EFFECTS OF INCREASING CLIMATE VARIABILITY ON HUMAN HEALTH: A CASE STUDY OF WEST NILE VIRUS MOSQUITOES IN CENTRAL ILLINOIS

BY

PRANJALI ANIL BORSE

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Civil Engineering
in the Graduate College of the
University of Illinois Urbana-Champaign, 2023

Urbana, Illinois

Advisers:

Assistant Professor Lei Zhao
Assistant Professor Deanna Hence

ABSTRACT

Climate change poses not only direct impact on human health due to frequent events like heat waves, flooding, droughts, etc., but also has surprising indirect effects, such as the strong influence of climatic variability on mosquito-borne disease transmission. West Nile Virus (WNV) is one such prevalent mosquito borne disease in United States. Various studies have reported different effects of climate on the vectors of WNV to humans, *Culex* spp. mosquitoes; the differences in findings are often due to environmental differences between locations. The effect of temperature is well documented, and findings reported by various studies are in consensus. Precipitation parameters, however, have more complex effects on mosquito ecology. Thus, my area of interest was the combined effects of temperature and precipitation on mosquito ecology and disease transmission in the Central Illinois region.

My findings indicate that the WNV infection rate in mosquitoes was associated with hot and low moisture conditions, whereas the *Culex* spp. larval abundance was associated with high temperatures over both longer and shorter timeframes. Accumulated rainfall over long periods showed strong positive effects on larval abundance, but high intensity rainfall was associated with low larval abundance in short timeframes. Temperature and precipitation were both strong predictors of larval abundance in summer of 2003, which was the wettest and hottest period over the study period. Classification tree analysis was used to find ranges of temperature and precipitation thresholds that classify larval abundance into Low, Medium or High abundance categories. Geospatial analysis of rainfall accumulation was used to determine the meteorological processes associated with the rainfall to investigate their impact upon larval populations. These findings could be crucial for policymakers to understand the effects of current and future climate on the risk of mosquito borne disease.

ACKNOWLEDGMENTS

I would like to thank my advisors Professor Deanna Hence, Professor Brian Allan and Professor Lei Zhao for providing me with constant guidance and insights throughout the study. I would also like to thank Chris Stone from the Illinois Natural History Survey for providing the laboratory larvae data and detailed information about collection methods. This work was part of I-GUIDE and was supported by the National Science Foundation under Award No. 2118329. I would like to thank I-GUIDE community for providing the opportunity, guidance, and support. Lastly, a big thank you to my family for supporting me through every stage of my career.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

Climatic conditions have been shown to significantly influence the spread of West Nile Virus (WNV) from mosquitoes to humans, and several studies have attempted to investigate the background meteorology conducive to mosquito survival and virus transmission (Mavrakis et al., 2021, Ruiz et al., 2010, Kronenwetter-Koepel et al., 2005). However, many studies have reported that the relationship between climate and WNV is complex and is challenging to unveil because of inconsistency in findings reported. The primary vector of WNV in the eastern U.S. is an urban-dwelling mosquito, *Culex pipiens*, which lays eggs in the standing water of city drains and catch basins (Epstein, 2001). Stormwater infrastructure conditions are subject to change by rainfall events, with heavy downpours resulting in inundation of drains or catch basins, leading to a flushing effect on the aquatic habitat necessary for mosquito larval development. This flushing effect is most pronounced in urban areas, as rural areas with fewer manmade catchments and more pervious surfaces tend to absorb the rainwater and provide stable pools of water for mosquito development (Nosrat et al., 2021). However, events occurring in urban areas fill the drains with water which remains stagnant for several days or until the next heavy rainfall, attracting the mosquitoes to lay eggs. Thus, mosquitoes occur in greater abundance after heavy flooding (Nosrat et al., 2021, DeGaetano, 2005). Occurrences of heavy rainfall or continuous rainfall over a longer period of time results in water accumulation in drains and make them a favorable breeding habitat for mosquitoes (DeGaetano, 2005). Thus, comparing the relative influence of factors like inflow, retention, and flow rate of water in stormwater infrastructure is crucial to understand mosquito dynamics in water.

The average lifespan of an adult *Culex pipiens* is nearly 1 week and therefore can be impacted by weather conditions over a 1-week period (Armando, 2006). The lifespan of an adult

Culex pipiens thus determines the crucial period for studying their oviposition dynamics (Armando, 2006, Reisen et al., 1992). Temperature has been well-documented to influence mosquito lifespans and various studies have indicated a strong relationship between temperature and mosquito abundance. One study indicates that temperature during the previous month was a significant predictor of mosquito occurrence in early and peak summer (DeGaetano, 2005). Results from a study by Lebl et al., (2013) indicate that the host-searching behavior of female mosquitoes and their survival is driven by favorable temperature ranges. Importantly, previous research indicates that mosquitoes typically develop faster and survive longer in high-temperature and low rainfall conditions (Epstein, 2001), although extremely high temperatures are fatal for their survival (Drakou et al., 2020).These studies suggest that although temperature is a dominant factor in mosquito reproduction, precipitation also influences mosquito development periods, survival, and infection rate.

To investigate this role of precipitation on mosquito larval abundance, Gardner et al., (2012) used a conditional inference tree modeling approach to determine an upper rainfall threshold of maximum rainfall of 3.48 cm over a 4 day lag period, beyond which the flushing effect was pronounced (Chaves & Kitron, 2011). Their findings also established a lower threshold of 0.406 cm, below which larvae were in highest abundance. Their research focused on the larval populations emerging through the oviposition in catch basins, unlike other studies which associated mosquito abundance with weather variables without explicitly studying their place of emergence (Nosrat et al., 2021**,** DeGaetano, 2005). These findings thus identify rainfall as critical in altering the habitat conditions and flight activity of mosquitoes. Overall, rainfall was identified to be critical in altering the habitat conditions and flight activity of mosquitoes, but there are several other environmental and climatic factors that influence mosquito development periods, survival, and infection rate.

Various studies have used correlational analyses to measure the strength of the relationship between various weather variables, such as temperature, precipitation, humidity, etc., and the abundance of mosquitoes over their lifetime (Mavrakis et al., 2021**,** Ruiz et al., 2010**,** Lebl et al., 2013**).** Studies have also considered the predictive modeling of mosquito occurrence or WNV risk by adapting statistical approaches such as hierarchical clustering, linear regression, Poisson regression, classification trees, etc. depending upon the nature of data and end goals of the studies (Ruiz et al., 2010**,** Lebl et al., 2013, Gardner et al., 2012). Fewer studies have explicitly considered the climatology of the study period before feeding data into their respective model or addressed the dependency of the oviposition behavior of mosquitoes on weather conditions.

In this study, I examine the effects of weather variability on larval abundance and WNV-positive adult mosquito activity in the Champaign-Urbana (C-U) region of east-central Illinois, United States. My objective is to identify how this variability could be linked to indirectly associated factors (oviposition behavior, flight activity of mosquitoes) in order to determine when and where the risk of WNV transmission is greatest across an urban-to-rural land use gradient. In this analysis I address two significant questions: 1) What is the seasonality of suitable juvenile habitat conditions across different human land-use types that employ a mixture of subterranean and surface stormwater infrastructure, and 2) Whether the seasonality of the climatic conditions for mosquito reproduction has shifted over time, which in turn affects the timing and severity of mosquito-borne disease transmission. This study thus aims to identify the timing and bounding thresholds of mosquito-borne disease risk, with an objective of informing policies to curb transmission, given that WNV continues to pose a yearly severe threat to human health with cases over a large geographical extent in the United States.

Towards these goals, I link climatological analysis of weather conditions through the 2002 to 2021 warm seasons with batch datasets of WNV positive adult mosquitoes to examine the impact longer timescale seasonal variability on mosquito activity. To examine the influence of short-term weather variability on mosquito abundance, I pair geospatial analysis of temperature and precipitation variability with larval abundance data collected from 2003-2005 in the C-U metropolitan area. This timeframe is of particular interest, given that the years 2002 and 2003 experienced large outbreaks of WNV in the United States, with 685 cases of the novel virus reported from Chicago, IL in 2002 alone. I additionally examined the type of meteorological systems associated with significant precipitation and temperature events, since meteorological processes like mesoscale convective systems and frontal systems bring heavy rainfall to the Midwest during the warm season, but their time of occurrence and atmospheric conditions significantly differ from each other and associations with the ecology of mosquitoes have not been explored.

Chapter 2 highlights the meteorological datasets and the collection methods used for the mosquito datasets, as well as the statistical techniques used for the analysis. Chapter 3 details the results of the two exploratory phases of research of the adult and larval mosquito datasets, and Chapter 4 provides a discussion of the significance of the relationships found. Chapter 4 also provides a summary of the findings and conclusions in the end.

CHAPTER 2: DATA AND METHODS

The analysis was divided into two phases, one for WNV positive adult mosquitoes sampled from 2002-2021 and another for larval mosquitoes sampled from 2003-2005. The exploration phase I of the study analyzed climate and WNV-infected mosquito batch data obtained from the website (*Illinois Department of Public Health*, n.d.). Municipality level data records ranging from 2002-2021 were available for C-U region in the form of the number of WNV positive mosquito batches. Figure 2.1 shows the distribution of batch counts over the May-October period each year. Records contained only the dates on which positive mosquito batches were found throughout the surveillance period of June 15th to October 31st of each year. Local health departments, mosquito abatement districts, municipalities, and Illinois Natural History Survey (INHS) along with the public health department engage in regional collection and testing of mosquitoes for WNV and these agencies later input this data in the department's database.

