Environ Ecol Stat (2008) 15:279–291
DOI 10.1007/s10651-007-0056-6

Predicting seasonal abundance of mosquitoes based on off-season meteorological conditions

Andrew S. Walsh · Gregory E. Glass · Cyrus R. Lesser · Frank C. Curriero

Received: 1 June 2005 / Revised: 1 October 2005 / Published online: 3 December 2007

Abstract Modeling mosquito population dynamics has become an important part of understanding the transmission of mosquito-borne arboviruses. Of these models, those including meteorological variables have mainly focused on conditions during or immediately preceding the mosquito breeding season. While these conditions are clearly critical biologically and statistically, it is also biologically plausible that conditions during the off-season may contribute to interannual variation in mosquito population size. To examine the effect of off-season factors, we develop a pair of Poisson regression models for July captures of *Aedes sollicitans* and *Culex salinarius*, two East Coast vector species of arboviruses including Eastern equine encephalitis virus and West Nile virus. Model results indicate that average maximum temperature, total heating degree-days, and the total number of days with a minimum temperature below freezing during the winter months was predictive of mosquito populations. In

A. S. Walsh (✉)
Language Technologies Institute, Carnegie Mellon University School of Computer Science,
5000 Forbes Avenue, Pittsburgh, PA 15213-3891, USA
e-mail: awalsh@cs.cmu.edu

G. E. Glass
W. Harry Feinstone Department of Molecular Microbiology and Immunology, Johns Hopkins Bloomberg School of Public Health, 615 North Wolfe Street, Baltimore, MD 21205, USA

C. R. Lesser
Mosquito Control Section, Maryland Department of Agriculture, 50 Harry S Truman Parkway,
Annapolis, MD 21401, USA

F. C. Curriero
Department of Environmental Health Sciences, Johns Hopkins Bloomberg School of Public Health,
615 North Wolfe Street, Baltimore, MD 21205, USA

F. C. Curriero
Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health,
615 North Wolfe Street, Baltimore, MD 21205, USA

addition, the average maximum relative humidity from the preceding fall and total rainfall and total heating degree-days during the preceding spring were also associated with vector population dynamics. The descriptive and predictive power of these models is discussed.

Keywords Climate · *Culex salinarius* · *Aedes sollicitans* · Poisson regression · Population dynamics

1 Introduction

Mosquito-borne arboviral infections are an annual public health concern in the United States. The recent emergence and spread of West Nile virus has brought media attention to the issue, but other arboviruses like St. Louis encephalitis and Eastern equine encephalitis have been endemic for a number of years and continue to remain a concern. A full understanding of the dynamics of these diseases has remained elusive; the involvement of multiple vector species and reservoirs with their own unique population dynamics complicate the system substantially. Each vector species may have unique ideal conditions for egg-laying, larval development, and adult activity that make ecology an important component of understanding the patterns of these diseases.

Considerable effort has been made to develop models that can predict the population dynamics of various mosquito vector species. Many of these models rely on meteorological and environmental data from the days and weeks preceding the capture of mosquitoes. For example, Focks et al. (1993a,b) created a model of the daily survivorship and development of *Aedes aegypti* based on multiple factors including the temperature and precipitation on that day. Evans et al. (1987) were able to predict the timing of emergence, but not the size, of broods of *Aedes taeniorhynchus* using rainfall, tide and temperature from the preceding week. These and other models have demonstrated that much of the variation in mosquito populations is influenced by prevailing conditions while the mosquitoes are developing.

Recently Curriero et al. (2005) introduced a graphical method for visualizing associations between mosquito populations and preceding meteorological conditions that allows these variables to extend over an interval of time, thus identifying the timing and duration of these potential meteorological effects. This concept was extended in Shone et al. (2006) and shown to be a successful tool for building regression models of daily trap counts of *Aedes sollicitans* on a year by year basis. It was observed that while the same set of meteorological variables and time intervals could be used for all years, different coefficient values needed to be estimated for each year to retain good model performance.

In temperate climates, mosquitoes must suspend their activity in order to survive the winter; the conditions during this period may also influence the population dynamics of mosquitoes and explain some of the interannual variation observed in the coefficient values. The effects of these off-season conditions on mosquito population growth have not been studied as extensively and the results have been mixed. Some have hypothesized that milder winters may lead to larger populations the following summer (Mogi 1996). On the other hand, a study by Wegbreit and Reisen (2000) demonstrated that higher levels of snow moisture and thus higher runoff from melting snow led to increases in the *Culex tarsalis* population. Further complicating the picture is another study that showed that droughts can actually lead to increased populations of mosquitoes through an effect on predation (Chase and Knight 2003); this is contrary to the notion that more water leads to more mosquitoes. These results suggest that the effects of off-season conditions are complex, with both direct and indirect mechanisms of action. They probably also differ by species, particularly since different species employ different strategies for surviving winter and rely on different cues to emerge. Clearly additional work needs to be done to fill in more of the details.

