



# Short-term prediction of *Culex quinquefasciatus* abundance in Central North Georgia, USA, based on the meteorological variability

Navideh Noori<sup>1</sup> · Latif Kalin<sup>2</sup> · B. Graeme Lockaby<sup>2</sup> · Krisztian Magori<sup>3</sup>

Received: 13 December 2021 / Accepted: 18 April 2022 / Published online: 11 May 2022  
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2022

## Abstract

*Culex quinquefasciatus* is the main vector of West Nile Virus (WNV) in the southeast USA, and inter-annual variation in this vector abundance is mainly related to meteorological variability. In this study, short-term effects of meteorological conditions on seasonal variation in the vector abundance in the central north part of the State of Georgia, USA, from 2002 to 2009 were assessed. Four weeks moving average temperature, precipitation, potential evapotranspiration, and available moisture in the surface layer of soil were considered as risk factors. Cross-correlation maps were developed to investigate influences of preceding environmental conditions during a time-lagged interval on mosquito count data. The Poisson regression model and Artificial Neural Network (ANN) model were used for prediction purposes. Two sets of predictors were used: (1) the interval lagged climate data with the highest correlation and (2) single time lag antecedent *Culex* mosquito abundance up to 10 weeks prior to the events combined with lagged climate data. Results revealed that both models predicted the seasonal cycle of vector abundance fairly accurately, with ANN performing better than the regression model. The addition of antecedent mosquito data as input improved the prediction power of both models. The developed predictive models can be helpful in informed decision-making when high WNV activities are anticipated.

**Keywords** *Culex* mosquito population · West Nile Virus · Artificial neural network · Meteorological variables

## 1 Introduction

*Culex* species are the major vector of West Nile Virus (WNV) in the USA. This virus was first identified in North America in New York City in 1999. By the end of 2019, more than 51,801 disease cases of WNV had been reported

to the Centers for Disease Control and Prevention (CDC) [1]. Recently in 2020, 664 human cases and 52 deaths were reported to the CDC. In the southeast USA, over 96% of the WNV positive mosquito pools reported to the CDC from 1999 to 2010 have been obtained from *Culex* mosquitoes, among which 64.6% were from *Culex quinquefasciatus* (CQ) [2]. The state of Georgia, and especially the Atlanta area, was a hotspot of WNV incidence in 2012 with 117 WNV human cases, 6 deaths, and 125 WNV positive mosquito pools (among which over 81% were from CQ) reported by the Georgia Department of Public Health.

Transmission of vector-borne diseases is influenced by a wide range of environmental factors. Among these, meteorologic variability is one of the most important drivers of inter-annual WNV transmission risk. Weather directly affects the vector population, pathogens and hosts distribution, and their abundances [3–5]. *Culex* species display a seasonal behavior. Their activity reaches its minimal level in the winter and then rises in the spring to the peak levels of summer and continues until mid-fall [6]. Females which emerge in late summer search for sheltered areas where

✉ Navideh Noori  
navideh.noori@gatesfoundation.org

Latif Kalin  
latif@auburn.edu

B. Graeme Lockaby  
lockabg@auburn.edu

Krisztian Magori  
kmagori@ewu.edu

<sup>1</sup> Institute for Disease Modeling, Bill and Melinda Gates Foundation, 500 5th Ave N, Seattle, WA 98109, USA

<sup>2</sup> School of Forestry and Wildlife Sciences, Auburn University, 602 Duncan Drive, Auburn, AL 36849-5126, USA

<sup>3</sup> Department of Biology, Eastern Washington University, 258 Science Building, Cheney, WA 99004, USA

they become inactive until spring. They become inactive when the temperature drops below 60 °F [7], while warm weather brings them out in search of water on which to lay their eggs [8]. Changes in meteorological conditions such as temperature, relative humidity, and wind speed can impact mosquito populations [9]. The greatest WNV transmissions during the epidemic summers of 2002–2004 in the USA were linked to above normal temperatures. Analysis of temperature deviations from the 30-year mean (1971–2000) during summer in the USA showed that during years with above normal temperatures, WNV always dispersed into new areas and the amplification occurred during summers with above or normal temperatures. Subsequent cool summers were associated with decreased or delayed virus activity, especially at northern latitudes [10].

Temperature influences the development rate and fitness of immature mosquitoes and the biting rate and survival of adult female mosquitoes [11]. Drought can lead to a decline in the number of mosquito predators, and it may encourage birds to gather near standing water, where the virus can circulate more easily. High temperatures also speed the development of viruses within the mosquito carriers [12–14]. During periods of drought between rainfall events, blood-fed and potentially infected mosquitoes digest blood meals and wait for a heavy rainfall that floods the temporary pools to oviposit [15]. Rainfall and the surface moisture first create temporary freshwater habitats and also maintain permanent aquatic habitats that are used as egg-laying sites by female mosquitoes. Subsequently, rainfall saturates the ground and increases near-surface humidity levels [16].

The increase in the relative rate of WNV human cases from 2001 to 2005 in the USA has been linked to warmer temperatures, elevated humidity, and heavy precipitation independent of season using conditional logistic regression [17]. WNV mosquito infection changes from year to year spatiotemporally. For the temporal scale, higher temperature and less rainfall are associated with more human cases and also with the highest WNV prevalence in the mosquitoes. WNV infection can also be negatively correlated with the previous year's precipitation [18]. For the spatial analysis, temperature plays a bigger role than precipitation in comparison to temporal patterns [18]. Drought followed by wetting of the land surface is associated with the spatiotemporal variability of human WNV cases [19]. Spring drought induces the amplification of WNV by concentrating vector mosquitoes in humid vegetated areas where nesting birds are present. This makes the virus transmission easier as birds are the natural host of WNV, and this virus is maintained in nature in a mosquito-bird-mosquito transmission cycle. Subsequent summer rainfall and wetting of the land surface enable the dispersal of infected

mosquitoes into the open, sparsely vegetated areas they had avoided during the drought [20].

To control mosquito populations and to prevent disease, understanding this vector–environment relationship is essential. It is also helpful to understand the responses of WNV transmission risk to meteorological variability for public health policies so they can be adapted based on the consequent impacts [21]. Predictive models can be helpful in this regard to enhance the warning of high-risk periods for WNV and to describe the variations in mosquito abundance over time. Extensive attempts have been made to develop mosquito abundance prediction models which mostly rely on meteorological and environmental data from the days and weeks preceding the capture of mosquitoes [22]. Such models can be designed to provide continuous daily or weekly estimates of mosquito populations under the impacts of different environmental conditions. Ahumada et al. [23] proposed a discrete-time population model to simulate the temporal dynamics of *CQ* abundance. The model incorporated temperature and rainfall dependence and breeding site density dependent competition. This model simulated the mosquito population growth through time and at different elevations in Hawaii. Temperature was the major driving force behind mosquito population growth and abundance in Hawaii, but precipitation dependence also constrained population size which was evident during dry years.