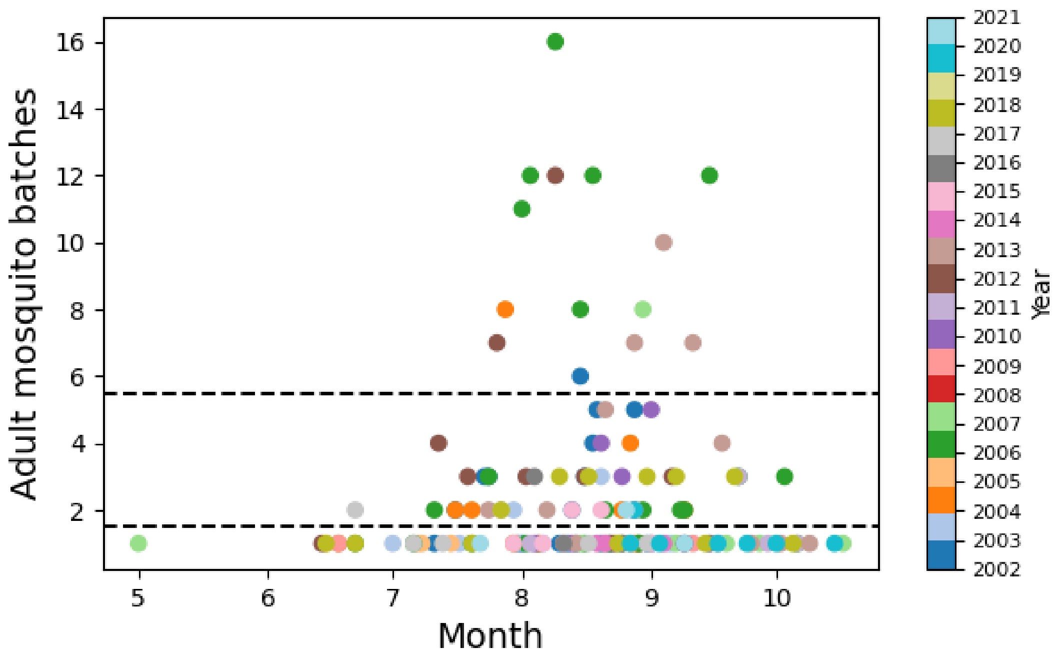


Fig. 2.1 Month wise time series of WNV positive mosquito batch counts for the period of 2002-2021. Black dotted lines represent the partition of batch counts into the three categories of batches=1, batches from 2 to 5, batches>=6.

Exploration phase II refers to the investigation of the relationship between larval abundance and climate. The daily counts of *Culex (Cx.)* species larvae were obtained from INHS and were available for the 3 years period of 2003 to 2005 for C-U region. INHS conducted larval surveillance by setting up oviposition traps at several locations in the area. These oviposition traps were made from white, 18.9-liter (5-gallon) plastic buckets, each with 6 equally spaced 7.65-cm diameter openings cut into the upper 3rd of the bucket to allow for entry of the gravid mosquitoes (Lampman & Novak, 1996b). Each trap was filled with water up to these openings and water levels were maintained over the course of the summer. To induce oviposition, about 75 g of rabbit food pellets were wrapped in cheesecloth and placed in each bucket. Rabbit chow infusion attracts both *Cx. restuans* and *Cx. pipiens* (Lampman & Novak, 1996a). At the start of each season, the rabbit chow was added for 2 to 5 days and then once every 10 to 14 days overnight, to recharge the infusions. The oviposition traps were placed in shaded areas with overhanging vegetation in residential yards, around municipal buildings, and in woodlots in both cities. The collection of egg rafts continued until either the 1st fall frost or until the majority of traps had no egg rafts for 3 consecutive days. A few trap locations changed yearly, but at least 6 sites were used each year. The number of egg rafts per trap was counted daily. For analysis purposes, the daily larvae count from all trapping sites were aggregated into total daily counts since the data were spatially inconsistent over the three years.

Meteorological data consisted of over 20 years of daily records of temperature obtained from the Illinois State Water Survey monitoring station located in Champaign (*Water and Atmospheric Resources Monitoring Program - Illinois Climate Network, Illinois State Water Survey*, n.d.) and NCEP/EMC 4KM Gridded (GRIB) Stage IV precipitation data obtained from the National Center for Atmospheric Research Earth Observation Laboratory data archive using

FTP (Du, 2011), an example of which is shown in Fig. 2.2a. Stage IV precipitation data are a combination of station observations and radar reflectivity collected from the National Weather Service WSR-88D operational radar network. This dataset was provided in a unique Hydrologic Rainfall Analysis Project (HRAP) grid coordinate system; to extract the data for desired location from the Stage IV data of the entire CONUS (Contiguous United States) region, it was necessary to know the value of grid points of C-U region polygon in HRAP coordinates. I thus used equations provided in CRWR Online Report 95-3 (Reed & Maidment, 1995) to convert latitude and longitude from degrees to HRAP coordinates. The spatial dataset of precipitation was then extracted for the C-U region using the coordinates of vertices of the C-U polygon. Unfortunately, the ovitrap locations shown in Fig. 2.2a were available only for 2003; for the time series analysis, the spatial dimension of larval data was discarded due to location inconsistencies between the three years of data. The resulting daily time series of aggregated larval counts are shown in Fig.2.2b. I thus obtained the daily time series of precipitation by spatially averaging the polygon data, which was required to match the spatial dimensions with the aggregated larval count data.

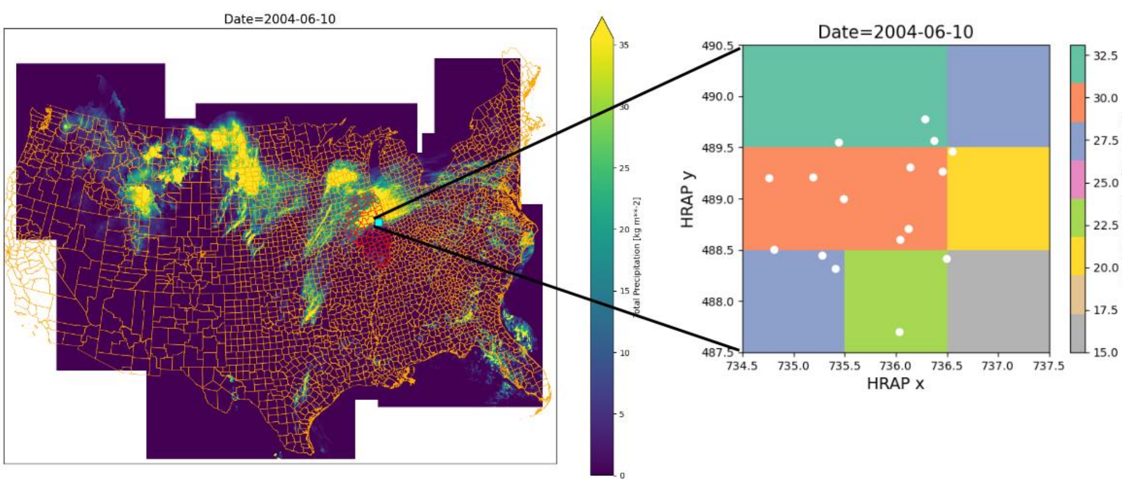
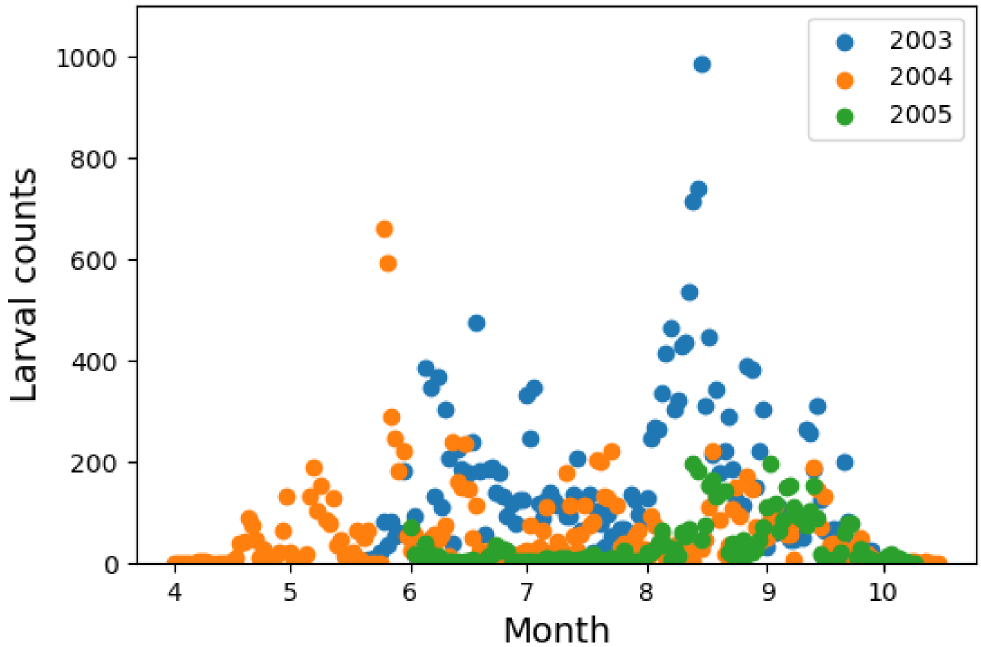


Fig. 2.2 a) An example map showing the CONUS county boundaries overlaid with one timestep of Stage IV precipitation, with Champaign County highlighted in teal. The inset highlights the 4×4 Km precipitation grid used for the C-U region and locations of ovitraps in 2003, with an example geospatial distribution of daily precipitation accumulation collected that year. **b)** Daily time series of total larval counts for the total collection months of study period (2003-2005)

Fig.2.2 continued



2.1 Exploration phase I

This phase of the analysis involved exploring the relationships between WNV infected mosquito abundance and two meteorological variables - precipitation and temperature. The mosquito data were a record of positive mosquito batch counts and not individual mosquito counts. Visualizing this data revealed the possible grouping of these batches according to their frequency, i.e., the frequency of finding a single batch a day was highest followed by that of 2-5 batches a day, with the lowest frequency being finding greater than 5 batches in one day. This classification helped in analyzing each group separately with the meteorological variables to identify the varied effects of weather on different groups of batches. These relationships are discussed in detail in the results section.

As the data being analyzed were for WNV infected mosquitoes, mosquitoes were presumed to be in the adult stage that would have already obtained a blood meal from an infected host. Thus, accounting for the development period of mosquito, infection, and virus incubation period, I considered a lag period of 3 weeks to analyze the weather dynamics that would have influenced the prevalence of WNV positive mosquito batches. After examining correlations with several meteorological variables, the resulting analysis focused on accumulated precipitation and mean temperature over continuous lag periods ranging from 1-21 days. I plotted the frequency distributions of the accumulated precipitation and mean temperature corresponding to the three batch groups. This analysis enabled me to identify the ideal ranges of these weather variables and their variation in the development period of mosquitoes. The analysis was extended to considering different lag windows to identify the most influential lag window as well as the weather variable magnitudes associated with it, but the results were not consistent over the study period which limited our ability to draw strong conclusions.

