccination available for this virus. The reason is that most people who develop this disease are elderly and immunocompromised people. So, taking into consideration the health safety measures development of the WNV vaccine has been discarded.

c) **Social and community interventions:** Some preventive measures may be taken against it, such as not letting standing water nearby and dumping the water away if found. Babies and pregnant women must be protected with proper care and use mosquito repellents registered by the environmental protection agency. Further, all types of adulticides, larvicides, truck sprays, and aerial sprays should be used properly. Other methods include genetically modified mosquitos, irradiated mosquitos, and mosquitos with Wolbachia. [2]

d) **Behavior:** Based on the data tests on the pools of mosquitos collected by the Chicago Department of Public Health and Environmental Health program following data were presented on the behavior of the virus outbreak in a small population over Chicago. Chicago Department of Public Health maintains an environmental surveillance program for West Nile Virus (WNV).

Figure 1: Visualization of weeks in the years when mosquito growth is expected to go high. Usually, during summer through fall.

The traps that were set throughout the city had collected different species of mosquitos that spread West Nile Virus. It can be seen that from figure 1 temperature and climate strongly affect the growth of the mosquito population and hence during these periods more infections can be found. It can also be seen from figure 2 that the mosquitos with a probability of having the West Nile Virus can be mostly found during the peak times of the year when the climate and the temperature conditions are the most favorable.

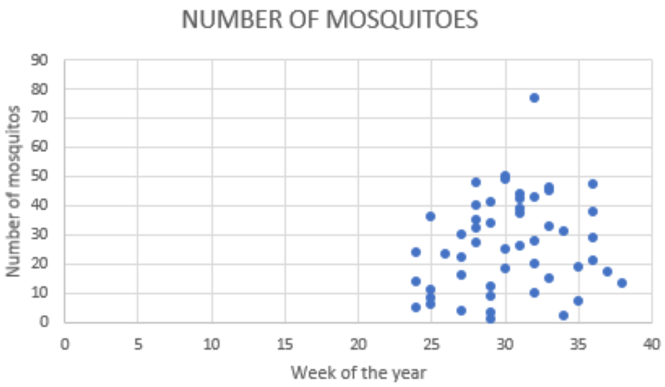


Figure 2: Distribution of detected or trapped mosquitos having the virus, during the peak periods of the year.

2. Problem Statement

Climate change is affecting the entire world. Many regions are projected to experience an increase in average temperatures and flooding, increasing mosquito populations. Seeing as mosquitos are the primary transmission vector of West Nile Virus, our team will be predicting the future trajectory of the virus in regions where it is currently endemic and in new areas. We will be using different mathematical modelling techniques to establish the relations between different entities involved in the spread of West Nile Virus.

3. Methodology

a. We will be using SEIR mathematical models to represent the future trajectory of West Nile Virus in endemic and new areas. We will be constructing this mathematical model with interventions and without intervention strategies and provide a study comparing the extent of this disease globally under different conditions. [1]
b. We plan to give a regression analysis to guide the future trajectory of the disease using python. We will train the model using the obtained datasets by using 80% of the dataset for training and then the remaining 20% for testing the models. We plan to use Logistic Regression which uses an activation function to draw the curve of future trajectory.
c. We plan to also design an agent-based model using Mesa that would capture the interactions between people, birds, and mosquitoes and show how the disease could spread across cities.

4. Preliminary Results

Mathematical Modelling

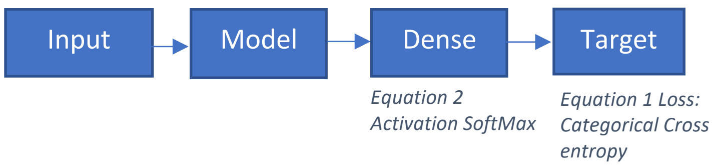
Due to West Nile Virus being a vector-borne illness, mathematical modeling of the spread of the disease can be quite complex. Such models must account for interactions between vectors, hosts, and reservoirs. Most animals, such as humans are dead-end hosts, whereas birds are the predominant reservoirs for the virus. In order to model the spread of the disease in both endemic and new emergence regions, we will use the modified SEIR models proposed in "Modelling and control of local outbreaks of West Nile virus in the United States" [3]. These models are as follows:

$$\begin{aligned}
\frac{dS_M}{dt} &= \lambda_M - a_B\beta_{BM}\frac{S_M I_B}{N_B} - \delta_M S_M \\
\frac{dE_M}{dt} &= a_B\beta_{BM}\frac{S_M I_B}{N_B} - \delta_M E_M - \gamma_M E_M \\
\frac{dI_M}{dt} &= \gamma_M E_M - \delta_M I_M \\
\frac{dS_B}{dt} &= \lambda_B - a_B\beta_{MB}\frac{S_B I_M}{N_B} - \delta_B S_B \\
\frac{dE_B}{dt} &= a_B\beta_{MB}\frac{S_B I_M}{N_B} - \delta_B E_B - \gamma_B E_B \\
\frac{dI_B}{dt} &= \gamma_B E_B - \delta_B I_B - \mu_B I_B - \nu_B I_B \\
\frac{dR_B}{dt} &= \nu_B I_B - \delta_B R_B \\
\frac{dS_H}{dt} &= \lambda_H - a_H\beta_{MH}\frac{S_H I_M}{N_H} - \delta_H S_H \\
\frac{dE_H}{dt} &= a_H\beta_{MH}\frac{S_H I_M}{N_H} - \delta_H E_H - \gamma_H E_H \\
\frac{dI_H}{dt} &= \gamma_H E_H - \delta_H I_H - \mu_H I_H - \nu_H I_H \\
\frac{dR_H}{dt} &= \nu_H I_H - \delta_H R_H.
\end{aligned}$$

Figure 3 SEIR Modelling of WNV interactions between reservoir, vector, and host

Regression Modelling

We have used regression analysis to predict whether a tested mosquito will result in positive or negative results in a given region and season conditions. The model score was 92.6 % accuracy in predicting the results using categorical cross entropy as a loss function. Categorical cross-entropy is a loss function that is used mainly in multi-class classification. Where an example can belong to one of many classes and the model decides which classification to be done.



The cross-entropy function calculates loss of any given record by computing the following formula:

$$\text{Loss} = -\sum yi \cdot \log yi$$

Where the Yi inside the log function is the target value and the Yi outside is the scalar value of the module. Categorical cross-entropy is best used in most classification models were any record is believed to belong to a specific category with a probability of 1 and belongs to another category with a probability of 0.

We will also be using the SoftMax function as an activation function after the logistic regression. As we speak of the output of the model needs always to be positive so that the *Log* of that value always exists. This is the main reason why *SoftMax* activation is used which rescales the model output so that it stays on the positive property side.

In addition to all the above methods in designing our model we also use the Adam Optimizer with a learning rate of 0.01. We choose Adam Optimizer over the classical stochastic gradient descent algorithm. We use this to update the weights based on the training data. Adam stands for adaptive moment estimation. This optimizer is not only straightforward forwards implement and has low memory requirements but also is computationally efficient.

Adam is considered to have advantages over two other extensions of the stochastic gradient descent algorithm AdaGrad and RMSProp. This is the reason even its authors recommend using it over all other optimization algorithms.

Next is the data, we have used the data from the Chicago data portal. A list of locations and test results for pools of mosquitoes are tested through the Chicago Department of Public Health Environmental Health program. Environmental surveillance for West Nile Virus (WNV) is maintained by the Chicago Department of Public Health. An important part of this program is the collection of mosquitoes from traps located around the city, their identification and sorting, as well as testing of specific mosquito species for West Nile virus.

We collected data available in CSV format which contains strings and blank data. We have done data cleaning for this which includes removing redundant data from the CSV file. This was easily achieved using Pandas Library which is used here to import the data, clean the data and use it as data frames. Using Pandas, we also Normalized the data using the below formula.

$$\mathbf{x = (x - min(x)) / (max(x) - min(x))}$$

where x is each individual attribute of the CSV data. We have about eight features that we will be needing for the result prediction study. These eight features include season year, week, trap type, result, number of mosquitoes, species, latitude, and longitude. We will be splitting the data into training and testing data with 80 % of the dataset for training and the remaining 20 % for testing. Our input features for the result prediction will be a set of seven features season year, week, trap type, number of mosquitoes, species, latitude, and longitude. Before creating the logistic regression model, we converted the trap type, result, and species to Integer and Normalized the values.

```
!python -c 'import keras; print(keras.__version__)'
```

```
2.9.0
```

For the model creation, we have used the Keras (2.9.0) Library available in Google Collaboratory. We have used the Google Collaboratory for our model training since we did not require more computation time for our training. The model was created as a Sequential model and then a Dense layer of size 1 was added to it with its input size (7,:). We add the SoftMax Activation function to this and compile the model using Adam Optimizer. This model is then subject to fit with 50 epochs. The predicted values were passed to a scoring function and around 92.6 % accuracy was obtained. The further experiment is required.