The purpose of this study is to determine what factors from the off-season, if any, impact the population size in subsequent seasons of two important vector species in Maryland: *Ae. sollicitans* (Walker) and *Culex salinarius* (Coquillett). *Aedes sollicitans* is an major vector of Eastern equine encephalitis and Venezuelan equine encephalitis (O'Meara 1992) and has shown some potential as a vector for West Nile virus (O'Leary et al. 2002, Turell et al. 2001b). *Culex salinarius* is considered a bridge vector for Eastern equine encephalitis, St. Louis encephalitis and West Nile virus (Nayar et al. 1986, Nasci et al. 1993, Turell et al. 2001a,b). *Aedes sollicitans* females lays eggs in the fall which enter diapause to endure the winter (O'Meara 1992), while *C. salinarius* is thought to overwinter as inseminated adult females that may even be periodically active (Slaff 1990).

2 Methods

2.1 Mosquito and meteorological data

Data on daily mosquito trap captures carried out between 1956 and 1989 were made available by the Mosquito Control Section of the Maryland Department of Agriculture. Mosquitoes were collected each night in New Jersey light traps and identified to species; the numbers of males and females of each species were recorded. This surveillance was done at a number of sites in the Chesapeake Bay area; further details on these data were reported by Shone et al. (2006). The site in Cambridge, Maryland (76.079°W, 38.558°N) was used for this analysis since the ultimate goal is to extend analyses already performed on the data from that site. Daily totals for the species of interest were summed over each of the months of the trapping season (June through September).

The meteorological data were obtained from the nearest observation point with available records for each variable. Total daily precipitation (in hundredths of an inch) and maximum and minimum daily temperature (in degrees Fahrenheit) were obtained from a station at the Cambridge Water Treatment Plant, 1.44 km from the light trap location. These data are part of the Cooperative Summary of the Day TD3200 data set provided by the National Climatic Data Center (NCDC). Average daily temperature was calculated as the mean of the maximum and minimum daily temperature. Heating degree-days were computed as the minimum of 65°F less the average daily temperature and 0. Indicators of a cool day and a cold day were defined as 1 when the minimum and maximum temperatures, respectively, were below freezing (32°F) and

Table 1 The off-season meteorological variables, their associated summary functions when aggregated over a range of days, and their notation used in modeling July capture totals

Variable	Summary function	Notation
Maximum daily temperature	Mean	TMAX.AVG
Average daily temperature	Mean	TAVG.AVG
Minimum daily temperature	Mean	TMIN.AVG
Daily precipitation	Sum	PRCP.SUM
Maximum daily relative humidity	Mean	MXRH.AVG
Minimum daily relative humidity	Mean	MNRH.AVG
Heating degree-days	Sum	HTDD.SUM
Maximum daily temperature	Sum of observations $< 32°$F	COLD.SUM
Minimum daily temperature	Sum of observations $< 32°$F	COOL.SUM

0 otherwise. Daily values for the minimum and maximum relative humidity (in percent) were measured at Baltimore-Washington International Airport, 86.09 km from the trapping site; these measurements were obtained from the Solar and Meteorological Surface Observation Network (SAMSON) data set also provided by the NCDC. Relative humidity observations were only available from 1961 to 1988. Table 1 provides a full list of the variables considered, the summary functions applied to these variables when aggregated over a range of days, and their corresponding notation.

2.2 Exploratory analysis

Cross correlation maps (Curriero et al. 2005) provide a graphical method for visualizing bivariate time series associations where the lag effect of one variable is allowed to extend over an interval of consecutive time points. Letting Y_i and X_i represent two time series with time index i, cross correlation maps display corr$(Y_i, f(X_{i-j,i-k}))$, for positive lags j, k with $j \geq k$, representing the start and end dates respectively, and some aggregate summary function $f(\cdot)$. In the current application Y_i represents the time series of monthly mosquito trap totals and $f(X_{i-j,i-k})$ a summary of a preceding off-season meteorological variable. For example, Y_i would represent July trap totals, $X_{i-j,i-k}$ would be the maximum daily temperature between December 15 and December 26, with j, k set to represent this lag interval in December, and $f(\cdot)$ would be the mean of those temperatures. As a model building tool, we extend this approach as in Shone et al. (2006) to consider reduction in Poisson model deviance (RMD) as the measure of association instead of correlation.