The observations from frequency distributions were a motivating factor to investigate the relationship between the length of the immediately preceding dry period and the mosquito batch frequency. Pertaining to this, I identified the duration of continuous dry period before every batch occurrence and the frequency of each duration before every batch occurrence over the years of data.

2.2 Exploration phase II

As is evident from the literature, it is challenging to identify the lag window of weather conditions that are most influential for mosquitoes. The above correlation analysis on mosquito abundance and climate variables strongly suggested relationships to both precipitation and temperature, yet the likely influence of other both ecological and environmental factors over such long time periods prevents drawing more specific conclusions. The daily record of larval counts and climate variables opened options to expand and perform more detailed analysis.

It was important to characterize the study years based on background meteorology before initiating further analysis as it was the basis of the interpretations for the results. The analysis of climatology and estimation of climate anomalies, i.e., monthly accumulated precipitation and monthly mean temperature anomalies as in Fig.2.3 allowed me to do the intra- and inter-annual comparisons of climate variability. This initial climate analysis set enabled me to conduct the analysis to identify the relationships between larval abundance and the weather variables in different lag windows.

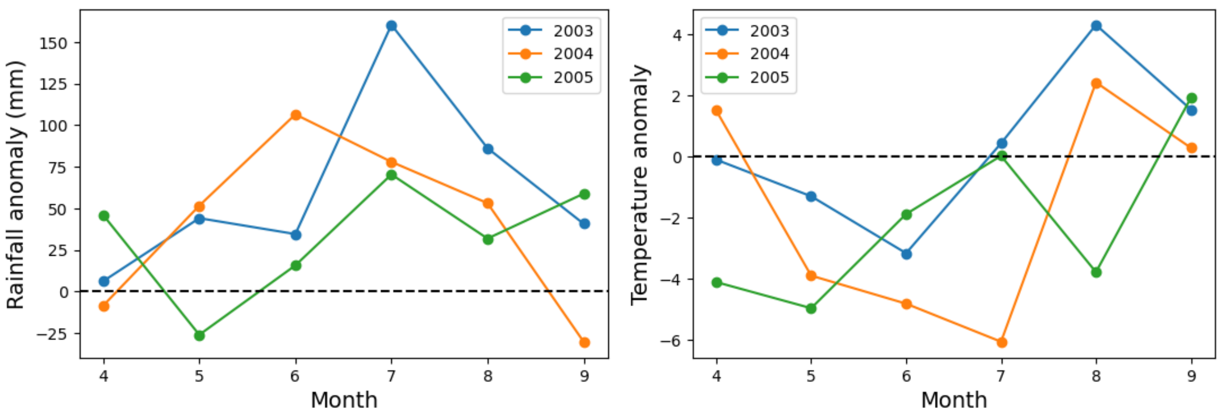


Fig 2.3 Observed monthly anomalies of precipitation for April-September calculated using 20-year monthly **a)** rainfall anomaly and **b)** temperature anomaly. Black dotted line is the zero line to differentiate positive and negative anomalies.

I computed the spearman's correlation coefficients to quantify how strongly the larval abundance was related to the weather in different lag windows ranging from 1-84 days. The upper limit of the lag window was determined by computing the spearman's coefficient until the highest correlation was reached. This correlation analysis helped in identifying the influential lag windows in each year during early and peak summer for both the climate variables. I hypothesized that there exists a common influential lag period among all years in which climate is strongly related to the larval abundance, and the intent of correlation analysis was to identify this lag period.

2.2.1. Predictive modelling and probability approach

Based on the outcomes of this test, I attempted to use a machine learning model to translate the relationships between climate and larval abundance into a predictive model. I used a random forest regression model for our preliminary analysis with modeling, selected based on the nature of relationship between variables. Input features were climate anomalies in influential lag periods identified from correlation analysis. Data were split into training and testing datasets to evaluate the performance of the model.

The lag analysis mentioned above captured the differences between the relationships during immediate and longer lag periods, suggesting the need of separating out the analysis into longer duration and immediate effects analyses. Accounting this, phase II was divided into long-term and short-term analysis. Visual inspection of a dual Y-axis plot of time series of larval abundance and immediately lagging rainfall hinted to the relationship between simultaneously occurring rainfall spikes and troughs in the larval population.

To statistically quantify this relationship, after reviewing literature, I settled on using the Bayesian probability approach. This approach helped me in determining the probability of having local peaks in larval counts given the rainfall exceeds a certain value. A similar methodology wasn’t suitable to determine temperature thresholds as temperature has much slower effects over a longer period, unlike the episodic nature of rainfall which could alter the habitat conditions immediately and so the abundance of larvae. The conditional probability of finding local peaks in larval population given the rainfall exceeds certain threshold was calculated as

$$\mathrm{P(peaks|\ R > i)} = \frac{\mathrm{P(R{>}i\ |\ peaks){*}P(peaks)}}{\mathrm{P(R{>}i)}} \tag{1}$$

where, $P(peaks)$ is probability of finding peaks in the larval population, $P(R > i)$ is the probability of rainfall exceeding i (in mm) before the peak occurrence and $P(R > i \mid peaks)$ is the probability of rainfall exceeding i that is followed by peak abundance in larvae. These separate probability components could be approximated as

$$\mathrm{P(peaks)} = \frac{\mathrm{N(peaks)}}{\mathrm{N(Total\ observation\ days)}} \tag{2}$$

$$\mathrm{P(R > i)} = \frac{\mathrm{N(R{>}i)}}{\mathrm{N(Total\ rainy\ days)}} \tag{3}$$

$$\mathrm{P(R > i \mid peaks)} = \frac{\mathrm{N(peaks,R{>}i)}}{\mathrm{N(peaks)}} \tag{4}$$

Here,

$N(peaks)$ is the number of occurrence of peaks in time series of larval count data,

$N(Total\ observation\ days)$ is the total number of days when larvae samples were collected,

$N(Total\ rainy\ days)$ is the number of total rainfall events in the given time reference,

$N(R > i)$ is the number of days when rainfall exceeds *i* mm,

$N(peaks, R > i)$ is the number of days peaks occurred in larvae when rainfall exceeded *i* mm

The limitation of this approach is that it yields consistent results for all three years only if local maxima in the larval counts time series are considered as peaks. However, it gives an indication that rainfall immediately preceding the larval occurrence i.e., rainfall with time lag of one day is crucial in determining the larval abundance.

2.2.2 Classification approach: Tree modeling

A study by Gardner et al., (2012), which focused specifically on *Culex* species from stormwater catch basins to investigate its relationship with weather variability, was also aimed at finding weather-related thresholds. Its approach of conditional inference tree modeling motivated me to set up the model in a similar fashion; however, due to poor performance of the Gardner et al., (2012) categorization system during testing of model, I instead ran the model directly using the larval count data. Since the aim was to focus on short term weather conditions, the lag period of weather was limited to 7 days. Features like maximum daily temperature for 7 days, maximum temperature over the range of 2 to 7 lag days period, and mean temperature over the range of 2 to 7 lag days period, were considered. Rainfall features were established in similar ways except accounting for an accumulated rainfall over the range of 2 to 7 lag days instead of the mean. Including a feature set of rainfall and temperature variables together in the conditional tree model resulted in a tree split based solely on temperature; since rainfall did not play a role in splitting of the tree, this initial result suggested that rainfall played a much smaller role compared to temperature and thus, was not a part of the tree. As the major focus of this study is investigating relationships with rainfall, the conditional tree approach did not seem to be the right fit.

As there exist several other classification techniques, I considered it important to redo the analysis using other classification techniques to find out if there was consistency between the outcomes from two different methods and if rainfall variables influenced the classification. Decision trees and random forests are two such widely used methods for classification problems. Random forest usually works well when the sample size of the data is large as it yields an average of outcomes from decision trees trained on the numerous ensembles of the data. In this study, the timeframe wasn’t enough to pursue this approach, so a single decision tree modeling approach was chosen with an aim to find weather thresholds.

Decision trees are known to have worked well with categorical dependent variables. Also, in some cases like this one, categorization serves the purpose of determining relative abundance and determining the precise numbers is not necessary. The next step was to convert the larval counts into categories. Evident from the histogram in Fig. 2.4a, the larval abundance data were an imbalanced dataset, with 50 percentiles of the values covering only the range of 0 to 67 larvae while the maximum count value was 986. In such situations, it is necessary to choose the categorization that results in a comparable sample size of all the categories to avoid falsely inflated performance estimates of the model. Trying different categorizations based on percentiles, absolute counts, re-binning of histogram and ranking approach, concluded that re-binning of the histogram was the most suitable for categorization amongst all, which resulted in a relatively balanced classification into 'Low', 'Medium' and 'High' categories. Counts with seemingly lower values belonged to the 'Low' category, which ensured true assignment of category to abundance levels. According to this categorization, counts less than 57 percentiles were defined as 'Low (L)', between 57 to 90 as 'Medium (M)' and greater than 90 as 'High (H)' category. The sample sizes were 209, 119, and 38, respectively.

As the original data were highly imbalanced, it was not possible to obtain ideal balanced categorization. In such cases, there are techniques to obtain the implicit balance while modeling using the scikit learn (sklearn) package in Python (Pedregosa et al., 2011). The most suitable technique for the larval dataset was the class weighting approach which enables me to bias the model to compensate for those categories that are less well represented in the data. Thus, I used an additional hyperparameter, 'class_weight= balanced' in our model simulations. This parameter creates an implicit balance in the data by automatically adjusting weights inversely proportional to class frequencies in the input data. Thus, including this parameter increased the purity of the tree nodes by representing well the minority class which was the 'H' abundance category in this case.