Agent-Based Modelling

In an ever-changing world, where everything is in constant motion and there are too many variables to mathematically know what precisely is going on, we then rely on numerical and computational simulations to give a concrete and visual representation of a process. As such, we will simulate the spread of the West Nile Virus using Agent-Based Modelling. However, even this task in itself is difficult, as knowing the correct context of a specific phenomenon is very tedious to know precisely. As our main task is to show temperature, humidity, and precipitation affect the spread of the West Nile Virus, we must make some simplifications:

1. The Birds have no migratory behavior and move at random.
   a. The main hosts of the West Nile Virus are Corvidae, and Corvidae is shown to have no specific migratory behavior. However, they are shown to wander around, especially when temperatures get extremely cold [4]. As such, Corvidae does slightly tend to locate in cities as the temperatures are slightly warmer due to a phenomenon is known as Urban Heat Islands [5]. We will not be considering this when it comes to modeling. (Especially since most Corvidae can already survive very harsh arctic weather.)
2. Mosquitoes slowly expand towards areas that become more habitable to them.
   a. The main effects we will consider when it comes to what is more "habitable" will only rely on three factors, being temperature, humidity, and precipitation. Features such as elevation levels (i.e., "mosquito line") are either still debated or disregarded when it comes to how it affects Mosquito Population Rates [6], and as such will not be considered. Some intervention strategies relating to the eradication of mosquitoes (i.e., a cleaner environment) will be considered later. We will also assume that a city is uniformly susceptible and vulnerable, even if mosquito population rates do vary within cities, depending on factors such as average neighborhood income levels [7].
3. All humans remain within their respective cities.
   a. One of the key features of West Nile Virus is that it does not depend specifically on humans, due to its vector-borne nature. The main circulation method of West Nile seems to follow a mosquito-bird-mosquito pattern [8], being completely independent of humans. As such, we will not be considering ideas such as gathering places (i.e., a local grocery store, school, etc.) or immigration rates (i.e., driving/flying between cities). Rather, humans will remain within their cities, and will randomly float within them.

The specific methodology we will use to measure newly susceptible areas is to initialize a space of 9 different cities. Each of them will contain a variable amount of people living within it. Since we have noted that human-to-human interaction is not a factor, we do not care about specific habits of the movements of the people, and they will be wandering around the city borders aimlessly. As for the mosquitoes, we will assume that they start within a variable region of the plane, whether it be within the lower-left most city or the right half of the map. As for the birds, they will initially be starting out in random locations and will move around randomly. In this case, a bird within the model could represent a singular bird, or it could represent a flock of birds moving together.

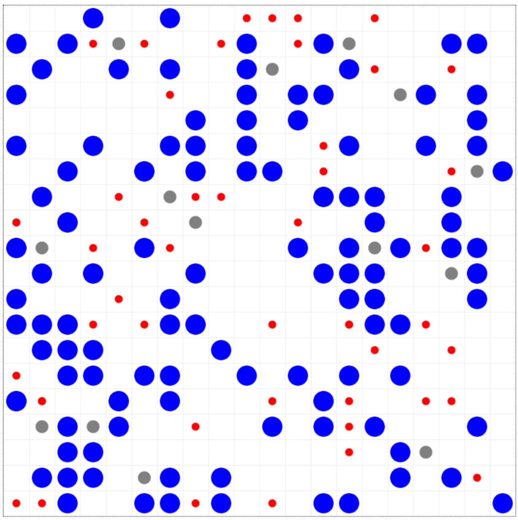


Figure 4 Representation of a Supposed City. The large, blue circles represent people, while the medium-sized, grey circles represent birds, and the small, red circles represent mosquitoes.

$$(1) \quad B(t) = .0943 + .0043t \quad \text{for } 21^\circ\text{C} < t < 32^\circ\text{C}$$
$$B(t) = 0 \text{ for } t < 21^\circ\text{C} \quad \text{for } t > 32^\circ\text{C}$$

$$(2) \quad I(t) = .0729(t - .9037) \quad \text{for } t < 26.1^\circ\text{C}$$
$$I(t) = 1 \quad \text{for } t > 26.1^\circ\text{C}$$

Figure 5 The Temperature Ranges Deciding Mosquito Habitability[9]

We will then analyse how the infection spreads. We will start with the condition that a single mosquito has the virus. If a mosquito is within a certain radius of another agent, the mosquito will "feed" upon that agent. If a healthy mosquito "feeds" into a human (whether healthy or infected), nothing will occur; however, if a healthy mosquito "feeds" into an infectious bird, it will have a probability of becoming infectious itself. Likewise, if an infectious mosquito "feeds" into an agent, whether human or bird, the bird then has a likely chance of becoming infected (unless already infected). There will be no direct interactions between humans and birds.