A climate-based model was developed by [24] to predict mosquito abundance of WNV *Culex* species. Temperature, rainfall, evaporation, and photoperiod were used as inputs to the model. A moisture index was also created based on 7 days cumulative rainfall and evaporation. The model was developed on temperature-dependent functions including development rate and survival rate, a moisture index dependent function, and daily egg-laying rate.

The Dynamic Mosquito Simulation model (DyMSiM) developed by Morin and Comrie [25] was used in simulating *CQ* population dynamics in Florida and California. This model breaks up the larval phases into separate instar stages. The model used daily temperature and precipitation to drive population simulations throughout the year. This model revealed that dry conditions in California reduced mosquito populations due to loss of immature mosquito habitats, while drier late summer conditions in Florida decreased late-season adult mosquito populations.

In most of the previously mentioned analyses, the impact of meteorological conditions on mosquito abundance was limited to single point lags which consider the conditions at a certain time prior to trapping. Curriero et al. [9] introduced cross-correlation maps (CCMs) as a graphical method for visualizing the influence of preceding environmental conditions during a time lagged interval on the abundance of *Ochlerotatus sollicitans* species. Since

then, this tool has been used to identify the timing and duration of potential meteorological effects on mosquito populations [22, 26–28]. In this study, to investigate the correlation between meteorological variables and inter-annual and seasonal variation in *Culex* mosquito population carrying WNV, CCMs were developed for mosquito data from central north Georgia (GA). The main goal was to develop an improved predictive model of *CQ* populations by extending effects of meteorological conditions over a range of time rather than a single point in time. Two modeling approaches were applied in this study, multi-regression and Artificial Neural Network (ANN); lagged meteorological data were fed into these models for prediction purposes. In addition, as there is a correlation between any two observations of the time series of mosquito count data, antecedent conditions of response variable up to 10 weeks prior to the event were added to the models as predictors. It was hypothesized that addition of past values of mosquito count data to the model improves the model performance and increases the prediction accuracy.

## 2 Materials and methods

### 2.1 Mosquito and meteorological data

In this study, effects of meteorological variation on female *CQ* abundance per trap night were explored for central north Georgia. The weekly meteorological data including mean weekly precipitation, temperature, potential evapotranspiration (PET), and available moisture in surface layer from 2002 to 2009 were downloaded from the National Weather Service, Climate Prediction Center (CPC) ([http://www.cpc.ncep.noaa.gov/products/monitoring\\_and\\_data/drought.shtml](http://www.cpc.ncep.noaa.gov/products/monitoring_and_data/drought.shtml)). Climatic divisions were defined for the state of GA by CPC, and as the Atlanta metropolitan area is located in division 2, the central north part of GA, weekly climatic data were obtained for this division (Fig. 1). Soil moisture is estimated by a one-layer hydrological model [29, 30]. The model takes observed precipitation and temperature and calculates soil moisture, evaporation, and runoff. Potential evapotranspiration is computed from observed temperature and using Thornthwaite method [31]. Mosquito data were obtained from 2002 to 2009 for the counties located in division 2. Mosquitoes had been collected, classified by species, pooled by date, location, species and trap type, and tested for WNV infection [32]. Collections were done using paired CO<sub>2</sub>-baited CDC light traps [33] and gravid traps [34].

As *Culex* species either hibernate or become inactive during winter, no traps were set during winter and active mosquito counts were assumed to be zero for this period

[28]. Figure 1 shows the average weekly precipitation, temperature and female *CQ* abundance over the period 2002–2009 for counties located in the study area.

### 2.2 Statistical analysis

Female *CQ* mosquito time-series data and its potential relationship with each meteorological variable were analyzed using cross-correlation maps (CCMs). This graphical approach characterizes the temporal structure of mosquito population size in association with meteorological variables. Using this method, the key antecedent environmental conditions, their timings, and durations were identified which can improve the ability of developing predictive models of vector abundances.

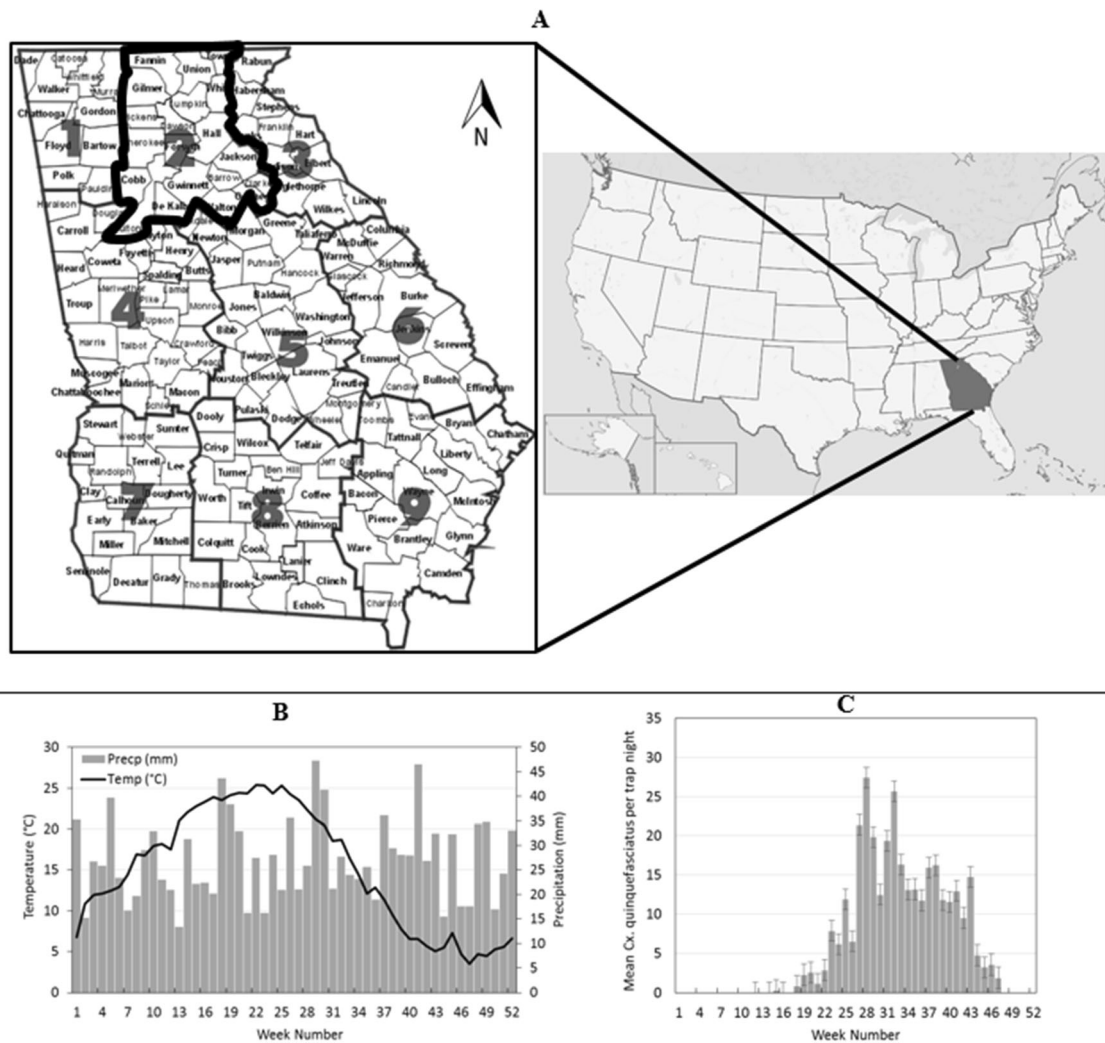
Assume  $Y(t)$  and  $X(t)$  represent two time series with time index  $t$ , CCMs illustrate the correlation coefficients ( $r$ ) between  $Y(t)$ , in this study number of captured female *Culex* mosquitoes at time  $t$ , and a meteorological variable  $X$ , averaged over a time period starting at time  $t - j$  and ending at time  $t - k$  with  $j \geq k$ :