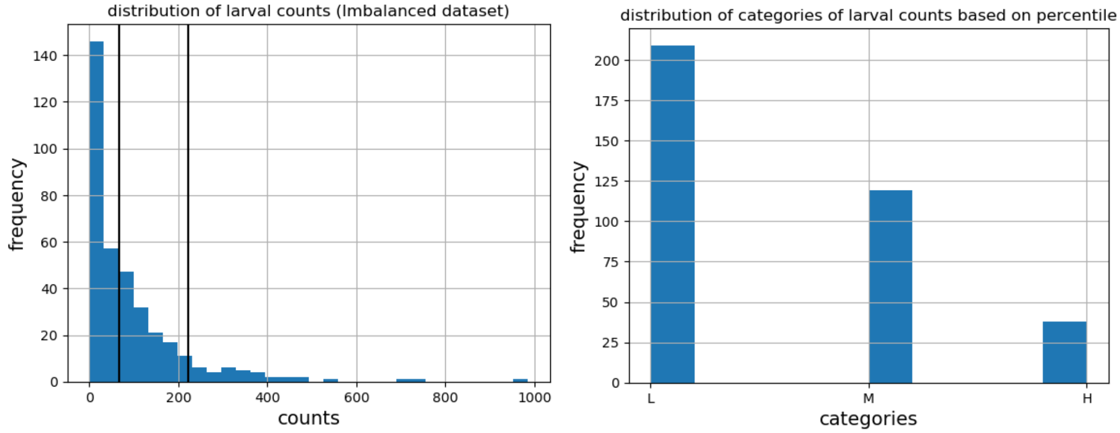


Fig 2.4 a) Histogram showing distribution of larval counts and vertical black lines representing separation of counts in categories. Vertical black lines indicate the partition of counts into larval abundance categories **b)** Frequency distribution of larval counts into Low, Medium and High frequency categories.

Although the balance was obtained, it was observed in the initial stages of analysis that the model was overfitting. Decision trees are prone to overfitting and need to be pruned to eliminate it. Therefore, the next step was tuning of the hyperparameters to obtain the set of hyperparameters resulting in balanced accuracy between training and testing dataset. I used a pruning algorithm which passed a range of learning rates in the model and obtained training and testing accuracy for each of the learning rates. To do this, the data was split into 80% training and 20% testing using the 'train_test_split' package of the sklearn Python package (Pedregosa et al., 2011). Visualizing these training and testing accuracies plotted against the learning rate allowed me to identify the learning rate which minimized the difference between training and testing accuracy that protected the trees from overfitting. I was able to obtain the learning rates for different combinations of temperature and precipitation variables which resulted in greater than 60% accuracy for both training and testing dataset with minimum difference between the accuracies. A 60% benchmark was chosen to filter out the models resulting in good to best performance.

The class weighted decision tree modeling also confirmed that rainfall variables, even though significant, do not play a dominant role in classification of larval counts. It is likely due to the episodic nature of rainfall consisting of outliers with extremely heavy rainfall on some days, unlike the much more gradual variation of temperature over longer periods. With over hundreds of simulations using different combinations of temperature and rainfall as features, the resulting decision tree always yielded temperature as the deciding factor for first splitting. This exercise also allowed me to identify that oviposition trap conditions do not replicate stormwater infrastructure conditions as the results from my analysis indicated that rainfall was not the important parameter in splitting the tree. Thus, they were different than what was reported for the study using stormwater infrastructure conditions (Gardner et al., 2012). Thus, it cannot be used as a direct proxy to study effects of catch basin conditions on mosquitoes.

Categorization of larval data turned out to be an effective approach as it let me visualize and understand the distributed effects of weather conditions on different larval categories. Even though the decision tree classifier could not yield a strongly supported rainfall threshold, categorization of larval counts helped in analyzing rainfall variables separately. The boxplots of rainfall distribution associated with larval categories (Fig. 3.9 a,b) showed a distinct difference between rainfall ranges for different larval categories.

2.2.3. Geospatial analysis of precipitation distribution

Since the rainfall data were spatially averaged, it was important to go back and check on spatial variability of rainfall to identify if rainfall was uniformly distributed over the region during heavy rainfall events or if it was concentrated in certain areas. Spatial variability analysis was carried out for the average rainfall values greater than 20 mm day^{-1}. I selected 20 mm day^{-1} based on the frequency distribution of rainfall over the period of three years; the frequency of events with rainfall greater than 20 mm day^{-1} is much lesser compared to the frequency of smaller magnitude events. In addition, the types of warm season meteorological systems that deliver high rainfall rates frequently have highly localized rainfall with tight precipitation gradients. So, it was important to investigate both why such events are rare and how the rainfall is spatially distributed. Visualizing the frequency distribution with boxplots (Fig.3.9a) allowed me to identify the events with high and low spatial variability. This analysis segregated events with high spatial variability and determined ranges of larval counts observed after high and low spatial variability events. The difference in the ranges is discussed in the results section.

CHAPTER 3: RESULTS

3.1 Lag analysis of climate with WNV-positive mosquitoes

Mosquito batch data were analyzed with the accumulated precipitation and mean temperature over lag periods ranging from 1-21 days from the date of occurrence of mosquito batches (Section 2.1). The analysis results indicate that high temperatures in the range of 21-32◦C (Fig.3.1a) were optimal for the adult infected mosquito occurrence as evident from Fig.3.1a where highest frequency of batches is associated with high temperature range. Precipitation, however, did not have a defined range (Fig.3.1b). The occurrence of a single batch occurred more frequently than the other two batch size categories and the analysis revealed that the single batches spanned a wider temperature range compared to the other two. Batch counts of 6 or greater were the least frequent in occurrence and were confined by a narrow moderate to high temperature range of 20-32◦C. However, peaks in the distributions in Fig 3.1b generally shows that the mosquito batches were most frequent in the low rainfall ranges below 40 mm of accumulated rainfall. Interestingly, the distribution of the batch counts of 6 or greater was significantly flatter through that range.

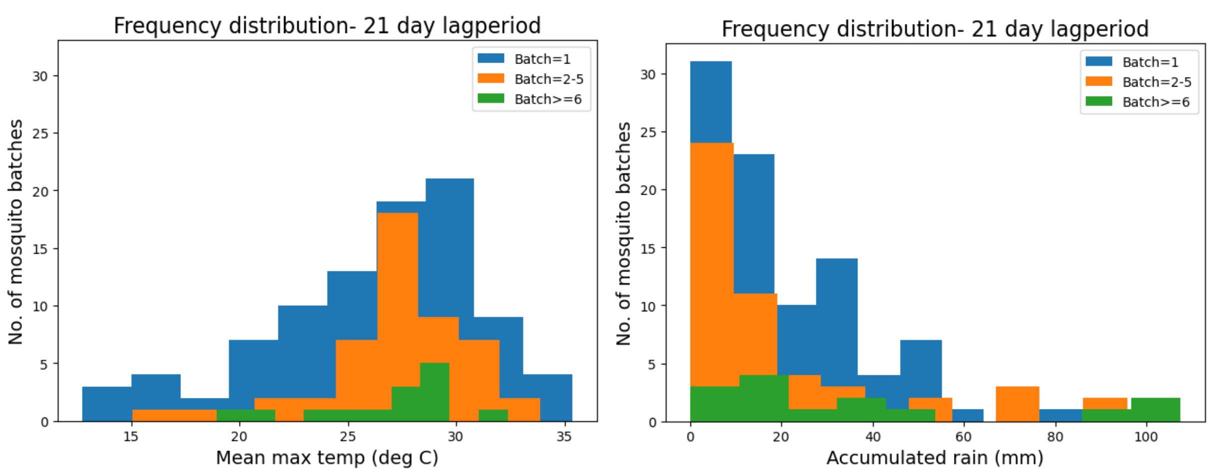


Fig. 3.1 a) Frequency distribution of WNV positive mosquito batches according to mean max. temperature over preceding 21 days **b)** Same as a) but according to Accumulated rainfall.

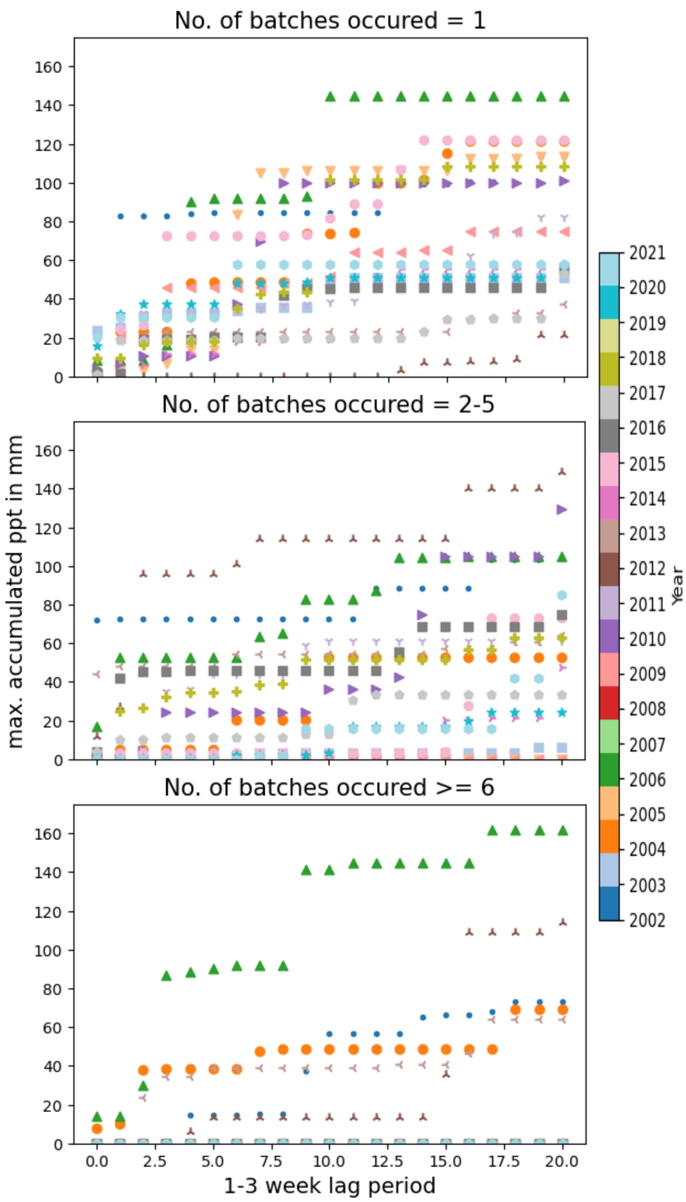


Fig. 3.2 a) Time Series of accumulation of rainfall over 1 to 21 days lag period for occurrence of single batch **b)** 2-5 batches **c)** more than 5 batches

Batches were mainly found to occur following 2- 3 days of a dry period and several dry spells in the 3-week period. Figure 3.2 shows the progression of maximum accumulated rainfall amount that was associated with the three categories of batch number. The plateaus in the intervals define the dry spells and thus it suggests that mosquitoes only need small amount of rainfall to develop and survive. Dry spells could be an indication of an increased level of water searching activity by female mosquitoes to lay eggs and possibly coming in proximity to the infected birds.