To explore the effects of off-season conditions, this reduction in model deviance was computed for each of the variables listed in Table 1 for univariate models. It was assumed that conditions on a single day during the off-season would not be as biologically meaningful as a summary of conditions over intervals of consecutive days, although information from both interval and single day lag structures are included in the cross-correlation maps and the RMD alternatives. Some exploration was also done to determine the most appropriate way to summarize seasonal mosquito totals. This involved calculating the RMD for each meteorological variable over a range of mosquito outcomes as well; every 4-week period from June 1 to September 30th was analyzed in this way.

2.3 Model building

A Poisson regression model of July trapping totals was built through an iterative process using conditional RMD. In each iteration, the RMD values were calculated for all of the off-season variables conditional on other covariates already selected in prior iterations to be included in the model. The covariate and lag interval yielding the greatest reduction in deviance when added to the model was included in subsequent iterations. For example, after the first iteration for *Ae. sollicitans*, the variable TMAX.AVG (Dec. 15–26), the maximum temperature averaged over Dec. 15– 26, was selected as having the largest deviance reduction. For the second iteration, the RMD value for all the variables was calculated conditional on the fact that TMAX.AVG (Dec. 15–26) was already a covariate in the model. This process was repeated until negligible reductions in deviance were obtained for all variables.

To assess the relative statistical importance of the covariates chosen by the iterative process, a Poisson model was calculated for every possible subset of the chosen covariates. The lowest Akaike's Information Criterion (AIC) (Sakamoto et al. 1986) among models with a given number of covariates was plotted against the number of covariates in the model. As an example, assume a three covariates model had been selected by the iterative process. There are seven possible subsets of these covariates, excluding the null set: the three covariates individually, three pairwise combinations, and the full set of three. The AIC for all seven of these models is computed. The models are then grouped by the number of covariates they contain (one, two, or three) and the lowest value from each group is chosen. These values are then plotted against the number of covariates for their respective groups. This plot was used to determine the minimal set of covariates needed to maintain a model fit comparable to a maximal model, that is the model with the full set of covariates. That determination was made by looking at where the largest change in AIC occurred. This minimal model was then examined in more detail along with the maximal model.

2.4 Model evaluation

To check for overdispersion in the models, the Pearson χ^2 goodness-of-fit statistic was calculated as

$$P_{\chi^2} = \sum_{i=1}^{n} \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i}$$

where y_i and $\hat{\mu}_i$ represent the yearly July totals and fitted Poisson regression values, respectively. The ratio of P_{χ^2} to the residual degrees of freedom should be approximately 1; otherwise there may be over- or under-dispersion. To correct for over-dispersion when determining the statistical significance of the regression coefficients, the standard error of the coefficients was multiplied by the square root of this ratio before computing a z-score (McCullagh and Nelder 1983).

A Monte Carlo approach was taken to assess the predictive power of the models. In each step, 21 of the 28 years of data were randomly selected and coefficients were estimated from those data using the covariates and summary interval lags previously selected. These coefficients were then used to predict July totals for the remaining seven years. This process was repeated 2,000 times so that each year would be a part of the seven predicted years an average of 500 times. The mean of these approximately 500 predicted values was computed for each year and the fit of this model summary was determined. Model fit was assessed using the R^2_{KL} measure defined by Cameron and Windmeijer (1997) as

$$R^2_{KL} = 1 - \frac{\sum y_i \log(y_i/\bar{\mu}_i) - (y_i - \bar{\mu}_i)}{\sum y_i \log(y_i/\bar{y})}$$

where $y_i, i = 1, \ldots, 28$, is the observed July total and $\bar{\mu}_i$ is the mean over the 500 predicted values for the jth year, and $y_i \log y_i = 0$ when $y_i = 0$. This R^2_{KL} statistic was developed analogous to the R^2 from standard linear regression by taking the Kullback–Leibler divergence as a generalized measure of variance. Although the interpretation of the statistic as the percent of variance explained does not carry over, it is still true that values closer to 1 indicate greater model fit. A 95% prediction interval was also computed for the predictions from each year by calculating the 2.5 and 97.5 percentiles of their empirical distribution.

All plotting and analysis was performed using the R statistical computing environment (R Development Core Team 2004).

