

Spatial autocorrelation of West Nile virus vector mosquito abundance in a seasonally wet suburban environment

P. R. Trawinski · D. S. Mackay

Received: 29 January 2008 / Accepted: 29 September 2008 / Published online: 24 October 2008
© Springer-Verlag 2008

Abstract The objective of this study is to quantify and model spatial dependence in mosquito vector populations and develop predictions for unsampled locations using geostatistics. Mosquito control program trap sites are often located too far apart to detect spatial dependence but the results show that integration of spatial data over time for *Cx. pipiens-restuans* and according to meteorological conditions for *Ae. vexans* enables spatial analysis of sparse sample data. This study shows that mosquito abundance is spatially correlated and that spatial dependence differs between *Cx. pipiens-restuans* and *Ae. vexans* mosquitoes.

Keywords Geostatistics · Spatial autocorrelation · Classification and regression trees (CART) · West Nile virus · Landscape epidemiology · Mosquito vectors

JEL Classification C · C21

1 Introduction

West Nile virus (WNV), a mosquito-borne *flavivirus*, causes an infection in humans that is usually asymptomatic or a mild febrile illness but may also cause severe disease or even death (CDC 2003). The mild form is known as West Nile fever while more severe neurological forms are termed West Nile encephalitis and West Nile meningitis (Kramer et al. 2007). Its distribution encompasses Australia (subtype Kunjin), most of Africa and Europe, parts of Asia, and has recently extended into North America (Hubalek and Halouzka 2000; Rappole et al. 2000; Campbell et al.

P. R. Trawinski (✉) · D. S. Mackay
Department of Geography, State University of New York at Buffalo,
Wilkeson Quad, Buffalo, NY 14261, USA
e-mail: geopatti@gmail.com

2002). Outbreaks have occurred in Israel (1950s, 1980, 2000), South Africa (1974), Bucharest (1996–1997), Romania (1996), Russia (1999), and the United States (2002–2003) (Hubalek and Halouzka 2000; Rappole et al. 2000; Platonov et al. 2001; Weinberger et al. 2001). Since the emergence of the virus in the United States in 1999, neuroinvasive disease has been reported in 11,016 human cases, with 1,071 case fatalities (9.7% case–fatality ratio) (CDC 2008). All reported cases in the US number 27,485 but serosurveys have documented that less than 1% of people infected with the virus will develop the severe form of the disease, so estimates of actual cases are much higher (Mostashari et al. 2001; O’Leary et al. 2004). West Nile virus spreads rapidly among the bird and mosquito populations across the continental United States and now is firmly established in most areas of the United States.

There is no vaccine or other treatment specific for infection with West Nile virus. Vector control and public education are the primary means of fighting infection. Habitat elimination (“source reduction”) and larvicide application are the preferred methods of controlling nuisance mosquito populations. Aerial spraying of insecticides is generally the intervention of last resort but is used more often when mosquito populations are determined to be a public health threat. Aerial applications are generally used conservatively because of cost, potential environmental damage, and potential development of resistance (Spielman et al. 2001). Control with larvicides and adulticides requires larva and adult population surveillance and most states require these public health pesticide applicators to be certified (CDC 2003). Many municipalities and counties lack the resources to staff dedicated mosquito control personnel and certified insecticide applicators.

Better understanding of vector population structures in space is vital to improving knowledge of population dynamics and ability to control these populations (Srividya 2002; Ryan et al. 2004; Diuk-Wasser et al. 2006). Analysis of vector populations in space is based upon a simple but powerful premise by Tobler (1970) that “everything is related to everything else, but near things are more related than distant things”. The key is to find the distance below which values are more alike, or spatially autocorrelated. Detection, modeling, and estimation of spatial autocorrelation can be accomplished with geostatistics, a branch of statistics that originated from mining applications. The semivariogram is the tool used in geostatistics to find the range of autocorrelation. Small values of semivariance correspond to data that are more similar and large values reflect data that are more dissimilar. As the distance between a value and its neighbor increases, the value of its semivariance increases and should level off at some point that is equal to the conventional variance. This point where the variogram levels off indicates that the variance is now independent of distance and is termed the range. The sill is the actual plateau of the variance at that point. The discontinuity at the origin of the variogram is the nugget effect, a result of sampling error and microscale variability (Isaaks and Srivastava 1989; Cressie 1993; Bailey and Gatrell 1995; Piegorsch and Bailer 2005). The parameters of the semivariogram model provide important information about ecological systems and may be used with the data from the sample locations to produce kriged maps. Ecosystems would not be able to function without the spatial structure inherent in them. Organisms are not distributed randomly or even uniformly throughout ecosystems but are found clustered where the environment is most conducive to their survival (Legendre