While more batches tend to occur following a dry period, a continuous, long dry period is detrimental for the mosquitoes, evidenced by the decrease in batches as the continuous lagged dry period increases in Fig.3.3. These results suggest the need of rainfall in intermediate periods.

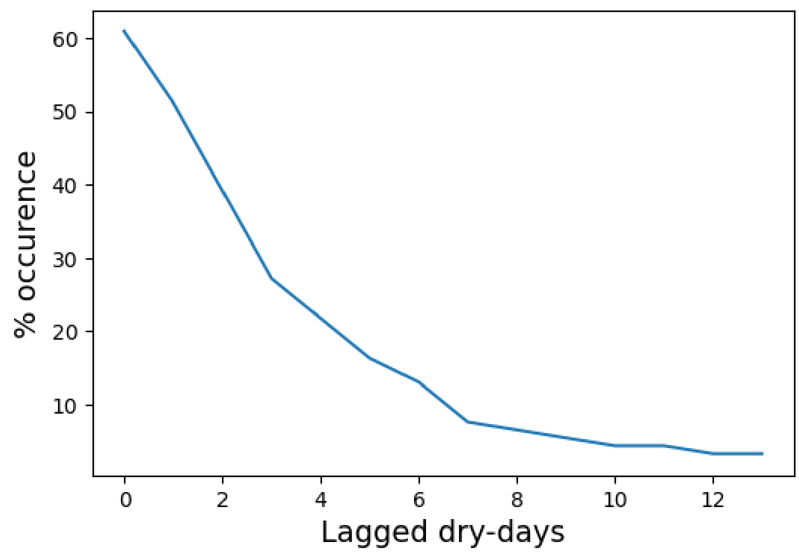


Fig. 3.3 Variation in percent occurrence of mosquito batches with increasing lagged dry period

These were the notable and consistent results from the analysis for WNV positive mosquitoes. Overall, the data of WNV positive mosquito batches also demonstrated the seasonality of WNV as batches were at peak occurrence in mid-August almost every year. This peak period was preceded by the hottest period of July every year which further supports the lag relationship between temperature and mosquitoes. The main findings suggest that prevalence of hot and semi-dry weather in summer could increase the risk of WNV spreading to mosquitoes and in turn to humans. Thus, such weather is not only detrimental from the context of heat stress to human health but is also associated with the spread of WNV.

3.2 Lag analysis of climate with mosquito larvae

3.2.1 Longer lag analysis

Exploratory correlation analysis suggested a strong nonlinear relationship between larval counts and climate variables at certain lag periods, in that larval abundance possess moderate to strong positive and negative correlation with climate variables at different time lags. The results from separate analysis with precipitation and temperature are discussed in the following subsections:

3.2.1.1 Relationship 1: Larvae & Accumulated Precipitation

Spearman's coefficients were determined for any association between larvae and accumulated precipitation over different time lags over the warm season. Fig.3.4a shows that larvae occurring in June-July period possess moderate negative relationship with accumulated precipitation in the time lag of 3-5 weeks in 2003 and in the time lag of 3-4 weeks in 2005, whereas the relationship is weakly positive for almost at all time lags in 2004. Larvae occurring in August-September period possess a moderate to strong positive relationship with accumulated precipitation in the time lag of 5-9 weeks in 2003 and a moderate positive relationship in time lag of 8-9 weeks in 2005. For 2004, however, there is a moderate negative relationship between larvae and accumulated precipitation in a time lag of 4-5 weeks.

3.2.1.2 Relationship 2: Larvae & Mean Temperature

Similarly, Spearman's coefficients were determined for any association between larvae and mean temperature over different time lags for two sub seasons - Early and Peak summer. Figure 3.4b shows the variation in the relationship between larvae and mean temperature over the period of 1-11 weeks. Larvae occurring in June-July have a moderate positive relationship with mean temperature over the first 3 weeks in 2003 and a moderate negative relationship in the same period in 2004. The relationship was weak and negative throughout 11 weeks in 2005. In the August-September period, the relationships were positive throughout the 11-week period but only 2003 was found to show a moderate to strong positive relationship in 1-6 weeks period and 8–12-week period and the year 2004 showed a moderate positive relationship in 2–7-week period.

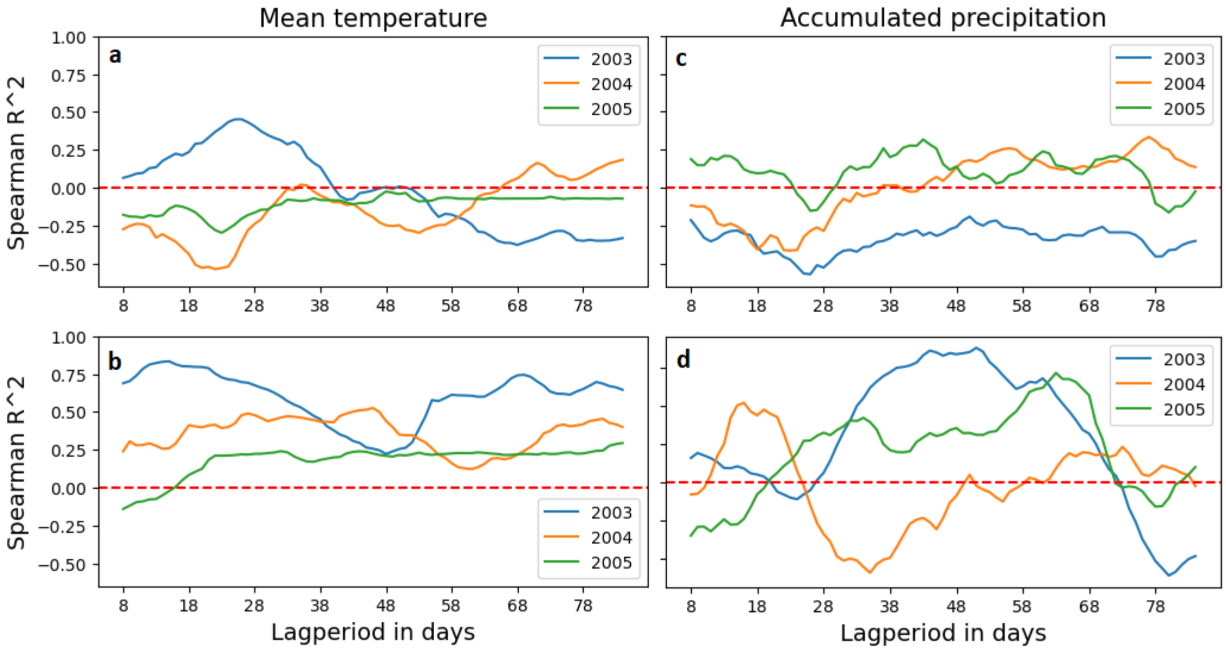


Fig. 3.4 a) Correlation between larval counts and mean temperature for Jun-Jul timeframe for 2003 (blue), 2004 (yellow) and 2005 (green) over 8 to 84 days lag period. **b)** same as a), but for Aug-Sept timeframe **c)** same as a), but accumulated rainfall **d)** same as c) but for Aug-Sept timeframe

The purpose of this correlation analysis was to identify the most influential lag period among all three years where the climate variables were crucial in altering mosquito activity. As described in methodology, these climate variables in crucial lag periods were then fed to the model as features in an attempt to create a predictive model that could inform disease risk based on the long-term climate.

3.2.1.3 Predictive modeling

Initially, Random Forest model was chosen to do the predictive analysis and the separate models were fitted for June-July and August-September period. These models yielded the best results when the climate features in a lag period of 1-5 weeks were used for the early summer period and 3-8 weeks were chosen for the peak summer period. However, the model did not perform well on unseen datasets as the influential lag period differed every year, so it was not feasible to generalize results of this model due to the small timeframe of the study.

3.2.2 Shorter lag analysis

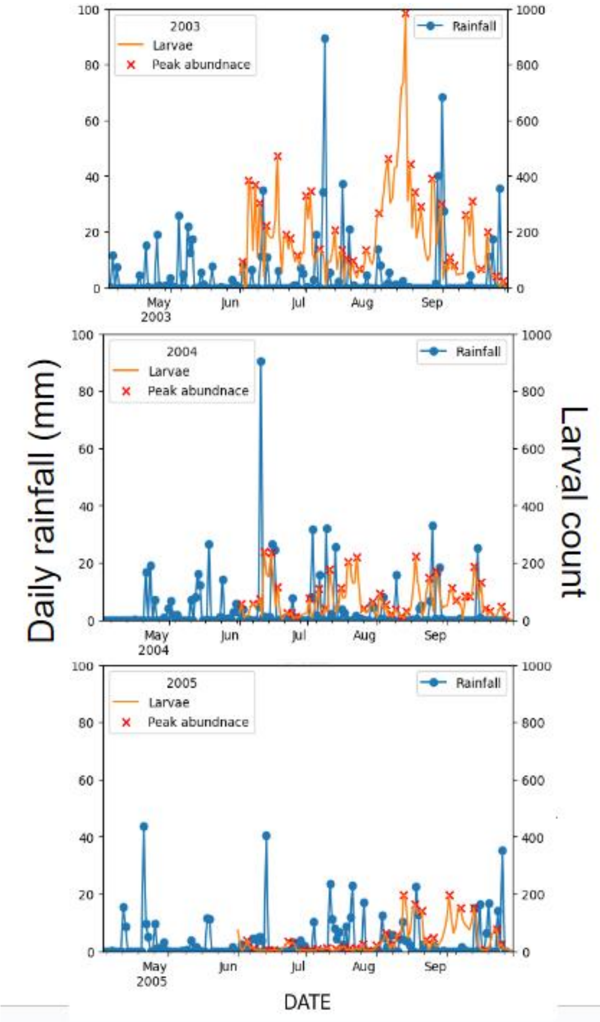


Fig 3.5 a) Time Series of rainfall and larvae catch data with red crosses representing local peaks in larval data for 2003 **b)** 2004 **c)** 2005 resp.

The first lag week was excluded from the long-term analysis as climate variables in that period showed non-significant correlations with larval abundance, whereas the short-term analysis considered only the first lag week. The purpose was to investigate the immediate impacts of weather conditions. Figure 3.5 shows the simultaneously plotted time series of larvae and rainfall for the years 2003, 2004 and 2005, respectively. These figures helped in visualizing the negative impact of rainfall spikes on the larval population since the spikes in rainfall were always followed by large depressions in the larval population.