$$r(Y, X_{j,k}) = \text{cor}(Y(t), \overline{X(t-j, t-k)}) \quad (1)$$

In this study,  $t$  changes from 1 to 52 for a given year (i.e., weekly time interval). Spearman's rank-order correlation was applied to calculate the correlations as it makes no assumption about the distribution of the data and does not consider a linear relationship between mosquito abundance and meteorological data. The CCMs were developed for four weeks moving average scale. As preceding meteorological conditions up to 5 months prior to summer play a significant role on the life cycle of mosquitoes, and also to make sure that sufficient time lag is searched to capture high correlations, the maximum time lag was set to 20 weeks. Other studies used similarly long time lags [28, 35]. In addition, the sample autocorrelation function (ACF) was defined for the time series of mosquito count data to identify the time interval over which a correlation in the data series exists. All analyses were performed in R statistical software (version 3.0.2.) [36].

### 2.3 Predictive models

As the response variable is count data, Poisson regression model was selected for prediction purposes. Poisson regression model assumes that the log-transformed outcomes are linearly related to the count data, and the mean and variance of date are equal. To overcome the limitations of statistical models, and to capture the potential complex nonlinear relationships between meteorological variables and mosquito abundance, the ANN model was also used. ANN is a black box type lumped model that has the ability

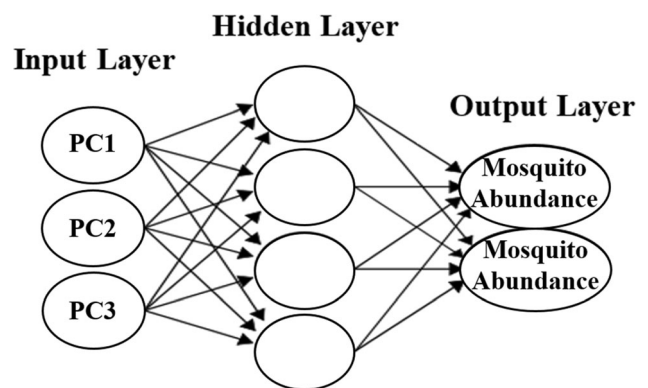


**Fig. 1** A Climatic divisions for the state of Georgia defined by Climate Prediction Center ([http://www.cpc.ncep.noaa.gov/products/monitoring\\_and\\_data/drought.shtml](http://www.cpc.ncep.noaa.gov/products/monitoring_and_data/drought.shtml)). B, C Average weekly weather

and mosquito abundance data over the period 2002 to 2009 for the central north of GA, respectively

to identify a relationship from given patterns which makes it possible to solve nonlinear models. ANNs can be categorized based on the direction of information flow and processing. In a feed-forward network, the connections between nodes are from an input layer, through one or more hidden layers, to an output layer [37] (Fig. 2). The most common method used to find the number of hidden layers and nodes is a trial-and-error approach [38]. In this study, the number of hidden neurons changed from 4 to 6, and number of hidden layer was set to 1 to build a parsimonious model and to avoid data overtraining. A neural network was constructed using MATLAB version 7.10.0 (2010) and was trained by adjusting the weights that link its neurons.

Some meteorological variables are highly correlated with each other (e.g., evaporation and temperature), which will cause high variance inflation in the Poisson regression



**Fig. 2** An example of feed-forward artificial neural network structure with three vectors as inputs, 1 hidden layer with 4 neurons and two output vectors



model. To handle such collinearities, Principle Component Analysis (PCA) was used. PCA is a variable reduction technique that uses orthogonal transformation to convert a set of observations of correlated variables into a set of values of linearly uncorrelated variables called principle components. PCA was applied to two sets of predictors: the first set included the interval lags of meteorological variables with highest correlation, temperature, precipitation, PET, and available moisture in surface layer, and the second set included the interval lags of meteorological variables and also antecedent conditions of *Culex* mosquito count. Components that explained the variability of observed data were fed into ANN and Poisson regression models. 70% of the time series data, selected randomly from the whole data set [39], was used for training and 30% of data was used for testing purposes into the ANN and regression models. Model performances were assessed with the coefficient of determination ( $R^2$ ), Nash–Sutcliffe efficiency ( $ENASH$ ) [40], and bias ratio ( $RBIAS$ ) [40].

### 3 Results

#### 3.1 Cross-correlation maps (CCMs)

Results of CCMs for each meteorological variable and *Culex* mosquito abundance (female *CQ* species) data are shown in Fig. 3. *Culex* vector abundance was positively correlated with temperature and PET, respectively, over 20 to 5 weeks prior to sampling,  $r(M, T_{20,5}) = 0.82$ , and 19 to 7 weeks prior to sampling,  $r(M, PET_{19,7}) = 0.82$ , and negatively correlated with four weeks moving average available moisture in surface layer over 16 to 8 weeks prior the capture event,  $r(M, \theta_{16,8}) = -0.75$ . Four weeks moving average precipitation over 20 to 13 weeks was positively correlated with mosquito abundance at week  $t$ ,  $r(M, P_{20,13}) = 0.1$ ; also precipitation one week prior to mosquito capture event was weakly and negatively correlated with vector abundance,  $r(M, P_{1,1}) = -0.09$ . As mosquito population density peaks in summer/early fall (Fig. 1C), counting back the lags with the highest correlation identifies the preceding late winter and spring as the most relevant time period.