1993). A hierarchy of geomorphic and meteorologic physical processes controls the spatial heterogeneity in the environment to which biologic systems respond. Within homogeneous geomorphic zones smaller scale processes act to induce further spatial structuring of the ecosystem (Legendre 1993). Geostatistics have been used to quantify and interpolate population densities of tick vectors of Lyme disease (Nicholson and Mather 1996; Estrada-Pena 1998), filariasis (Srividya 2002), and mosquitoes (Gleiser et al. 2000; Ryan et al. 2004).

The known primary vector of WNV in the northeastern US, *Culex pipiens*, is associated with human settlements and prefers water habitats with a high organic content (Turell et al. 2000; Kilpatrick et al. 2005; Turell et al. 2005). *Cx. pipiens* has similar feeding behavior to *Culex restuans* and the two species breed in similar habitats. *Cx. pipiens-restuans* combined data will be used because it is almost impossible to distinguish between the two similar species as adults (DeGaetano 2005; Diuk-Wasser et al. 2006). These two species have been grouped in previous studies (DeGaetano 2005; Kilpatrick et al. 2005; White et al. 2006; Trawinski and Mackay 2008). Another possible vector in the northeast is *Ae. vexans*, the inland floodwater mosquito, which is known to carry WNV and is a prolific human biter, allowing for transmission of the virus to humans (Hayes et al. 1985; Andreadis et al. 2004; ECDOH 2005; Turell et al. 2005). Larvae are found in practically all temporary pools of water and eggs are laid on the ground, hatching when flooded with fresh water (Hayes et al. 1985; Pratt and Moore 1993).

In this study, we use classification and regression tree (CART) analysis to classify vector abundance data into appropriate bins and geostatistical techniques to investigate the spatial structure of mosquito abundance of these bins for two West Nile virus vectors in the Town of Amherst, New York. Lacustrine silt and clay dominate this area, which is generally wet and prone to mosquito problems. As such, the Town of Amherst has a mosquito control program that is more than 30 years old. The Town of Amherst Mosquito Control Program works in conjunction with the Erie County Vector and Pest Control Program; sampling is done by the Amherst program and virus testing is provided by Erie County. Since virus testing began in 2002, 89 mosquito pools have tested positive for WNV in Amherst, of which 87 were pools of the *Cx. pipiens-restuans* complex of mosquitoes and two were *Ae. vexans* pools. Isolation of WNV from field-collected mosquitoes has been found to be a sensitive indicator of virus activity that is associated with the risk of human infection (Andreadis et al. 2004). We hypothesize that vector abundance will be spatially correlated, and that spatial autocorrelation will be higher in *Cx. pipiens-restuans* populations versus *Ae. vexans* populations because *Cx. pipiens-restuans* has a high affinity for human habitations and a very small flight range. *Ae. vexans* will travel much further for food or breeding purposes so clustering may not be as important for this species. Furthermore, we hypothesized that spatial autocorrelation would vary with time for *Cx. pipiens-restuans* and meteorological conditions for *Ae. vexans*, as these variables will affect the abundance of habitats available to each species. *Ae. vexans* prefer temporarily flooded areas and their abundance is highly dependent upon meteorological factors. *Cx. pipiens-restuans* is much less dependent upon weather conditions (Trawinski and Mackay 2008).