3 Results

3.1 Exploratory analysis

A total of 32,433 *Ae. sollicitans* and 13,299 *C. salinarius* female mosquitoes were collected at the Cambridge site between June 8, 1961 and September 15, 1988. July totals ranged from 0 to 7,581 for *Ae. sollicitans*, with a median of 91.5; the first and third quartiles were 34 and 287 respectively. For *C. salinarius*, July totals ranged from 0 to 871; the median total was 3 and the first and third quartiles were 0 and 72.

Reduction in model deviance (RMD) was plotted for each off-season meteorological variable under consideration; a representative plot is shown in Fig. 1. Pixels on the diagonal of this plot (where start date = end date) show the strength of the association between July totals of *Ae. sollicitans* and the average temperature averaged over a single day in the off-season. Moving to the left along the x axis shows how the strength of the association changes as the interval is widened to include earlier dates as well; moving downwards on the y axis adds later dates to the end of the interval. In general, the strongest and most stable (persisting even when a few days were added to or

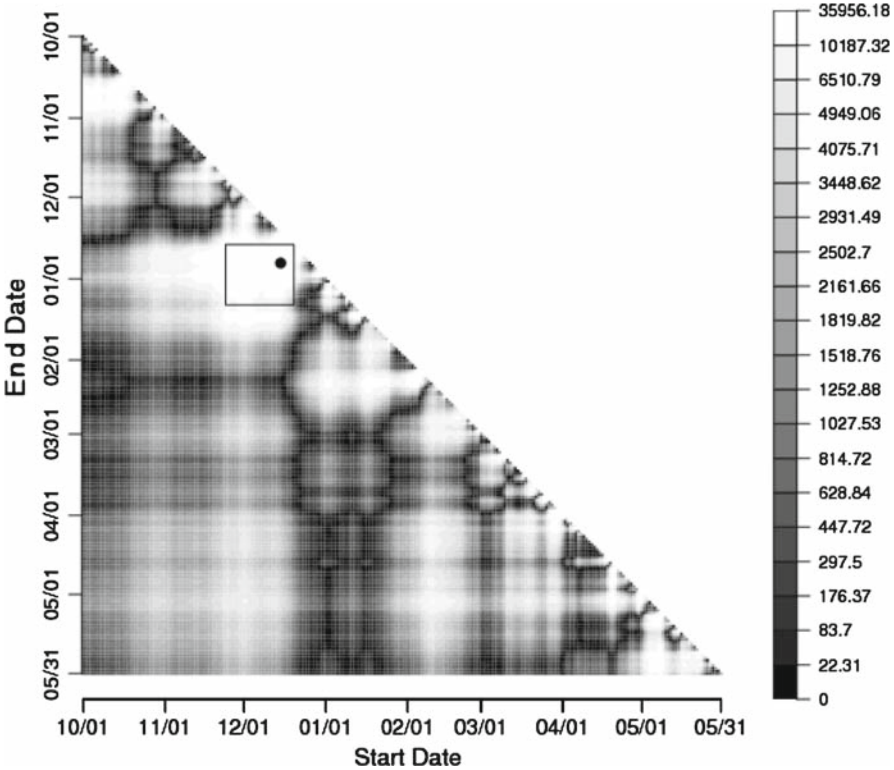

Fig. 1 Plot of the reduction in model deviance (RMD) for the average temperature (TAVG.AVG) averaged over the specified interval (Start Date, End Date) in a univariate model of July totals of *Ae. sollicitans*. The box highlights a region with a large reduction in variance that is stable when a few days are added to or removed from the interval

removed from the interval lag) effects were seen over two to 4 week intervals ranging over the entire off-season. The outlined box in Fig. 1 illustrates this phenomenon; the reduction in deviance remains large for the mean temperature averaged over intervals starting anywhere from Nov. 25 to Dec. 20 and ending anywhere from Dec. 20 to Jan. 10. Some strong effects were seen from very short intervals even down to one day; these were considered to be less biologically relevant and therefore not considered further. Most of the plots exhibit a pattern of striping as one moves either vertically or horizontally from the diagonal. This results from the fact that moving away from the diagonal implies that a larger period of time is being summarized. The smaller time intervals represented close to the diagonal begin to be grouped together, and they can either amplify each other's effect or diminish it. Alternating between this amplification and diminishment leads to the observed pattern of stripes.

Some exploration was done to determine the appropriate summary of a given trapping season. The strongest correlations for most of the meteorological variables were seen with the total number of mosquitoes caught during the 4 weeks following either June 15 or July 1. For the purposes of modeling, it was decided to use the July total as the outcome for both species. Since July is typically the beginning of the peak mosquito season in the region, it seems reasonable that the effects of the preceding off-season would be felt most strongly during that month.