#### 3.2 Principal component analysis (PCA)

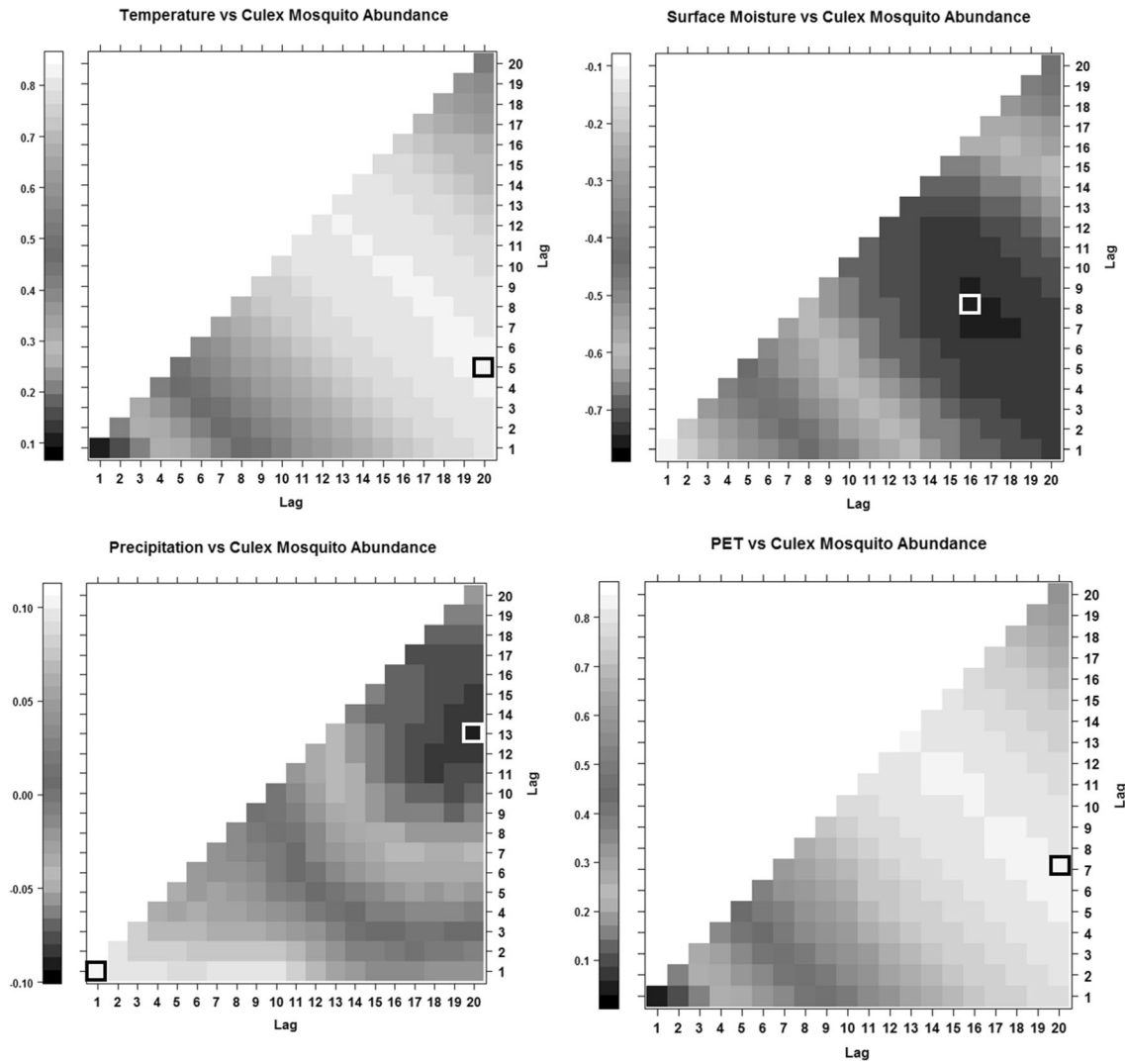
The weather data of interval lags with the highest positive or negative correlation were fed into PCA to eliminate collinearity. Table 1 shows the proportion of variance of each component and how much each variable contributed to that principal component. PCs 1, 2 and 3 together explained 97% of the variance in observed data. PC1 has

negative loadings for temperature and PET and positive loadings for surface moisture which considering summer/early fall as peak *Culex* species population corresponds to cold and moist late winter and spring. PC2 has strong negative loading for precipitation which reflects low precipitation in early spring and PC3 is positively related to precipitation one week prior the trapping event.

#### 3.3 Female *Culex quinquefasciatus* abundance prediction

Results of the Poisson regression model showed that all three PCs have a negative relationship with *Culex* mosquito count data. One unit increase in PC1 decreased mosquito abundance by 50% (0.50, 0.48–0.52, 95% C.L.), ( $p < 0.0001$ ), so half as many female *Culex* mosquitoes. PC2 is negatively related to mosquito data, and 1 unit increase in PC2 decreases mosquito abundance by 22% (0.78, 0.75–0.82, 95% C.L.) ( $p < 0.0001$ ) (Table 2). In addition, PC3 has a statistically significant negative relationship with mosquito data (0.94, 0.91–0.97, 95% C.L.) ( $p = 0.0004$ ) (Table 2).

Three PCs were randomized and fed to the ANN model as input for prediction purposes. Figure 4 compares ANN and the regression model performance versus the observed data, after sorting the randomized data (combined training and testing data). ANN predicted the four weeks moving average mosquito abundance more accurately with  $ENASH = 0.62$  and  $RBIAS = 9\%$  relative to the regression model with  $ENASH = 0.52$  and  $RBIAS = 18\%$ . To improve the model prediction accuracy, the antecedent four weeks moving average *Culex* mosquito abundance data up to 10 weeks prior to sampling was added to the PCA as a predictor. Table 3 shows the proportion of variance of components for each set of PCs. PCs 1 & 2 & 3 & 4 explained about 98% of variance in observed data. Components 1 & 2 & 3 had same interpretation as explained for Table 1. Component 4 corresponds to antecedent mosquito abundance condition. These components were fed into ANN and regression models for prediction purposes. Figure 5 compares the predicted versus observed data for testing and training periods. For all the data sets, ANN performed better, with higher  $ENASH$  values and smaller  $RBIAS$  values compared to the regression model, and as the lag interval increases, models performances gradually decrease (Fig. 6). This indicates that by combining interval lagged weather data and single time lag antecedent *Culex* mosquito abundance at the four weeks moving average scale, a stronger model with higher accuracy performance can be built for prediction purposes.