2 Methods

2.1 Study area: Amherst and Erie County overview

Amherst, Erie County, New York is a suburban community east of Buffalo, encompassing nearly 54 square miles (139 km²), and having a population of 126,253 (Census Bureau 2006) (Fig. 1). Its land uses are primarily residential and light industrial. Amherst is relatively flat as a result of being occupied by glacial lakes at the end of the last glaciation approximately 10,000 years ago. The surficial geology includes glacial till in the form of moraines, silt and clay deposited on the lake bottoms, and sand deposited on the beach ridges (USDA 1986). Soils range from hydric to non-hydric with a corresponding range in textures from silt to clay dominated to sand dominated. Both hydric soils and surficial geology dominated by lacustrine silt and clay occupy most of the northern portion of Amherst. In contrast, the southern section of the town is predominately variable till material and non-hydric soil, though there are many soils with the potential for hydric inclusions (USDA 1986). The main topographic feature of Amherst is the Onondaga escarpment

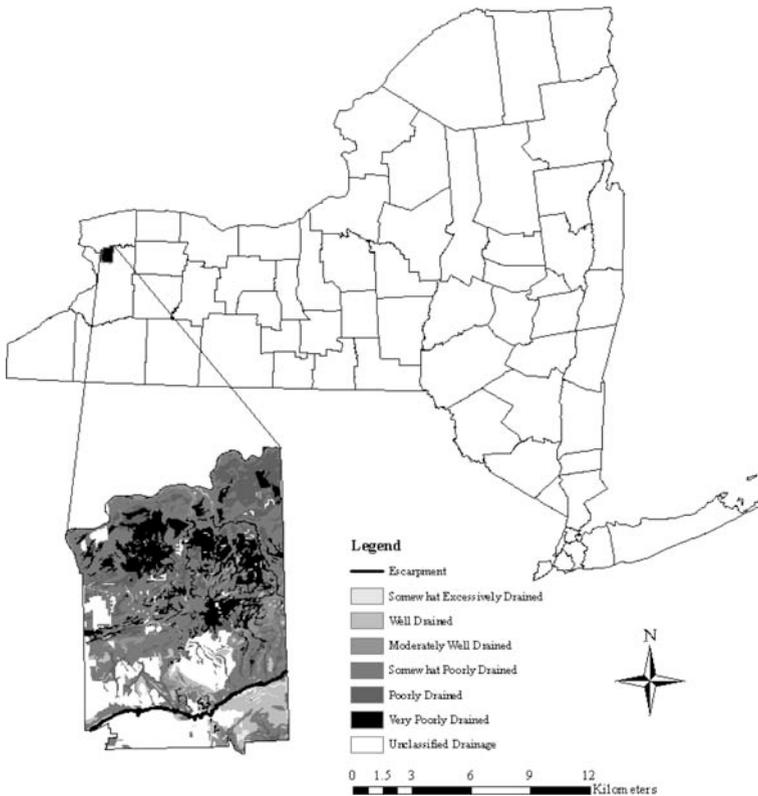


Fig. 1 Study area map of Amherst in Erie County, New York with Soil Survey Geographic drainage classes and the location of the Onondaga escarpment

escarpment that runs east–west through the southern part of town, with differences between elevation and other physical factors north and south of the escarpment. The only areas of significant slope occur along the Onondaga escarpment and Ellicott Creek. Amherst is susceptible to flooding during the late winter and early spring when snowmelt and runoff events coincide (Ebert 1988).

2.2 Mosquito vectors

Mosquito sampling by the Town of Amherst Mosquito Control Program from 2002 to 2005 utilized CDC light traps, CDC gravid traps, and “receiver” gravid traps to collect 112,551 mosquitoes. These traps often attract different species of mosquitoes. Light traps attract females seeking a blood meal while gravid traps attract females seeking to lay their eggs (“gravid females”) and are much more likely to capture *Culex* sp. mosquitoes. Generally, these two sources of data are not combined for analysis due to these differences (Trawinski and Mackay 2008). Ice trapping accounted for more than 72% of total collections in Amherst and constitutes the data used for this study.

Mosquito trapping by the Town of Amherst is not conducted on a systematic basis. Trapping occurs in brush or wooded areas throughout the town, including residential areas. Trapping may be performed in response to complaints about nuisance mosquitoes. In effect, some areas may be undersampled due to this bias in the sampling method. Initial mapping of trap locations indicated that the area south of the Onondaga escarpment has been undersampled and was therefore excluded from the spatial analysis.