3.2 Model development

Separate Poisson regression models were constructed iteratively for the total numbers of *Ae. sollicitans* and *C. salinarius* females trapped during July in the years 1961–1988. A sample plot of the conditional RMD used for generating these models is shown in Fig. 2. This plot can be read just as the plot in Fig. 1; each pixel corresponds to the conditional RMD for the COOL.SUM variable over the interval specified by the coordinates. The only difference is that now the reduction in deviance is conditional on TMAX.AVG (Dec. 15–26) already being included in the model. The value for the largest conditional RMD is highlighted in plot; it corresponds to the interval from Mar. 21 to Apr. 19. The variables and corresponding timeframes included in these models

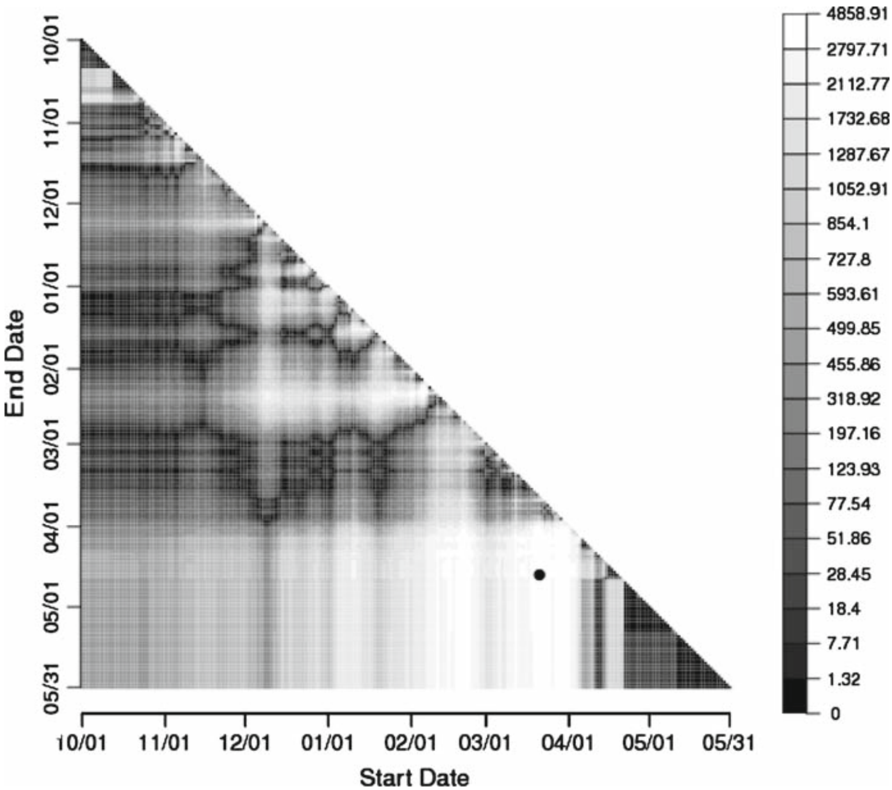

Fig. 2 A plot of the conditional RMD for the number of days with a minimum temperature below 32°F (COOL.SUM) over all possible intervals included in a model of July totals of *Ae. sollicitans* containing TMAX.AVG (Dec. 15–26). The black circle indicates the largest RMD and corresponds to Mar. 21–Apr. 19

Table 2 The meteorological variables and the interval lags over which they were summarized as selected by the model building scheme; also shown are the exponentials of the regression coefficients from the reduced term models. All coefficients were statistically significant even after correcting for over-dispersion. Estimates are only provided for meteorological variables included in the final models; other variables are denoted by NA

Variable	Summary interval	Three covariate model
Ae. sollicitans		
TMAX.AVG	Dec. 15–Dec. 26	0.795
COOL.SUM	Mar. 21–Apr. 19	1.35
PRCP.SUM	Apr. 18–May 18	0.996
COOL.SUM	Mar. 21–Apr. 13	NA
MNRH.AVG	Nov. 05–Feb. 02	NA
PRCP.SUM	May 03–May 19	NA
MNRH.AVG	Apr. 03–May 05	NA
HTDD.SUM	Feb. 10–Feb. 28	NA
Variable	Summary interval	Four covariate model
C. salinarius		
TMAX.AVG	Mar. 12–May 29	NA
COOL.SUM	Oct. 28–Nov. 12	2.74
TAVG.AVG	Jan. 06–May 27	NA
PRCP.SUM	Dec. 17–Jan. 17	NA
TAVG.AVG	Jan. 11–May 30	NA
MXRH.AVG	Oct. 01–Oct. 15	0.810
HTDD.SUM	Mar. 06–Apr. 13	1.016
HTDD.SUM	Jan. 10–Jan. 31	0.989

are shown in Table 2. The model for *Ae. sollicitans* had an AIC of 455.4 and an R^2_{KL} of 0.994; the model for *C. salinarius* had an AIC of 148.7 and an R^2_{KL} of 0.990. The excellent fit of these models is tempered by the fact that the number of covariates (8) is large relative to the total number of values to be fit (28 years).