**Fig. 3** CCMs of 4 weeks moving average *Culex* mosquito abundance versus meteorological variables. Black or white rectangles show the interval lags with the highest correlation

**Table 1** PCA for meteorological variables with highest correlation with four weeks moving average *Culex* mosquito abundance

	Spearman correlation	PC1	PC2	PC3	PC4	PC5
$T_{20,5}$	0.82	− 0.57	− 0.20		− 0.39	0.69
$P_{1,1}$	− 0.09		− 0.20	0.98		
$P_{20,13}$	0.1	0.14	− 0.92	− 0.18	0.33	
$PET_{19,7}$	0.82	− 0.58	− 0.19		− 0.31	− 0.73
$\theta_{16,8}$	− 0.75	0.56	− 0.20		− 0.80	
Proportion of variance	−	56%	21%	20%	2%	0%

T, temperature; P, precipitation; PET, potential evapotranspiration,  $\theta$ , soil moisture in surface layer. First column shows the Spearman’s rank order correlation coefficient between each meteorological variable and four weeks moving average mosquito abundance. Second to sixth columns, i.e., PC1 to PC5, show PC loadings

**Table 2** Poisson model analysis of four weeks moving average *Culex* mosquito abundance

Intercept	Estimate	95% CI	<i>P</i> value	
	4.22 <sup>a</sup>	3.93	4.54	< 0.0001
PC1	0.50	0.48	0.52	< 0.0001
PC2	0.78	0.75	0.82	< 0.0001
PC3	0.94	0.91	0.97	0.0004

<sup>a</sup>Odds ratios were derived using the link function  $e^{(\text{coefficient})}$  and coefficient estimates of the GLM model

## 4 Discussion

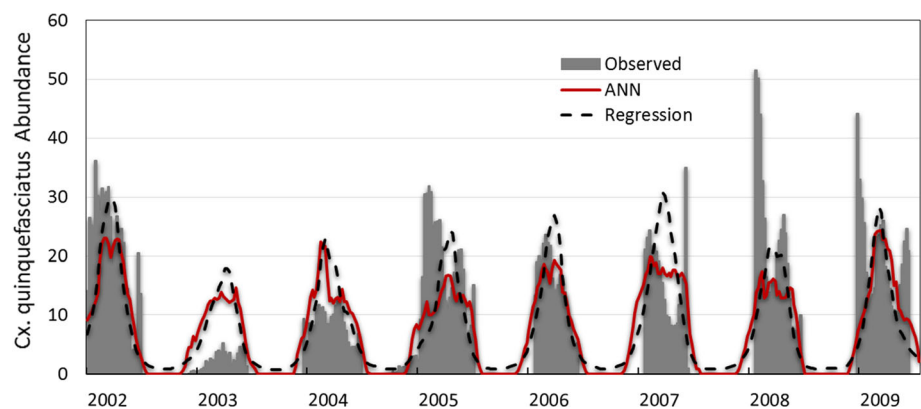
In this study, the associations of preceding meteorological conditions and *Culex* mosquito abundance were explored to enhance our understanding of mosquito ecology and disease risk of *CQ* vectors carrying WNV. To determine the maximum correlations between mosquito data and meteorological variables, cross-correlation maps (CCMs) were generated. The association of vector abundance with leading meteorological variables under specific time interval lags results in more robust inference than analyses that are restricted to single predefined time lags [27]. Using CCMs and considering interval lag structures, both the timing and duration of the meteorological effects are displayed [9]. The relationships revealed between interval-lagged environmental factors and the abundance of mosquitoes carrying WNV can be used as leading indicators of vector abundance. Predicting WNV activity is an essential requirement for vector control, and studying the *Culex* species population dynamics in relation to meteorological factors like ambient air temperature, surface moisture, and precipitation could help to improve the ability of predicting the WNV risk.

Using PCA, the collinearities among the meteorological variables were removed and the new components obtained at four weeks moving average scale were fed into the ANN

and Poisson regression models as explanatory variables. Considering the timing of peak mosquito abundance, which is summer and early fall, results of CCMs and the Poisson regression model reflected that elevated temperature and PET averaged over late winter and spring were closely associated with increased abundance of *CQ* in summer (considering mid-July as peak mosquito count, Fig. 1C). This is consistent with other field studies as larval and pupal developments are temperature dependent [27]. Also, drier than normal conditions during spring with low available moisture in surface layers creates favorable conditions for the development of *Culex* vectors in summer. Prolonged above normal temperature extends the duration of the mosquito season and vector activity. It also accelerates the development rate, influences the fitness of immature mosquitoes, the biting rate, and survival of adult female mosquitoes [13, 41]. Also, vector development conditions are facilitated and the frequency of transmission events is increased due to dry conditions by gathering hosts and vectors around nutrient-rich water bodies [19].

An increase in formation and persistence of mosquito development sites due to early period precipitation is associated with an increase in the abundance of *Culex* mosquitoes [42]. Extensive habitats of *Culex* mosquitoes can result from the heavy rains and associated flooding, especially in late winter and early spring, which is right before the mosquito life cycle starts. Habitats can include temporary ground pools, pools along receding river floodplains, or natural or man-made containers. However, the impact of precipitation on the mosquito population is controversial [3]. Generally, regions with lower seasonal variation in precipitation such as the southeastern USA have a lower probability of WNV mosquito cases [43]. Also, the southeastern USA receives sufficient precipitation to support mosquito populations throughout the year, making temperature the controlling variable affecting *Culex* mosquito population dynamics [40]. The CCMs obtained for precipitation versus mosquito count data support this statement. Due to exponential growth rates and

**Fig. 4** Comparison of predicted four weeks moving average *Culex* mosquito abundance by ANN and regression model versus observed data



**Table 3** PCA for meteorological variables and lagged four weeks moving average *Culex* mosquito abundance time series. Each set contains antecedent mosquito data changing from 1 to 10 weeks prior