Twenty-seven species of mosquito are known to inhabit Erie County (ECDOH 2005). *Cx. pipiens-restuans* and *Ae. vexans* are known vectors in the county capable of transmitting WNV to humans, horses and birds. These two species accounted for 5.31 and 36.97% of the 81,903 mosquitoes collected in CDC light traps by the Town’s Mosquito Control Division between 2002 and 2005. *Ae. vexans* is the most widespread species of *Aedes* in the United States. It is reported from every state and its abundance causes troubles in many areas (Pratt and Moore 1993). Adults of this species lay eggs on the ground in intermittently flooded areas and on the margins of water bodies. The eggs hatch after flooding and development may take from 10 to 21 days, depending on temperature. Migration of five to ten miles from their breeding places is not uncommon by adults (Pratt and Moore 1993).

Cx. pipiens-restuans, which is commonly known as the “northern house mosquito”, prefers water with a high organic content, often associated with storm water runoff, and permanent water sources and containers. These preferred sites are often associated with urbanization, and include rain barrels, tanks, tin cans, and almost any artificial container. Storm sewers, street gutters, neglected swimming pools, and effluent drains from sewage disposal are other important habitats for *Cx. pipiens-restuans*. Eggs are laid in rafts of 50–400 that hatch within a day or two in warm weather. Development time to adult stage may take as little as 8–10 days or may take several weeks during a cold spring (Pratt and Moore 1993). Adults migrate only short distances and it is usually assumed that when adults are present larvae are close (Turell et al. 2000, 2005; Kilpatrick et al. 2005).

Table 1 Descriptive statistics of log-transformed mosquito abundance data

	<i>Ae. vexans</i>	<i>Cx. pipiens-restuans</i>
Range	6.71	5.33
Minimum	0.00	0.00
Maximum	6.71	5.33
Sum	1608.12	859.90
Mean	2.07	1.11
SD	1.79	1.11
Variance	3.20	1.23
Skewness	0.44	0.88
Kurtosis	-0.79	0.19

Mosquito data from the Town of Amherst Mosquito Control Program contains information on 777 trap-nights of sampling at various times and locations throughout the season. The mean number of mosquitoes captured per trap-night is 38.8 for *Ae. vexans*, with a range from zero to 816. The mean is 5.6 for *Cx. pipiens-restuans*, with a range from zero to 206. Examination of probability–probability (P – P) plots of the original count data and the natural log of the data plus one revealed a significant improvement in normality and so the log transformed data were used for subsequent analyses (Table 1). Initial spatial analysis of all vector abundance data was problematic due to seasonal fluctuations of mosquito abundance in a temperate climate so the data were binned according to temporal and meteorological factors previously determined to be significant by time series analysis. (Trawinski and Mackay 2008).

2.3 Binning of mosquito data

Town of Amherst mosquito surveillance activities occur from mid-May to mid-October each year. Exploratory analysis of all the surveillance data from 2002 to 2005 revealed a need to bin the vector data according to time or meteorological factors in order to conduct spatial analysis. Spatial patterns of data from an entire season were masked by changing meteorological conditions while spatial patterns of data from only a single date or week in time were too sparse for spatial modeling. We used the results from a previous study on the meteorological variables significant for timing of *Ae. vexans* and *Cx. pipiens-restuans* population abundance in a regression tree analysis to determine the appropriate bins for the spatial data (Trawinski and Mackay 2008). Specifically, time-series analysis of mosquito abundance data from a location approximately 6 km from the current study area determined that there were three true meteorological predictors, accounting for 54.8% of the variation, for the abundance of *Ae. vexans*: an interactive cooling degree day base 65°F ($C_{DD_{65}}$)—precipitation variable at a lag of 2 weeks, evapotranspiration (E_T) \times E_T at a lag of 5 weeks, and $C_{DD_{65}} \times C_{DD_{65}}$ at a lag of 7 weeks (Trawinski and Mackay 2008). Three variables were significant for *Cx. pipiens-restuans* abundance: cooling degree days base 63 ($C_{DD_{63}}$) \times $C_{DD_{63}}$ and a

ponding index both at lag zero and C_{DD_63} at a lag of 8 weeks. However, these three variables only account for 16.4% of the variation in *Cx. pipiens-restuans* abundance and the model without any predictors performed nearly as well as the model with the meteorological predictors (Trawinski and Mackay 2008).