To determine if subsets of the eight covariates would yield models that performed nearly as well, additional models were built using all of the 255 possible subsets of the eight covariates. By comparing the models with the lowest AIC for a given number of covariates (Fig. 3), it was determined that a three covariate model would perform nearly as well for *Ae. sollicitans* ($R^2_{KL} = .961$) and a four covariate model would be adequate for *C. salinarius* ($R^2_{KL} = .941$). The variables used in these smaller models are indicated in Table 2. These reduced models were tested for over-dispersion; the model for *Ae. sollicitans* had P_{χ^2} to residual df ratio of 90.73 and the model for *C. salinarius* had a P_{χ^2} to residual df ratio of 11.18. The coefficients of the models remained statistically significant after correction for this over-dispersion. Autocorrelation plots of residuals from both species' models exhibit no residual temporal dependence, making this a valid correction for over-dispersion.

Additional regression models were built for both *Ae. sollicitans* and *C. salinarius* that included the smaller sets of covariates as well as second-order interaction terms.

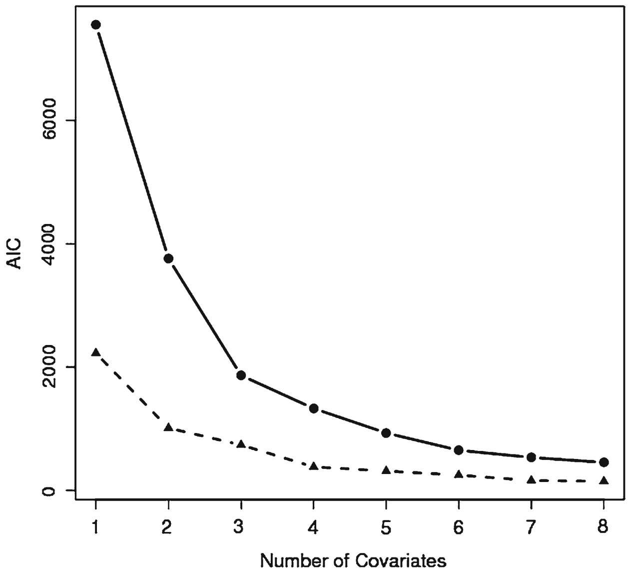

Fig. 3 Every combination of the eight variables in the *Ae. sollicitans* model (solid line) and the *C. salinarius* model (dashed line) was evaluated in a separate model. The lowest AIC for a model with the given number of covariates is depicted

These interactions did improve the fit of the model, but the effect was negligible. Furthermore, the coefficients for the interaction terms were not statistically significant after applying the correction for over-dispersion.

The exponentiated coefficients from the minimized models are shown in Table 2. For both species, the magnitude and sign of the coefficients for each variable generally remained consistent between the eight covariate models and the reduced-term models. Examination of the three-covariate model for *Ae. sollicitans* suggests some surprising connections between off-season conditions and the July captures. The coefficient for the average maximum temperature in late December is less than 1, suggesting that perhaps warmer weather in that interval leads to premature hatching, fewer eggs laid, or an increase in predator activity. Similarly, the number of days with a minimum temperature below freezing in late March to early April has a coefficient greater than 1, suggesting a positive relationship between cold weather and July totals. Early spring may also be a critical period for predator activity or premature hatching. The effect of total rainfall in late April to early May is negative; this may also be an effect on the timing of hatching or may be from flooding of egg sites. Changes in the timing of hatching could lead development of later stages in suboptimal climate conditions, or it could simply shift the timing of broods causing a peak in June or August rather than July.