The Bayesian probability approach was helpful in quantifying this relationship by determining the decreasing probability of peak larval abundance with increasing rainfall when local maxima were considered as peaks (Fig.3.5, 3.6). Figure 3.6 shows the consistent pattern of drop in probability of peak larval occurrence as the rainfall amount increases. But the value of rainfall when a sudden drop in probability is observed is different for all three years. Also, the approach did not yield similar results when peaks were defined as values above 90th percentile, which excluded local maxima with low larval counts. It was because of the very small sample size of peak abundance instances using the latter definition that resulted in randomized outputs.

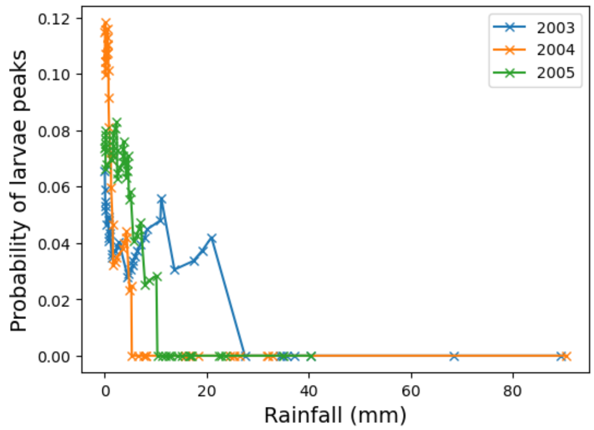


Fig 3.6 Variation in probability of peaks in larvae with increasing 1-day lagged rainfall amount for years 2003 (blue), 2004 (orange), and 2005 (green)

Such approaches brought me closer to quantifying the relationship between rainfall and larval abundance but were not strong enough to make a conclusion. As the probability approach yielded different results for different ways of defining peaks in larval counts, it was a motivation to pursue the classification approach to identify the weather characteristics associated with lower and higher abundances of larvae.

3.2.2.1 Weather thresholds estimation

The most significant finding from numerous model simulations in R and Python was that rainfall was not a significant predictor of larval abundance categories compared to temperature, as seen in Fig.3.7. Conditional inference tree modeling in R indicated that i) when any of the rainfall variables was included in the model with the temperature variable, temperature variable tended to be the dominant feature i.e., the tree was split according to temperature with no appearance of rainfall variable in the resulting tree (Fig.3.7a). The split was made at a high temperature of 34°C (93.3 F) and it wasn’t informative of the larval abundance in moderate temperature conditions.

ii) Only rainfall with a time lag of 1 day was significant when compared with other rainfall variables (Fig.3.7b) as they were never a deciding factor for splitting the tree.

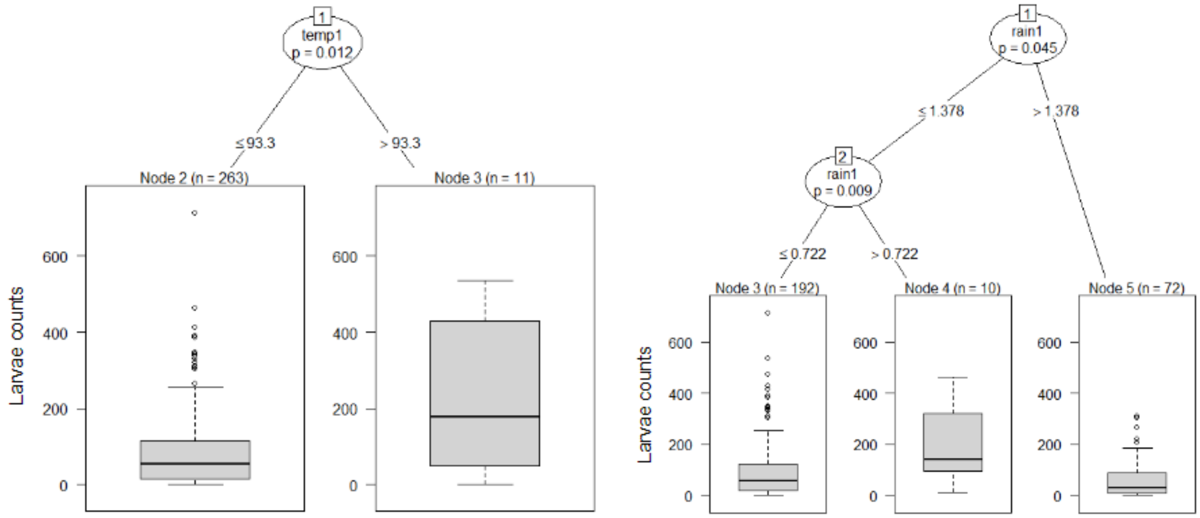


Fig. 3.7 a) Conditional inference tree model involving one day lagged temperature and rainfall as features, **b)** same as a) but only rainfall as a feature. The boxplots represent distribution of larval counts belonging to the tree nodes.

Further analysis using Decision tree modeling yielded additional information in comparison to the previous model. Initial simulations of the Decision tree were performed using the pairwise combinations of all rainfall and temperature variables as features which resulted in 19x19 i.e., 361 combinations. Amongst these, model accuracy was greater than 0.6 only for certain combinations which mainly included 'maximum' category of both rainfall and temperature variables. To statistically verify this conclusion, I performed a 'mutual information' feature selection test on the set of 38 rainfall and temperature variables. The results of this test were in consensus with the conclusion from accuracy scores. The first 25th percentile of the important features consisted of only the 'maximum' category variables. This helped me narrow down the feature space and number of simulations.

The most notable factor of the resulting decision trees was that the first split was governed by a temperature variable in 100% of simulations. It confirmed that the first splits of the conditional inference tree method and decision tree method were in consensus. Based on these results, I can conclude that rainfall is not as important as temperature in determining the abundances of larvae. However, unlike conditional inference tree models, decision trees considered further splits which involved rainfall variables. I attempted to optimize the model performance by using methods like pruning, cross validation, and nested cross validation recommended for small timeframe datasets, and settled on a model yielding the best results. Based on these techniques, a feature combination of 'maxrain6' and 'maxtemp7' yielded best results with the rainfall variable participating in secondary splitting. I achieved the outer and inner cross validation accuracy score of 0.61 by determining the learning rate obtained by pruning. The certainty of this accuracy score was also verified using the confusion matrix to confirm that the model is not biased towards the majority. Figure 3.8a, b shows the confusion matrices for training dataset and testing dataset respectively. It can be noticed that false predictions in minority class ('H') are very low meaning among five data points of 'H' category, only one is predicted falsely as of 'M' category. There was still lack of precision in predicting the 'M' category correctly. Thus, it's the category where other factors than climate could dominantly be playing role in influencing the mosquito activity.,

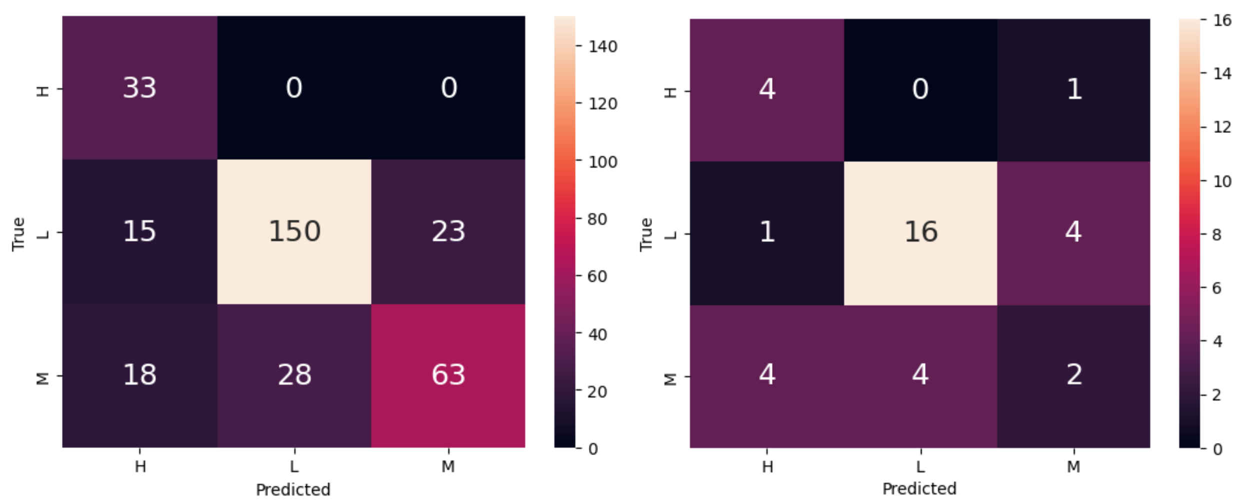


Fig. 3.8 a) Confusion matrix for results obtained by running model on training dataset **b)** same as a) but on testing dataset.

This model yielded the rainfall threshold that could have played a role in influencing the abundance of larvae. However, the first split was made at maximum temperature over a 7-day period (maxtemp7) <=35.9^{0}C. The second split was based on the maximum rainfall over 6 days <= 78.8 mm. Identifying the individual thresholds of temperature and rainfall was thus possible due to the modeling approach. However, the question about their interplay arose from their distribution associated with different categories of larvae as depicted by boxplots (Fig.3.9). The distributions indicate that while the high abundance category belongs strictly to the moderate to high temperature ranges, the low abundance category boxplot spanned from low to high temperature range. Such overlap in temperature ranges suggested that there could be other factors driving the mosquito activity. Also, the medium abundance category did not show a clear range of weather variables that differentiated it from high and low abundance categories. So, the boxplots in Fig.3.9 reflect the cause of model not accurately predicting the 'M' category which is evident from Fig.3.8a and b. It is, thus, crucial to determine the role of other biotic and abiotic factors to understand the larval abundance belonging to medium category.