Initially, the abundance data were graphed by the week of trapping (season week) in order to explore possible categories for binning. Visual interpretation of the graphs showed distinct changes in slope for the *Cx. pipiens-restuans* abundance data that would allow classification of the data for this species. Visual interpretation of the *Ae. vexans* data showed no clear groupings so further exploration of the abundance data was performed using classification and regression trees, modern statistical techniques used to explain variation in a response variable by one or more explanatory variables. Classification trees are used for categorical response variables while regression trees are used for quantitative response variables (De'ath and Fabricius 2000; Camp and Slattery 2002). Explanatory variables can be categorical or numeric. Trees are constructed by repeatedly splitting the data into groups that become more and more homogeneous with each split. Each split is defined by a simple rule based on a single explanatory variable, i.e. less than or greater than some chosen value (De'ath and Fabricius 2000). Each group produced by a split is mutually exclusive and as homogeneous as possible. The splitting procedure is then applied again to each group separately (De'ath and Fabricius 2000). The size of the tree is equal to the number of final groups or leaves of the tree, also termed terminal nodes. Trees are generally overgrown and then pruned back as needed (De'ath and Fabricius 2000; De'ath 2002).

We constructed regression trees with the *Ae. vexans* and *Cx. pipiens-restuans* abundance data as response variables and the significant meteorological variables and the temporal variable season week as explanatory variables. The final groups were used as bins with the node rules supplying the significant cut-off values (De'ath and Fabricius 2000). The tree was pruned as necessary and sometimes expanded when the number of observations was large enough to explore the spatial variation at a finer temporal scale. Regression tree analyses were performed with Systat (version 11, Systat Software Inc. (SSI), San Jose, CA, USA).

2.4 Spatial variance of mosquito density by species

The spatial distribution of each group of mosquito trap sites was examined and locations north of the escarpment were selected for further analysis of each group. A few of the mosquito groups contained lone trap sites located at a distance (greater than 3,000 m) from the rest of the sample and were excluded from analysis because they were likely to skew the results of the spatial analysis by virtue of their isolation from other trap locations. No points were excluded from the study because their z values were exceptionally high or low (outliers). All spatial distributions were visually inspected using ArcMap version 9.2 (ESRI, Redlands, CA).

Spatial variability of *Cx. pipiens-restuans* and *Ae. vexans* abundance in Amherst was quantified using geostatistical techniques. Specifically, we used variograms to describe the spatial dependence of the *Cx. pipiens-restuans* and *Ae. vexans* mosquito densities to determine if there was spatial structure to the variance of mosquito

abundance and as input to kriging analysis for interpolating mosquito abundance at unsampled locations. A semivariogram is a descriptive measure of spatial pattern that expresses variance (γ) as a function of distance between sample locations:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{(i,j)|h_{ij}=h} (v_i - v_j)^2 \quad (1)$$

where h is the lag distance, $N(h)$ the set of all pairs of observations separated by h , and v_i and v_j are the data values at locations i and j , respectively. We obtained estimates of the nugget, the sill, and the range of the variogram and determined how much of the sample variability was truly random, how much was spatially dependent, and the range of influence over which the spatial dependency was in effect. The results of this analysis can be used to derive prediction parameters specific to the spatial variability of the data. The Amherst data sampling was not designed with the intent of spatial analysis. Practically, this made semivariogram construction more difficult. We examined the number of lag pairs at each interval for a minimum of 30 pairs for the lag sizes of 250–500 m used in this study and, when necessary, we combined lags to achieve this minimum. We further examined the credibility of the data with the construction of confidence intervals around each average semivariance value for each lag. Ninety-five percent confidence intervals for the experimental semivariograms were calculated as (Cressie 1993):

$$CI_{95} = 1.96 \frac{\sqrt{2\gamma}}{\sqrt{N}}. \quad (2)$$

All data were plotted to 6,000 m in order to ensure appropriate representation of the whole study area (Liebhold et al. 1993). The width of the study was approximately 10,500 m.

Models were evaluated with the residual sum of squares, an exact measure of the model fit to the sample data (lower is better). R -squared is also reported as a common measure of overall explanation of the model. Semivariogram modeling also allows for the estimation of spatially structured variance [$C/(C_0 + C)$], which is the distance between the nugget and the sill (C) in proportion to the total variance of the sill ($C_0 + C$). This value equals zero when there is no spatial dependence at the scale measured and increases to one as the nugget variance decreases.