Given that *C. salinarius* overwinter as adults rather than eggs, it is not surprising that the most significant off-season factors are very different. The positive effect of the number of cool days in early November may have to do with the timing of entry to hibernation. *Culex salinarius* can be active into the late fall; if survival is low during this time, it may be advantageous to begin hibernating earlier. It may also have to

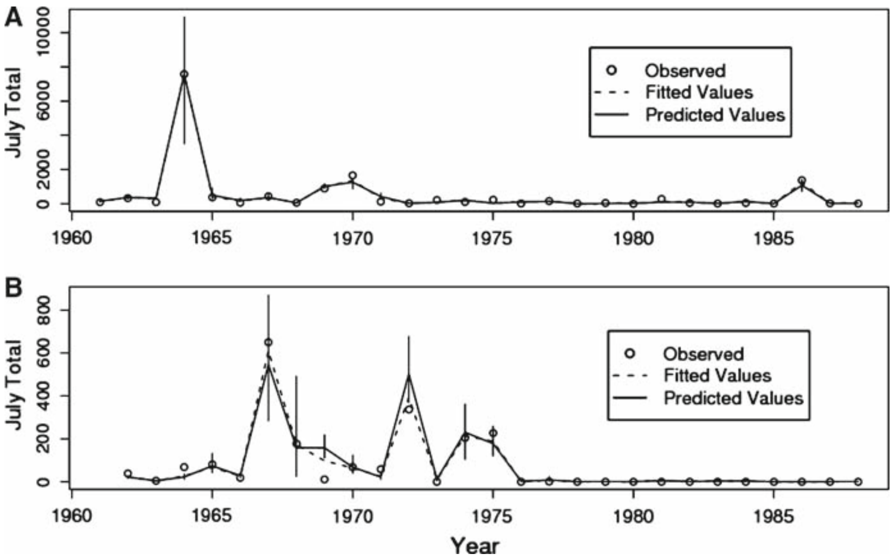

Fig. 4 The open circles represent the observed July totals of *Ae. sollicitans* (**a**) and *C. salinarius* (**b**). The dashed lines shows the fitted values from the final three- and four-covariate models, respectively. The solid lines depicts the mean predicted values from approximately 500 predictions with error bars for the empirical 95% prediction interval

do with activity of predators. The negative effect of humidity in October may also have to do with the timing of hibernation. Most interesting is the opposite effects of total heating degree days in January compared to March through early April. A colder January may reduce the survivorship of the adults, whereas a warmer March and April may encourage early emergence from hibernation leading to larger population amplification by July.

3.3 Model evaluation

The ability of these smaller models to predict totals was assessed by the Monte Carlo scheme. An average of 500 predictions for each of the observed July totals were made based on random groups of 21 other observations. The results of these simulations can be seen in Fig. 4 along with the fitted values from the models when all observations are included. The mean of the Monte Carlo predicted totals had a R^2_{KL} of 0.947 for *Ae. sollicitans* and 0.895 for *C. salinarius*; the prediction intervals are also fairly small for most years. These results suggest that the models are not only able to explain most of the interannual differences in population dynamics of these species but might also be useful for predicting future observations.

4 Discussion

Application of the interval-lagged reduction in Poisson model deviance analysis and visualization via the cross-correlation map approach to studying the effects of off-season conditions with *Ae. sollicitans* and *C. salinarius* population size revealed relationships that might not otherwise have been discovered. Many of the crucial time intervals were on a 2–4 week scale that did not correspond to calendar months. Analysis limited to monthly or total off-season summaries would not have identified these relationships.

Although not shown here, the possibility of using earlier monthly totals to predict the July totals was also explored. In particular, the totals from the preceding September (the population presumed to be initiating the overwintering process) and from June of the same year (the population presumed to give rise to the July population) were examined as covariates for the Poisson regression models. By themselves, neither of these variables nor both together yielded a good fit for models of either *Ae. sollicitans* or *C. salinarius*. When included in the models with the off-season meteorological covariates, the September and June totals did not substantially improve the fit; this is not surprising given how good the fit was already. More surprisingly, these prior totals slightly reduced the predictive power of the models as assessed by the Monte Carlo approach. This suggests that meteorological conditions may be better indicators of future population dynamics for these mosquito species than the present size of the population.

The same techniques used here could also be used to develop models for June totals, August totals, or other seasonal summaries. Initial efforts in these directions suggest that different variables might be involved, but the same quality of model fit could be obtained. Such models could have important implications for the control of arboviruses spread by these mosquito species. By gathering readily available meteorological data for the period between October and May, one can get a first approximation of the size of vector populations that are likely to be seen. This information could be used to prioritize the distribution of scarce mosquito control resources before the transmission season begins.

Acknowledgements We thank the employees of the Maryland Department of Agriculture's Mosquito Control section who collected the mosquitoes and supplied the data. Thomas Yu did much of the data entry of the collection records. Dr. Scott Shone gathered the meteorological data. This work was supported by funding from the National Oceanic and Atmospheric Administration (NA16GP2631) and the Defense Advanced Research Projects Agency (F3060201C0184). We thank the National Institute of General Medical Sciences MIDAS program.