	PC1	PC2	PC3	PC4	PC5		PC1	PC2	PC3	PC4	PC5
$T_{20,5}$	-0.52	-0.16		-0.31	-0.38	$T_{20,5}$	-0.52	-0.16		-0.32	-0.37
$P_{1,1}$		-0.20	-0.98			$P_{1,1}$		-0.16	-0.99		
$P_{13,20}$	0.11	-0.92	0.18		0.33	$P_{13,20}$	0.11	-0.93	0.15		0.33
$PET_{19,7}$	-0.52	-0.16		-0.27	-0.30	$PET_{19,7}$	-0.52	-0.15		-0.28	-0.30
$\theta_{16,8}$	0.50	-0.23		0.19	-0.81	$\theta_{16,8}$	0.50	-0.24		0.17	-0.81
$M_{t-1}$	-0.44			0.89		$M_{t-2}$	-0.44	-0.10		0.89	
%VAR	57%	18%	16%	7%	2%	%VAR	57%	18%	17%	7%	2%
$T_{20,5}$	-0.52	-0.15		-0.33	-0.37	$T_{20,5}$	-0.52	0.14		-0.33	-0.36
$P_{1,1}$		-0.15	-0.99			$P_{1,1}$		0.153	-0.99		
$P_{13,20}$	0.11	-0.93	0.14		0.33	$P_{13,20}$	0.10	0.92	0.14		0.33
$PET_{19,7}$	-0.52	-0.15		-0.28	-0.30	$PET_{19,7}$	-0.53	0.14		-0.29	-0.30
$\theta_{16,8}$	0.50	-0.24		0.15	-0.82	$\theta_{16,8}$	0.51	0.25		0.13	-0.82
$M_{t-3}$	-0.43	-0.13		0.89		$M_{t-4}$	-0.42	0.16		0.89	
%VAR	56%	18%	17%	7%	2%	%VAR	56%	18%	17%	8%	2%
$T_{20,5}$	-0.53	0.12		-0.32	-0.36	$T_{20,5}$	-0.53	-0.11	0.13	-0.30	0.37
$P_{1,1}$		0.21	-0.97			$P_{1,1}$		-0.29	-0.94	-0.18	
$P_{13,20}$	0.10	0.91	0.20	-0.12	0.33	$P_{13,20}$	0.10	-0.88	0.29	-0.14	-0.33
$PET_{19,7}$	-0.53	0.12		-0.28	-0.30	$PET_{19,7}$	-0.53	-0.10	0.12	-0.27	0.30
$\theta_{16,8}$	0.51	0.25		0.11	-0.81	$\theta_{16,8}$	0.51	-0.26			0.81
$M_{t-5}$	-0.41	0.20		0.89		$M_{t-6}$	-0.39	-0.24		0.88	
%VAR	55%	18%	16%	9%	2%	%VAR	54%	18%	17%	9%	2%
$T_{20,5}$	0.534		0.15	-0.27	-0.38	$T_{20,5}$	-0.54		0.17	-0.24	-0.38
$P_{1,1}$		-0.34	-0.91	-0.24		$P_{1,1}$		-0.40	-0.86	-0.32	
$P_{13,20}$	-0.11	-0.85	0.36	-0.15	0.33	$P_{13,20}$	0.11	-0.82	0.43	-0.16	0.33
$PET_{19,7}$	0.54		0.14	-0.25	-0.30	$PET_{19,7}$	-0.54		0.16	-0.23	-0.30
$\theta_{16,8}$	-0.52	-0.25			-0.81	$\theta_{16,8}$	0.53	-0.25			-0.81
$M_{t-7}$	0.37	-0.27	-0.10	0.88		$M_{t-8}$	-0.34	-0.31	-0.15	0.87	
%VAR	53%	19%	17%	10%	2%	%VAR	52%	19%	17%	11%	2%
$T_{20,5}$	-0.55		0.16	-0.22	-0.38	$T_{20,5}$	-0.56	-0.10	0.15	0.20	-0.39
$P_{1,1}$		-0.35	-0.88	-0.31		$P_{1,1}$		-0.31	-0.90	0.30	
$P_{13,20}$	0.12	-0.83	0.39	-0.18	0.33	$P_{13,20}$	0.12	-0.84	0.36	0.22	0.33
$PET_{19,7}$	-0.55		0.15	-0.22	-0.30	$PET_{19,7}$	-0.56		0.14	0.20	-0.30
$\theta_{16,8}$	0.53	-0.24			-0.81	$\theta_{16,8}$	0.54	-0.23			-0.81
$M_{t-9}$	-0.31	-0.33	-0.15	0.88		$M_{t-10}$	-0.26	-0.36	-0.15	-0.88	
%VAR	51%	19%	17%	12%	2%	%VAR	51%	19%	17%	12%	2%

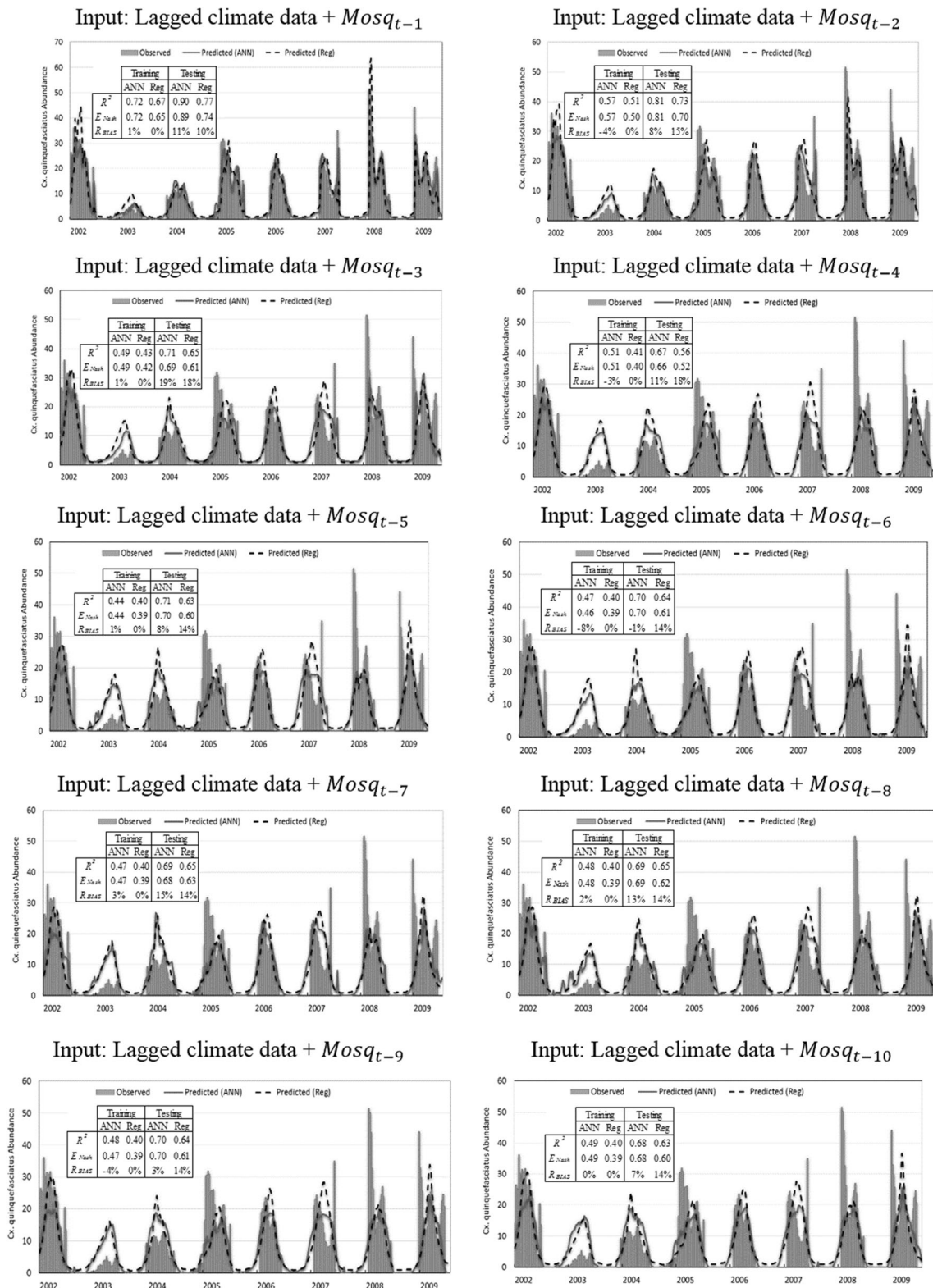
%VAR = proportion of variation

also complex interaction between mosquito abundance and rainfall, even small effects of weather conditions on a mosquito population could result in vast effects in future generations [3, 28]