We produced interpolated maps of mosquito abundance to help interpret the parameters of spatial structure estimated from fitting the semivariogram models. We used block kriging for interpolation of mosquito densities because it provides estimates for a discrete area around an interpolation point, which is more appropriate for mosquito densities than point kriging, which provide estimates for a precise point (Bailey and Gatrell 1995). Kriging is a linear interpolation that allows predictions of unknown values in the study area based on information from measurements made at sample locations (Isaaks and Srivastava 1989; Piegorsch and Bailer 2005). It produces a map in the form of a uniform grid of pixels across the spatial domain of the study area. Cross-validation analyses were performed for each kriged map with point-by-point exclusion. Geostatistical analyses were performed with GS+ (version 7, Gamma Design Software, Plainwell, MI, USA).

3 Results

Nine bins of data were developed for *Cx. pipiens-restuans* based on the regression tree and were used for semivariogram analysis (Table 2). Regression tree branches were only significant for the season week variable when analyzed with *Cx. pipiens-restuans* population abundance as the response variable and season week is the only explanatory variable used to partition the abundance data for further analysis. These results agreed with the visual interpretation of the graph of abundance by week of trapping (season week).

The *Ae. vexans* regression tree was more complex (Table 3). The first split was based on seasonal week and split the abundance data at week 33 into early summer and late summer primary branches. The early summer branch was split again by seasonal week into very early summer and mid-early summer branches at week 24. The mid-early summer branch was split into two terminal nodes based on the $C_{DD_{65}}$ -precipitation variable. The late summer primary branch was split based on the $C_{DD_{65}} \times C_{DD_{65}}$ variable—cool 7 weeks prior and warm 7 weeks prior. The warm branch was further split based on $E_T \times E_T$, resulting in a total of seven branches. The third group of the branch of the tree contained almost 40% of the observations and was further split into bins 3a–3d by season week to further explore the data temporally.

Empirical semivariograms are presented for ten bins of *Ae. vexans* based on time and meteorological factors (Fig. 2). Most of the sampling intervals had empirical semivariogram values with wide confidence intervals that did not allow modeling of a valid semivariogram. Only semivariograms from groups 1, 3b, and 7 showed spatial autocorrelation, in a range from 3,252 to 3,750 m (Table 4). All semivariograms were isotropic and best modeled by a spherical model. Plots of drift (regional trend) showed no discernible slope. *Ae. vexans* groups 1 and 7 contained the most spatially structured variance. Group 1 was comprised of mosquito sampling from very early in the season, from weeks 20 to 23, and correspondingly had the lowest mean value of all the groups (0.13). The estimated sill was 0.165 with an estimated nugget of 0.062 and range of 3,252 m. The model explained 33.2% of the variation in mosquito samples with 62.2% of the variation as spatially structured variance. Group 7 consisted of samples taken later in the summer with

Table 2 Bins of *Culex pipiens-restuans* mosquitoes based on time series and regression tree analyses

Group	Weeks	N
1	20–23	100
2	24–25	114
3	26–27	82
4	28–29	89
5	30–31	75
6	32–33	69
7	34–35	77
8	36–37	46
9	38–42	125

Confidence intervals are shown for empirical semivariograms

Table 3 Bins of *Aedes vexans* mosquitoes based on time series and regression tree analyses

Group	First Branch	Second Branch	Third Branch	Fourth Branch	N
1	Week < 33	Week < 24	–	–	100
2	Week < 33	Week ≥ 24	cdd65-precip ≥ 90.526	–	89
3	Week < 33	Week ≥ 24	cdd65-precip < 90.526	–	309
4	Week ≥ 33	CDD65 x CDD65 ≥ 729	Week < 39	ET x ET < 3.161	51
5	Week ≥ 33	CDD65 x CDD65 ≥ 729	Week < 39	ET x ET ≥ 3.161	99
6	Week ≥ 33	CDD65 x CDD65 ≥ 729	Week ≥ 39	–	50
7	Week ≥ 33	CDD65 x CDD65 < 729	–	–	79

cooler temperatures occurring 7 weeks earlier and had the highest mean of all the *Ae. vexans* groups (4.67). The estimated sill was 1.489, estimated nugget was 0.601, and range was 3,630 m. This model explained 43.1% of the variation in the mosquito sampling from this group with 59.6% being spatially structured. Estimated parameters from the semivariogram analysis are presented in Table 4. All models had residual sum of squares (RSS) values in the acceptable range, but RSS was by far the lowest for the group 1 model and, therefore, this was considered the best model of the set.