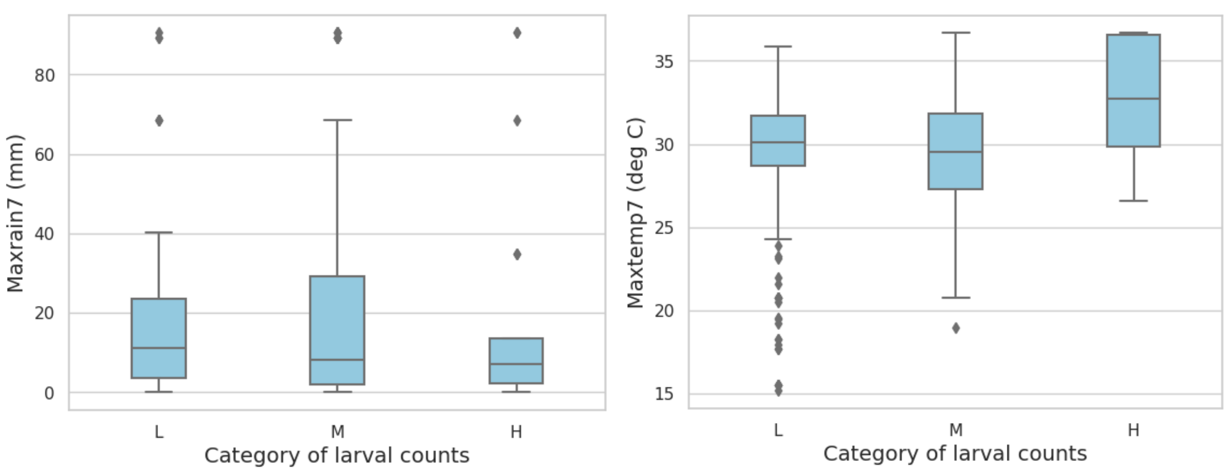


Fig. 3.9 a) Distribution of maximum rainfall over 7 days associated with the larval count categories L(Low), M(Medium), H(High), **b)** same as a) but maximum temperature.

3.2.2.2 Geospatial analysis of precipitation distribution

Overall, modeling covered the quantitative aspects of research but qualitatively, it was crucial to investigate the meteorology behind the resulting magnitudes of weather variables to make conclusions meaningful. Determining spatial distribution of rainfall was helpful in identifying the meteorological processes that would likely have driven certain precipitation events. High rainfall had been identified to be detrimental for mosquitoes in an earlier analysis, but there were some outliers e.g., high counts were observed even after high amounts of rainfall. So, checking the spatial distribution was necessary to find out possible reasoning behind these outliers and to attribute such heavy rainfall events to the meteorological processes common to the warm season in this region, like mesoscale convective systems (MCSs) in the mid-summer or frontal systems in the late spring and early fall.

Figure 3.10 shows the spatial distribution of rainfall events and how the distribution is related to the standard deviation (Fig 3.10a) as well as spatially averaged rainfall (Fig 3.10b). These figures indicated that very high rainfall events were also associated with high variability in both spatial distribution and intensity, although some moderate rainfall events showed high variability as well. These events more encapsulate precipitation commonly from convective storms, which would have a mixture of pocketed high-intensity rain regions surrounded by steep gradients to low or no precipitation. Low or moderate intensity, low variability events would be highly uniform light to moderate precipitation characteristic of widespread stratiform.

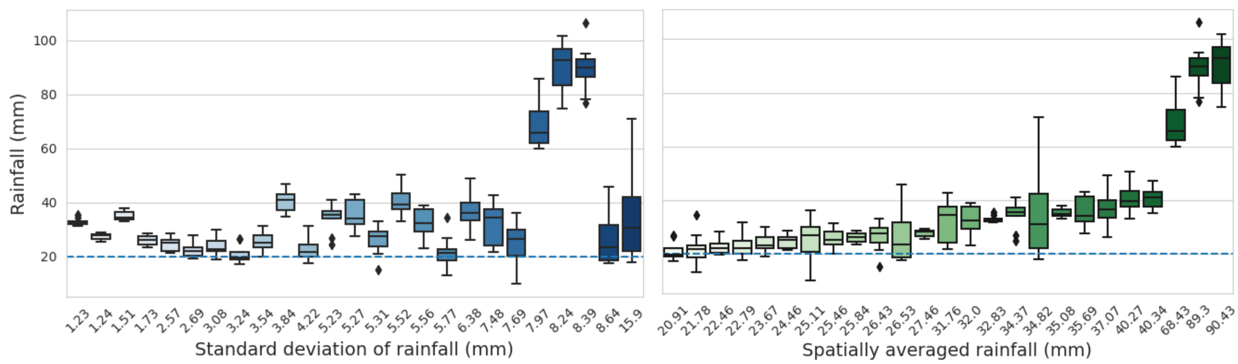


Fig.3.10 a) Spatial distribution of rainfall events with spatial average greater than 20 mm plotted against standard deviation of the rainfall events, **b)** same as a) but against spatially averaged values of rainfall events.

Strong convective systems such as MCSs bring heavy downpours with windy conditions, these facts could temporarily inhibit mosquito activity. Thus, the spatial analysis was aimed at identifying whether areas that experienced exceedingly high rainfall accumulations in this class of storms correlated with depressed larval counts. The mesoscale characteristics of the systems that these rainfall events belonged to were verified from the spatial extent of the data as well as NOAA's hourly precipitation radar data. The spatial analysis indicated that three very heavy rainfall events with high spatial variability result in 100th percentile of larval counts belonging to low abundance category the next day. These extreme rainfall events specifically belonged to the mesoscale complexes, providing support for my hypothesis. Fig. 3.11 presents one such rainfall event with its spatial distribution which was identified as having the scale of an MCS.

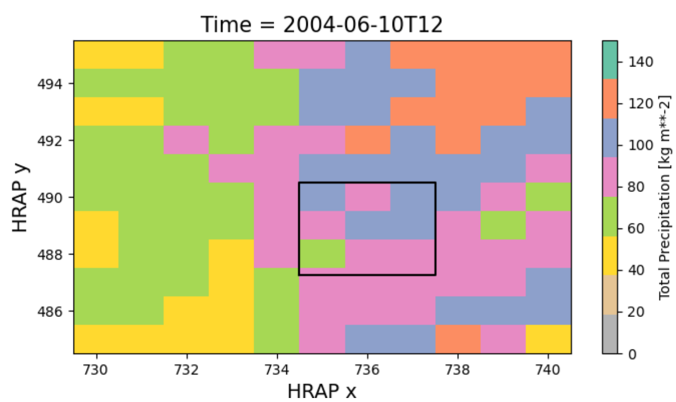


Fig. 3.11 Spatial distribution of an extreme rainfall event characterized as being mesoscale in nature. The black box represents the study region of C-U.

Figure 3.12 depicts the ranges of larval counts associated with different categories of rainfall determined using Fig. 3.10. The very heavy and high variability rainfall events are the three events distinctively seen in Fig. 3.10b upper right corner, heavy high variability rainfall events are the ones with spatial average rainfall greater than 35.8 mm which is mean spatial average rainfall and variability greater than 5.2 mm which is the mean variability. Figure 3.12 shows that moderate rainfall events with low variability category spans the much wider range of larval counts, while the high variability events are concentrated in low larval counts. The moderate high variability rainfall category is not discussed here as the sample size of events belonging to this category was not enough to make any conclusion based on it.

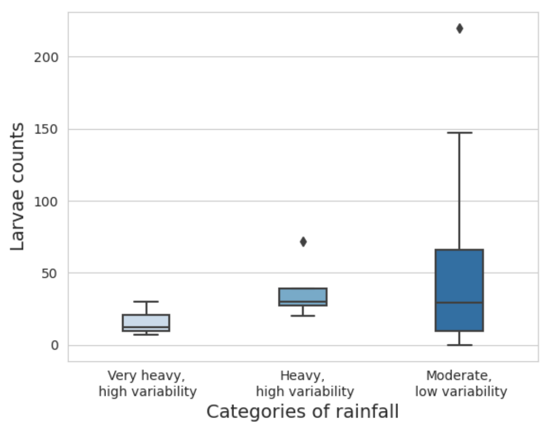


Fig. 3.12 Distribution of larvae counts associ-
-ated with different rainfall categories

The results from short term lag analysis in a whole highlighted the relationship between occurrence of larvae and immediately lagging precipitation. All the results were consistent with the finding of negative effect of heavy rainfall on larval abundance. However, the finding in Fig. 3.12 suggests that heavy rainfall could also have varied effects on larval abundance. The very heavy, high variability rainfall is most detrimental to mosquitoes and thus, temporarily favorable for humans in eliminating short term disease risk. However, as correlation analysis indicated, such rainfall events were also responsible for accumulation of water, creating habitats for mosquitoes. Thus, in the long term, the same rainfall events could be responsible for elevating disease spread risk.

CHAPTER 4: DISCUSSION AND CONCLUSION

4.1 Lag analysis of climate with WNV-positive mosquitoes

Despite considerable interest in the relationship between WNV transmission and climate conditions, previous studies present a range of outcomes for different locations and time periods. This may be in part because WNV's incidence is also dependent on factors like the abundance of infected hosts, virus incubation period, human population density in particular regions, and availability of suitable habitats for mosquitoes, mosquito growth period, etc. As it is not easy to track all these parameters consistently and simultaneously, it's often difficult to establish a clear relationship between WNV spread and these parameters and to identify the contribution of meteorological conditions among them.

In this study, the roots of the analysis were based on the consideration of a 3-week lag period to investigate the influence of meteorological conditions. The 3-week period was considered because it provides a sufficient window to account for mosquito growth, virus infection, incubation period, etc. Some of the consistent results over the study period suggest that warmer temperatures and some amounts of rainfall are necessary for the survival and reproduction of mosquitoes. The incidence of positive mosquito batches mainly followed dry conditions for about two days, but a dry period more than that saw declines in mosquito abundance, suggesting the need for moist conditions. High abundance of batches was strongly associated with sparse rainfall events combined with abundant the dry spells in the 3-week period. This result indicates the possibility of mosquitoes and hosts (birds) coming in proximity in search of water during the dry spells causing more infected mosquitoes (Paz, 2015).

4.2 Lag analysis of climate with mosquito larvae

4.2.1 Longer lag analysis

The main insight from correlation analysis was that in 2003, larval abundance was highly dependent on meteorological conditions, and it was also the year with highest larval abundance. 2003 also happened to be the hottest and wettest year in the sample. Overall, these insights suggested that the abundance of mosquitoes is driven by wetter conditions in the early summer and hotter conditions in the late summer.

The strong negative relationship of larvae occurring in early summer with the accumulated precipitation over the preceding one month (late spring) indicated the possible flushing effect that could have wiped out the mosquito population responsible for laying eggs in buckets. The strong positive relationship with mean temperature over the similar timeframe was suggestive of the requirement of high temperatures for higher abundance. Overall, the early summer relationships suggested that hotter and drier conditions were conducive to mosquito activity in that period. It also turned out that the relationship of larval abundance in early summer and late summer with rainfall were contradictory to each other.