The ANN and Poisson regression model predicted the seasonal cycle of mosquito abundance fairly accurately. The predictions improved significantly when antecedent conditions of mosquito count data up to 10 weeks prior to

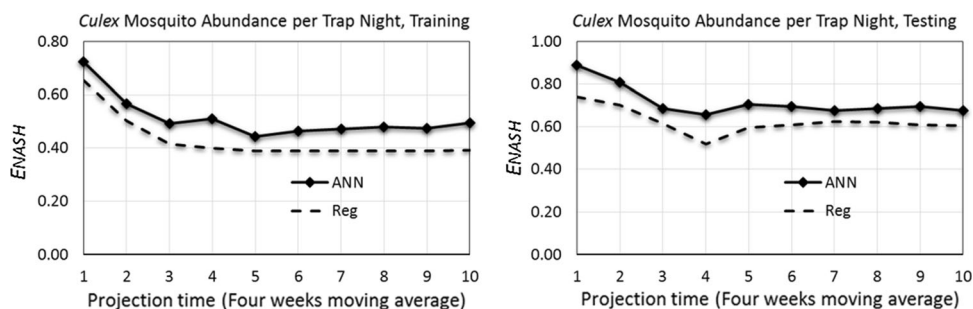
point of interest were added as predictors to the models. Addition of 1 week antecedent mosquito count data to the ANN model as a predictor increased the  $E_{NASH}$  value from 0.62 to 0.89 during the testing period. Also, the addition of 10 weeks antecedent mosquito abundance data to the ANN model improved the model performance during the testing period by increasing  $E_{NASH}$  from 0.62 to 0.68 (Fig. 6). ANN predicted the mosquito abundance slightly better than





**Fig. 5** Time series of predicted 4 weeks moving average *Culex* mosquito abundance by ANN and regression model versus observed data using different antecedent values of mosquito abundance and lagged weather data as predictors for training and testing periods

**Fig. 6** ANN and regression models performances built using interval lagged weather data and antecedent 4 weeks moving average *Culex* mosquito abundance for **a** training period, **b** testing period



the regression model, which could be due to the high non-linearity of ANN in comparison with Poisson which is a log-linear model. Generally, including the antecedent mosquito count to the model increased the predictive power of both ANN and regression models. This suggests that meteorological conditions and mosquito data from preceding weeks may be better indicators of future population dynamics for *Culex quinquefasciatus* mosquito species and the WNV risk than just the present size of the population.

## 5 Conclusions

The findings of this study and the developed ANN and Poisson regression models for predictions could have important implications for the control of West Nile Virus spread by *Culex* mosquito species. Most other studies developed mosquito abundance regression based models with single time lag antecedent weather data up to 2 months as explanatory variables without fixing the collinearities among meteorological variables and without extending effects of weather conditions over a range of time. Multi-collinearity can increase the variance of the coefficient estimates and reduce the statistical power of the analysis. In addition, single time lag might not capture meteorological effects on mosquito abundance if preceding conditions contributed to breeding and survival over weeks to months [27]. By collecting rigorous weather and mosquito data during important seasons, between February and June, and also addition of any antecedent mosquito count data 1–10 weeks prior, the size of vector populations that are likely to be seen in summer can be estimated and the possible abnormalities in the increase in rates of WNV infection can be monitored. These meteorological factors can be modeled under future warming conditions so that long-term predictions of shifts in risk can be estimated [44]. Such information could be used for planning of mosquito control strategies and to prioritize the distribution of scarce mosquito control resources before the transmission season begins. Also, it can help in early detection of virus circulation in mosquitoes and provide early

warning for WNV outbreaks. In years with warm spring and mild late winter, control operations such as applying insecticides can be initiated late in the winter to prevent rapid development of mosquitoes early in the spring and summer as a result of increase in survival rate of *Culex* mosquitoes throughout the winter. Although weather is the main driver of WNV risk and meteorological factors increase predictive power of determining risk associated with WNV, further studies are needed to explore whether other environmental factors such as socio-economic conditions, and landscape and mosquito habitat characteristics should be accounted for a better understanding of disease risk and for developing a more comprehensive *Culex* mosquito dynamic simulation model.

**Acknowledgements** We gratefully acknowledge the Southeast Cooperative Wildlife Disease Study (SCWDS) at the College of Veterinary Medicine at University of Georgia, Athens, for providing the mosquito data.

**Funding** This project is funded by USDA Forest Service, National Urban & Community Forestry Council and Center for Environmental Studies at the Urban–Rural Interface, School of Forestry and Wildlife Sciences, Auburn University.

## Declarations

**Competing interests** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

- ArboNET, Arboviral Diseases Branch, Centers for Disease Control and Prevention (2019) <https://www.cdc.gov/mosquitoes/mosquito-control/professionals/ArboNET.html>. Accessed Mar 2020
- Andreadis TG (2012) The contribution of *Culex pipens* complex mosquitoes to transmission and persistence of West Nile Virus in North America. *J Am Mosq Control Assoc* 28(4 Suppl):137–151. <https://doi.org/10.2987/8756-971X-28.4s.137>
- Roiz D, Ruiz S, Soriguer R, Figuerola J (2014) Climatic effects on mosquito abundance in Mediterranean wetlands. *Parasit Vectors* 7:1–13
- Paull SH, Horton DE, Ashfaq M, Rastogi D, Kramer LD, Diffenbaugh NS, Kilpatrick AM (2017) Drought and immunity determine the intensity of West Nile virus epidemics and climate