Empirical semivariograms for nine groupings of *Cx. pipiens-restuans* based on time are presented in Fig. 3. Seven of the groups showed significant spatial autocorrelation in the range of 1,971–6,540 m, and the model RSSs were low for all groups (Table 5). With the exception of groups 2 and 3, all groups showed evidence of spatial correlation distinguishable from pure nugget. *Cx. pipiens-restuans* group 6 contained the least spatially structured variance and group 8 contained the most. Group 6, containing samples from weeks 32 to 33, had a mean of 1.660, an estimated sill of 1.303, an estimated nugget of 0.980, and a range of 5,200 m. The model accounted for 21.2% of the variation in the sample, spatially structured variance was low at 24.8%. Group 8, consisting of weeks 36–38, had a mean of 0.972, an estimated sill of 0.995, an estimated nugget of 0.219, and a range of 5,840 m. The spatially structured variance was high at 78.0%. This model explained 68.6% of the variation in mosquito samples. All semivariograms were isotropic and best modeled by a spherical equation. Plots of drift showed no discernible slope. Random variations and measurement errors were manifested by oscillations in gamma along the sill but very few points were outside the confidence intervals. A graph of the variation in range for all groups of *Cx. pipiens-restuans* is presented in Fig. 4.

The maps in Figs. 5 and 6 show the interpolated mosquito abundances for *Ae. vexans* and *Cx. pipiens-restuans*. Kriged maps that incorporate both the model of spatial variance derived from observations and the individual measurements often aid in the interpretation of the variograms. Regression coefficients for the cross-validation of each kriged map indicate how well the model performs at estimating a sample point from the rest of the data, excluding one sample point at a time. The regression coefficient represents a measure of the goodness of fit for the least-squares model describing the linear regression equation. Regression coefficients for each group of mosquito abundance are presented in Table 6.