References

Cameron AC, Windmeijer FA (1997) An R-squared measure of goodness of fit for some common nonlinear regression models. J Econ 77:329–342

Chase JM, Knight TM (2003) Drought-induced mosquito outbreaks in wetlands. Ecol Lett 6:1017–1024

Curriero FC, Shone SM, Glass GE (2005) Cross correlation maps: A tool for visualizing and modeling time lagged associations. Vector Borne and Zoonotic Dis 5:267–275

Evans HT et al (1987) Use of surveillance results for predicting the emergence of salt marsh mosquito broods. J Fla Anti-Mosquito Assoc 58:28–33

Focks DA et al (1993a) Dynamic life table model for *Aedes aegypti* (Diptera, Culcidae) - simulation and validation. J Med Entomol 30:1018–1028

Focks DA et al (1993b) Dynamic life table model for *Aedes aegypti* (Diptera, Culicidae) - analysis of the literature and model development. J Med Entomol 30:1003–1017

McCullagh P, Nelder JA (1983) Generalized linear models. Chapman and Hall, London

Mogi M (1996) Overwintering strategies of mosquitoes (Diptera: Culicidae) on warmer islands may predict impact of global warming on Kyushu, Japan. J Med Entomol 33:438–444
Nasci RS et al (1993) Eastern equine encephalitis virus in Ohio during 1991. J Med Entomol 30:217–222
Nayar JK, Rosen L, Knight JW (1986) Experimental vertical transmission of Saint Louis encephalitis virus by Florida mosquitoes. Am J Trop Med Hyg 35:1296–1301
O'Leary DR et al (2002) West Nile virus activity – United States, 2001. MMWR 51:497–501
O'Meara GF (1992) The eastern saltmarsh mosquito *Aedes sollicitans*. Wing Beats 3:5
R Development Core Team (2004) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org
Sakamoto Y, Ishiguro M, Kitagawa G (1986) Akaike information criterion statistics. Kluwer, Dordrecht
Shone SM et al. (2006) Characterizing the population dynamics of *Aedes sollicitans* (Diptera: Culicidae) using meteorological data. J Med Entomol 43:393–402
Slaff M (1990) The biology of *Culex salinarius*. Proceedings of the New Jersey Mosquito Control Association, pp. 71–72
Turell MJ et al (2001a) Potential North American vectors of West Nile virus. Ann NY Acad Sci 951:317–324
Turell MJ et al (2001b) Vector competence of North American mosquitoes (Diptera: Culicidae) for West Nile virus. J Med Entomol 38:130–134
Wegbreit J, Reisen WK (2000) Relationships among weather, mosquito abundance, and encephalitis virus activity in California: Kern County 1990-98. J Am Mosq Control Assoc 16:22–27

Author Biographies

Andrew S. Walsh completed this work as part of his Ph.D. research in the W. Harry Feinstone Department of Molecular Microbiology and Immunology of The Johns Hopkins Bloomberg School of Public Health. After receiving undergraduate training in molecular biology, he turned his attentions to the application of statistical modeling to infectious disease ecology. He is now attempting to combine his experience with statistical methodology with his knowledge of molecular biology as a postdoctoral fellow in the Language Technologies Institute at Carnegie Mellon University.

Gregory E. Glass is a Professor in the W. Harry Feinstone Department of Molecular Microbiology and Immunology Department of The Johns Hopkins Bloomberg School of Public Health. His laboratory is involved in studies of the maintenance and transmission dynamics of infectious agents, especially zoonotic agents. He oversees both laboratory and field research of animal reservoir and arthropod vector populations as well as epidemiologic studies of affected human populations. His current focus involves developing integrated statistical spatial models for disease risk assessment.

Cyrus R. Lesser is the Chief of the Mosquito Control Section for the Maryland Department of Agriculture. He oversees projects for mosquito surveillance, source reduction, biological control initiatives, and public education for 22 counties and Baltimore City.

Frank C. Curriero is an Assistant Professor in the Department of Environmental Health Sciences and the Department of Biostatistics at The Johns Hopkins Bloomberg School of Public Health. His primary research focus is in spatial statistics and GIS applications of public health. Dr. Curriero takes a broad approach to environmental statistical method development and application. Study of the environmental influences on health is critical in a wide number of fields and research endeavors. The environment in a specific component of Dr. Curriero's work has ranged from the social and behavioral environment, the built environment, the soil, water, and air environment, the ecological environment, and the infectious disease environment.