The agent will continue to be in the infected state until after their period has passed, and then they enter the removed category. Once the agent has entered the recover category, they will gain life-long immunity and continue on with their life.

5. Discussion

Regression Modelling

The machine learning model needs more refining on learning the positive results since the most relevant model that we could get have 90 % of data showing negative results. Although the model is less precise in predicting the outcome of the mosquitos in other words we can say the model is less sensitive and more specific in the prediction.

Mathematical Modelling

From looking at the parameters in the equation derived for the reproduction number, it becomes apparent that the most important aspects of preventing the overall spread of West Nile Virus are increasing the mortality rate of mosquitos and decreasing the birth rate of mosquitos, meaning that mosquito population control should be the main focus of preventative measures in the case of an outbreak.

6. Limitations

Data-Related Limitations in Regression

Some of the Limitations mainly related to the field of Machine Learning and Predictions of the Disease spread is having enough data to train the models to learn the behavior of the West Nile Virus. Due to the limited number of cases that WNV has throughout the world and also due to the limited number of features that each dataset is available it becomes extremely difficult to predict the trajectory. Our model is using Categorical Cross entropy for the Loss function and SoftMax activation. But still, the curve seems to be miscalculated and predictions with an accuracy of 92 % seem to be deviating from actual data when observed with test data. Further experiments need to be done on this. For this purpose, this report does not involve the future trajectory curve as stated in the methodology.

Accuracy of Mathematical Models

Some limitations of using an SEIR model for an arboviral disease such as WNV is that reservoir–host–vector interactions are extremely complex and could vary drastically depending on the region and season that they are taking place. Additionally, these types of models rely on certain per capita bite rates and other rates that typically must be estimated based on data that is gathered, and these can also vary based on the different species involved, regions, and climates. These models also do their best to account for these types of systems being continuous and not discrete.

Sensitive nature of Agent-Based Modelling

When it comes to Agent-Based Modelling, the limitations lie in its very core, in which the very simplifications it makes to model. It may be the case that the assumptions are either too reductive or may be wrong altogether. For example, if it is the case that there is the possibility, even if it is less than 1%, that the West Nile Virus could be transmitted from human to mosquito, then this may have chaotic consequences that would cause the whole model to collapse. As such, we must keep in mind whether our assumptions capture the entire essence of the problem at hand, no more and no less.

References

[1] Kushmaro A, Friedlander TA, Levins R (2015) Temperature Effects on the Basic Reproductive Number (R0) Of West Nile Virus, Based on Ecological Parameters: Endemic Vs. New Emergence Regions. J Trop Dis: S1-001. doi:10.4172/2329891X.S1-001

[2] Centers for Disease Control and Prevention. (2022, June 2). West Nile virus. Centers for Disease Control and Prevention. Retrieved September 24, 2022, from https://www.cdc.gov/westnile/index.html

[3] Modeling and control of local outbreaks of West Nile virus in the United States,Discrete and Continuous Dynamical Systems - B,21,8,2423,2449,2016-9-1,Jing Chen,Jicai Huang,John C. Beier,Robert Stephen Cantrell,Chris Cosner,Douglas O. Fuller,Guoyan Zhang,Shigui Ruan,1531-3492_2016_8_2423, sensitive analysis, transmission dynamics.,West Nile virus, mathematical modeling, basic reproduction number. doi: 10.3934/dcdsb.2016054

[4]https://www.tandfonline.com/doi/pdf/10.1080/00063657109476300

[5] https://keep.lib.asu.edu/_flysystem/fedora/2022-05/1-s2.0-s1001074208600194-main_0.pdf

[6] http://www.lomborg-errors.dk/MalariaKenya.htm

[7] https://www.caryinstitute.org/news-insights/podcast/poor-neighborhoods-and-mosquitoes

[8] https://www.cdc.gov/westnile/transmission/index.html

[9] https://curiosity.scholasticahq.com/article/17130-an-agent-based-model-of-west-nile-virus-predicting-the-impact-of-public-health-agents-and-vaccinations-on-horses