Relationships in the August-September period were indicative of the hotter and wet conditions over preceding two months for high larval abundance. A possible explanation for this was that the hotter period could lead to drying out of the larval habitat and thus enough accumulation of water was necessary for the larval habitat and in turn adult mosquito emergence. As these meteorological conditions were not so influential in 2004 and 2005, it was difficult to generalize the conclusions and create a predictive model.

4.2.2 Shorter lag analysis

Unlike longer term analysis, results from short term analysis were consistent throughout the study period. The most intuitive explanation behind the relationship between rainfall spikes and larval population depressions was the flushing effect; however, due to the bucket setup and unavailability of data about rain rate and direction, the relationship could not be conclusive of flushing effect. Thus, another factor could be restricted mosquito movement to oviposit due to heavy rainfall events in the Midwest which likely occur with high winds (Hoffmann & Miller, 2003). It was challenging to quantify the negative relationship and threshold rainfall amount that results in flushing or restricting the mosquito activity due to the highly variable nature of rainfall.

4.2.2.1 Weather threshold estimation

Results obtained from the tree modeling approaches indicated that rainfall wasn't the primary variable in deciding the classification. The primary split was always based on temperature variables. It suggested that higher temperatures could be favorable for the high mosquito abundance as the node with splitting condition 'False' i.e., 'maxtemp7' > 35.9°C in the best performing model contained occurrences of only the high abundance category. The secondary split suggested that high rainfall amounts of around 78 mm which are not immediately preceding the larvae occurrence could favor high mosquito abundances. However, this split wasn't informative enough to get the threshold of rainfall where the larval abundance starts declining. It also indicates the presence of other factors playing a role in mosquito dynamics and suggests the need of finer analysis involving non-environmental factors and weather parameters other than rainfall and temperature. The results from these approaches did not yield the hypothesized output and that's likely because of the episodic nature of the rainfall with extreme rainfall events acting as outliers in the data.

However, boxplots of the rainfall and temperature associated with the larval categories helped in identifying the outliers and the interquartile range. Boxplots also distinctively showed that high abundance is favored by high temperature and low rainfall whereas high rainfall and low temperature result in low abundance in the short term. Thus, it confirms that mosquito survival is driven by hot weather with low moisture conditions.

More importantly, as this study does not directly address the larval abundance in stormwater infrastructure, it’s not the larval abundance itself that is driven by weather variables but it’s the factors like mosquito flight and oviposition activity. Hence, the results provide insights into these factors. As the high temperature and low moisture conditions favor high abundance, it means that mosquitoes are more active in these conditions and that has direct implications for WNV spread. Since adult female mosquitoes are the ones ovipositing, it’s certain that those are the ones responsible for infecting humans while taking the blood meal necessary for oviposition. Thus, in such a way, WNV spread is directly influenced by short term weather conditions. Therefore, WNV activity could potentially be controlled if the short-term weather conditions along with other environmental and non-environmental variables are considered which are beyond the scope of this analysis.

As the oviposition buckets were always filled with water with manually controlled water levels, it implies that the rainfall parameter here did not play a role in creating habitats for mosquitoes to breed or for the eggs to hatch. So, the short-term analysis was based on the assumption that eggs hatched within one day after oviposition. Thus, the weather conditions with time lag of one day explain the flight and oviposition activity of adult mosquitoes on that day. Therefore, the results from spatial analysis indicate that the heavy rainfall was responsible not only for flushing but also for inhibiting the mosquito flight and oviposition activity.

4.3 Summary and conclusions

Overall, this study identified the two crucial timeframes where weather may impact mosquito populations differently. The analysis for WNV infected mosquitoes revealed that hot and dry conditions within the 3-week period before infection are most favorable for WNV incidence among mosquitoes and thus alarming for human beings.

Long term analysis for larvae suggested that the impacts of accumulated precipitation and mean temperature over the timeframe of 2-12 weeks before larval occurrence depend on the climatological factors. The climatology analysis of the study period revealed that all three years were meteorologically different with2003 being the wettest and hottest. 2003 also was the year with highest larval abundance which was strongly correlated with precipitation and temperature in the long-term timeframe.

Short term analysis, however, yielded consistent results over the study period meaning that it neglected the background meteorology and focused on very short-term individual weather patterns. The main conclusion from various techniques is that the hotter and low rainfall conditions are most conducive for the larval abundance and high rainfall amounts drastically impact the larval population following it. Geospatial analysis which particularly focused on high rainfall events revealed that heavy rainfall with high spatial variability systems like MCS which are very frequent to Midwest are the most harmful for the larval emergence.

Thus, such analyses in different timeframes are very crucial to create a well-informed risk framework of WNV incidence. Long term analysis could help in informing WNV risk well before and could help in implementing proper actions to avoid the risk. Short term analysis on the other hand could inform immediate risks and could provide an exact timeframe for human beings to stay alert and aware of the disease spread risk.

REFERENCES

1. Armando, E. (2006). Gonotrophic cycle and survivorship of Culex quinquefasciatus (Diptera: Culicidae) using sticky ovitraps in Monterrey, northeastern Mexico. *Journal of the American Mosquito Control Association*, *22*(1). https://doi.org/10.2987/8756-971X(2006)22[10:GCASOC]2.0.CO;2

2. Chaves, L. F., & Kitron, U. D. (2011). Weather variability impacts on oviposition dynamics of the southern house mosquito at intermediate time scales. Bulletin of Entomological Research, 101(6), 633–641. https://doi.org/10.1017/S0007485310000519

3. DeGaetano, A. T. (2005). Meteorological effects on adult mosquito (Culex) populations in metropolitan New Jersey. International Journal of Biometeorology, 49(5), 345–353. https://doi.org/10.1007/s00484-004-0242-2

4. Drakou, K., Nikolaou, T., Vasquez, M., Petric, D., Michaelakis, A., Kapranas, A., Papatheodoulou, A., & Koliou, M. (2020). The Effect of Weather Variables on Mosquito Activity: A Snapshot of the Main Point of Entry of Cyprus. *International Journal of Environmental Research and Public Health*, *17*(4). https://doi.org/10.3390/ijerph17041403

5. Du, J. (2011). *NCEP/EMC 4KM Gridded Data (GRIB) Stage IV Data. Version 1.0* (1.0, p. 295313 data files, 2 ancillary/documentation files, 18 GiB) [GRIB-2: Gridded Binary Edition 2 (GRIB)]. UCAR/NCAR - Earth Observing Laboratory. https://doi.org/10.5065/D6PG1QDD

6. Epstein, P. (2001). West Nile virus and the climate. *Journal of Urban Health : Bulletin of the New York Academy of Medicine*, *78*(2). https://doi.org/10.1093/jurban/78.2.367

7. Gardner, A. M., Hamer, G. L., Hines, A. M., Newman, C. M., Walker, E. D., & Ruiz, M. O. (2012). Weather Variability Affects Abundance of Larval Culex (Diptera: Culicidae) in Storm Water Catch Basins in Suburban Chicago. Journal of Medical Entomology, 49(2), 270. https://doi.org/10.1603/me11073

8. Hoffmann, E., & Miller, J. (2003). Reassessment of the role and utility of wind in suppression of mosquito (Diptera: Culicidae) host finding: stimulus dilution supported over flight limitation. *Journal of Medical Entomology*, *40*(5). https://doi.org/10.1603/0022-2585-40.5.607

9. *Illinois Department of Public Health*. (n.d.). Retrieved May 1, 2023, from https://dph.illinois.gov/en.html

10. Lampman, R., & Novak, R. (1996a). Attraction of Aedes albopictus adults in sod infusion. *Journal of the American Mosquito Control Association*, *12*(1). https://pubmed.ncbi.nlm.nih.gov/8723268/

11. Lampman, R., & Novak, R. (1996b). Oviposition preferences of Culex pipiens and Culex restuans for infusion-baited traps. *Journal of the American Mosquito Control Association*, *12*(1). https://pubmed.ncbi.nlm.nih.gov/8723254/

12. Lebl, K., Brugger, K., & Rubel, F. (2013). Predicting Culex pipiens/restuans population dynamics by interval lagged weather data. Parasites & Vectors, 6(1), Article 1. https://doi.org/10.1186/1756-3305-6-129

13. Mavrakis, A., Papavasileiou, C., Alexakis, D., Papakitsos, E. C., & Salvati, L. (2021). Meteorological patterns and the evolution of West Nile virus in an environmentally stressed Mediterranean area. Environmental Monitoring and Assessment, 193(4), 1–11. https://doi.org/10.1007/s10661-021-09011-3

14. Mosquito-borne01-042716.pdf. (n.d.). Retrieved April 11, 2023, from https://dph.illinois.gov/content/dam/soi/en/web/idph/files/publications/mosquito-borne01-042716.pdf

15. Nosrat, C., Altamirano, J., Anyamba, A., Caldwell, J. M., Damoah, R., Mutuku, F., Ndenga, B., & LaBeaud, A. D. (2021). Impact of recent climate extremes on mosquito-borne disease transmission in Kenya. PLOS Neglected Tropical Diseases, 15(3), e0009182. https://doi.org/10.1371/journal.pntd.0009182

16. Paz, S. (2015). Climate change impacts on West Nile virus transmission in a global context. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1665). https://doi.org/10.1098/rstb.2013.0561

17. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*(85), 2825–2830.

18. Reed, S. M., & Maidment, D. R. (1995). *A GIS procedure for merging NEXRAD precipitation data and digital elevation models to determine rainfall-runoff modeling parameters*. https://repositories.lib.utexas.edu/handle/2152/6735

19. Reisen, W., Milby, M., Presser, S., & Hardy, J. (1992). Ecology of mosquitoes and St. Louis encephalitis virus in the Los Angeles Basin of California, 1987-1990. *Journal of Medical Entomology*, *29*(4). https://doi.org/10.1093/jmedent/29.4.582

20. Ruiz, M., Chaves, L., Hamer, G., Sun, T., Brown, W., Walker, E., Haramis, L., Goldberg, T., & Kitron, U. (2010). Local impact of temperature and precipitation on West Nile virus infection in Culex species mosquitoes in northeast Illinois, USA. *Parasites & Vectors*, *3*(1). https://doi.org/10.1186/1756-3305-3-19

21. *Water and Atmospheric Resources Monitoring Program—Illinois Climate Network, Illinois State Water Survey*. (n.d.). Retrieved May 1, 2023, from https://www.isws.illinois.edu/warm/datatype.asp