- change impacts. *Proc R Soc B* 284:20162078. <https://doi.org/10.1098/rspb.2016.2078>
5. Mordecai EA, Caldwell JM, Grossman K (2019) Thermal biology of mosquito-borne disease. *Ecol Lett* 2(10):1690–1708. <https://doi.org/10.1111/ele.13335>
  6. Strickman D (1988) Rate of oviposition by *Culex Quinquafasciatus* in San Antonio, Texas, during three years. *J Am Mosq Control Assoc* 4(3):339–344
  7. Rey JR (2011) The mosquito ENY-727 (IN652), one of a series of the Entomology and Nematology Department, Florida Cooperative Extension Service, Institute of Food and Agricultural Sciences, University of Florida
  8. American Mosquito Control Association (2014) <http://www.mosquito.org/>
  9. Curriero FC, Shone SM, Glass GE (2005) Cross correlation maps: a tool for visualizing and modeling time lagged associations. *Vector-Borne Zoonotic Dis* 5:267–275
  10. Reisen W, Fang Y, Martinez VM (2006) Effects of temperature on the transmission of West Nile Virus by *Culex tarsalis* (Diptera: Culicidae). *J Med Entomol* 43:309–317
  11. Dohm DJ, O'Guinn ML, Turel MJ (2002) Effect of environmental temperature on the ability of *Culex pipiens* (Diptera: Culicidae) to transmit West Nile virus. *J Med Entomol* 39:221–225
  12. Epstein PR (2001) West Nile Virus and the climate. *J Urban Health Bull N Y Acad Med* 78(2):367–371. <https://doi.org/10.1093/jurban/78.2.367>
  13. Patz JA, Olson SH, Uejio ChK, Gibbs HK (2008) Disease emergence from global climate and land use change. *Med Clin N Am* 92:1473–1491
  14. Wang G, Minnis RB, Belant JR, Wax ChR (2010) Dry weather induces outbreaks of human West Nile virus infections. *BMC Infect Dis* 10:1–7
  15. Day JF (2001) Predicting St. Louis encephalitis virus epidemics: lessons from recent, and not so recent. *Outbreaks Annu Rev Entomol* 46:111–138
  16. Shaman J, Stieglitz M, Stark C, Blancq SL, Cane M (2002) Using a dynamic hydrology model to predict mosquito abundances in flood and swamp water. *Emerg Infect Dis* 8:6–13
  17. Soverow JE, Wellenius GA, Fisman DE, Mittleman MA (2009) Infectious disease in a warming world: how weather influenced west nile virus in the United States (2001–2005). *Environ Health Perspect* 117:1049–1052
  18. Ruiz MO, Chaves LF, Hamer GL, Sun T et al (2010) Local impact of temperature and precipitation on West Nile virus infection in *Culex* species mosquitoes in northeast Illinois, USA. *Parasit Vectors* 3:1–16
  19. Shaman J, Day JF, Stieglitz M (2005) Drought-induced amplification and epidemic transmission of West Nile Virus in Southern Florida. *J Med Entomol* 42:134–141
  20. Day J, Shaman J (2008) Using hydrologic conditions to forecast the risk of focal and epidemic arboviral transmission in Peninsular Florida. *J Med Entomol* 45:458–465
  21. Wang J, Ogden NH, Zhu H (2011) The impact of weather conditions on *Culex pipiens* and *Culex restuans* (Diptera: Culicidae) abundance: a case Study in Peel Region. *J Med Entomol* 48:468–475
  22. Walsh AS, Glass GE, Lesser CR, Curriero FC (2008) Predicting seasonal abundance of mosquitoes based on off-season meteorological conditions. *Environ Ecol Stat* 15:279–291
  23. Ahumada JA, Lapointe D, Samuel MD (2004) Modeling the population dynamics of *Culex quinquefasciatus* (Diptera: Culicidae), along an elevational gradient in Hawaii. *J Med Entomol* 41:1157–1170
  24. Gong H, DeGaetano A, Harrington LC (2007) A climate based mosquito population model. In: Proceedings of the world congress on engineering and computer science, October 24–26, San Francisco, USA
  25. Morin CW, Comrie AC (2010) Modeled response to the West Nile virus vector *Culex quinquefasciatus* to changing climate using the dynamic mosquito simulation model. *Int J Biometeorol* 54:517–529
  26. Shone SM, Curriero FC, Lesser CR, Glass GE (2006) Characterizing population dynamics of aedes sollicitans (diptera: culicidae) using meteorological data. *J Med Entomol* 43:393–402
  27. Chuang TW, Ionides EL, Knepper RG, Stanuszek WW, Walker ED, Wilson ML (2012) Cross-correlation map analyses show weather variation influences on mosquito abundance patterns in Saginaw County, Michigan, 1989–2005. *J Med Entomol* 49:851–858
  28. Lebl K, Brugger K, Rubel F (2013) Predicting *Culex pipiens/restuans* population dynamics by interval lagged weather data. *Parasit Vectors* 6:1–11
  29. Huang J, Huug M, Dool VD, Georgakakos KP (1996) Analysis of model-calculated soil moisture over the United States (1931–93) and application to long-range temperature forecasts. *J Climate* 9(6):1350
  30. Van den Dool H, Huang J, Fan Y (2003) Performance and analysis of the constructed analogue method applied to US soil moisture over 1981–2001. *J Geophys Res* 108:1–12
  31. Thornthwaite CW (1948) An approach toward a rational classification of climate. *Geogr Rev* 38:55–94
  32. Vazquez-Prokopec GM, Vanden Eng JL, Kelly R, Mead DG, Kolhe P, Howgate J, Kitron U, Burkot TR (2010) The risk of west nile virus infection is associated with combined sewer overflow streams in urban Atlanta, Georgia, USA. *Environ Health Persp* 118:1382–1388
  33. Silver JB (2008) Sampling adults with light-traps. In: *Mosquito ecology: field sampling methods*, Springer, New York, pp 845–946
  34. Reiter P (1983) A portable, battery-powered trap for collecting gravid *Culex* mosquitoes. *Mosq News* 43(4):496–498
  35. Stoddard ST, Wearing HJ, Reiner RC Jr et al (2014) Long-term and seasonal dynamics of Dengue in Iquitos. *Peru PLoS Negl Trop Dis* 8:1–15
  36. R Core Team (2013) R: a language and environment for statistical computing. Reference Index. R Foundation for Statistical Computing, Vienna, Austria. Version 3.0.1
  37. Dawson CW, Wilby RL (2001) Hydrologic modeling using artificial neural networks. *Progs Phys Geogr* 25:80–108
  38. Kalin L, Isik S, Schoonover JE, Lockaby BG (2010) Predicting water quality in unmonitored watersheds using artificial neural networks. *J Environ Qual* 39:1429–1440
  39. Rezaeian Zadeh M, Amin S, Khalili D, Singh VP (2010) Daily outflow prediction by multi layer perceptron with logistic sigmoid and tangent sigmoid activation functions. *Water Resour Manage* 24:2673–2688
  40. Nash JE, Sutcliffe JV (1970) River flow forecasting through conceptual models, part I: a discussion of principles. *J Hydrol* 10:282–290
  41. Salas JD, Markus M, Tokar AS (2000) Streamflow forecasting based on artificial neural networks. In: *Artificial neural networks in hydrology*, Water Science and Technology Library, vol 36, Springer, pp 23–51

42. Morin CW, Comrie AC (2013) Regional and seasonal response of a West Nile virus vector to climate change. *PNAS* 110:15620–15625
43. Rosa R, Marini G, Bolzoni L, Neteler M et al (2014) Early warning of West Nile virus mosquito vector: climate and land use models successfully explain phenology and abundance of *Culex pipiens* mosquitoes in north-western Italy. *Parasit Vectors* 7:1–12
44. Harrigan RJ, Thomassen HA, Buermann W, Smith ThB (2014) A continental risk assessment of West Nile virus under climate change. *Glob Change Biol* 20:2417–2425

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.