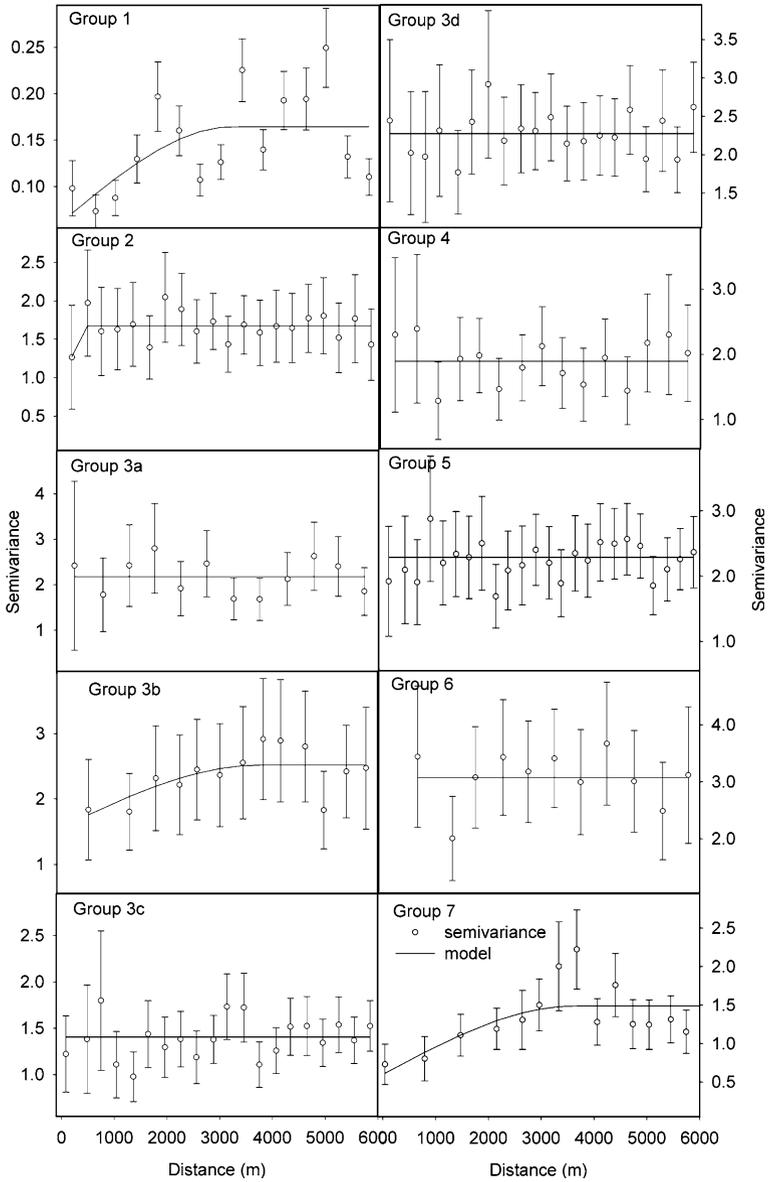


Fig. 2 Empirical and modeled semivariograms for *Aedes vexans*

4 Discussion

As hypothesized, evidence of spatial dependence was found in both species of vector mosquitoes analyzed; however, more spatial dependence was evident in *Cx. pipiens-restuans* populations than in *Ae. vexans* populations. Semivariograms for

Table 4 *Ae. vexans* semivariogram model parameters for groups based on time and weather variables

Group	Model	C_0	$C_0 + C$	A	Residual SS	R^2	$C/C_0 + C$	N	Mean
1	Spherical	0.062	0.165	3,252	0.026	0.332	0.622	99	0.13
2	Nugget	1.707	1.707	–	–	–	–	83	1.12
3a	Nugget	2.652	2.652	–	–	–	–	55	1.72
3b	Spherical	1.569	2.519	3,750	0.964	0.446	0.377	54	1.39
3c	Nugget	1.405	1.405	–	–	–	–	97	2.26
3d	Nugget	2.274	2.274	–	–	–	–	87	2.29
4	Nugget	1.893	1.893	–	–	–	–	78	1.81
5	Nugget	2.285	1.310	–	–	–	–	73	2.03
6	Nugget	3.074	3.074	–	–	–	–	47	3.63
7	Spherical	0.601	1.489	3,630	1.260	0.431	0.596	72	4.67

The nugget is represented by C_0 , the sill is represented by $C_0 + C$, the range is represented by A, and the spatially structured variance is represented by $C/C_0 + C$

Ae. vexans indicated that many of the samples were pure nugget, and the nuggets were relatively large with wide confidence intervals, implying that mosquito abundance was not spatially correlated at the scales measured for these mosquito groups (Fig. 2). Spatial autocorrelation of *Ae. vexans* abundance data did not show a significant change during periods of detectable spatial dependence. The range only varied from 3,250 to 3,750 m for the three valid models representing groups with different meteorological characteristics (Table 4). Our data support the concept that the placement of physical features, such as wetlands and other areas prone to ponding water, is important in the life history of *Aedes vexans*. More testing is needed to determine if this range of spatial dependence occurs for all *Ae. vexans* populations in varied landscapes.

Spatial autocorrelation of *Cx. pipiens-restuans* was not detected in the same range but was shown to vary with groups based on time, as is illustrated in Fig. 4. Most groups of *Cx. pipiens-restuans* exhibited spatial autocorrelation in the 5,000–6,500 m range, but spatial dependence was only evident to about 2,000 m for a brief period (Table 5). At the beginning of the mosquito season in Amherst, the range of spatial dependence was about 5,300 m. During the next 4 weeks (24–27) spatial dependence was not detectable at the scale examined. Spatial dependence was detectable again in weeks 28–29 at a much smaller range of about 1,970 m and only increased slightly through week 31–2,200 m. The range of spatial dependence then increased sharply to 5,200 m, back to the range found very early in the season, and then gradually increased late in the season to over 6,500 m. This pattern of change in spatial dependence over time supports our hypothesis that spatial autocorrelation varies temporally for *Cx. pipiens-restuans* and may be due to variations of breeding sites and adult mosquito habitats with meteorological conditions. However, a range between 5,000 and 6,000 m was encountered often and may be significant for this species.

In general, the regression coefficients for *Cx. pipiens-restuans* estimates are much better than those for *Ae. vexans* (Table 6). The interpolated estimates of *Cx. pipiens-restuans* are best for groups 7, 8, and 9, with regression coefficients ranging

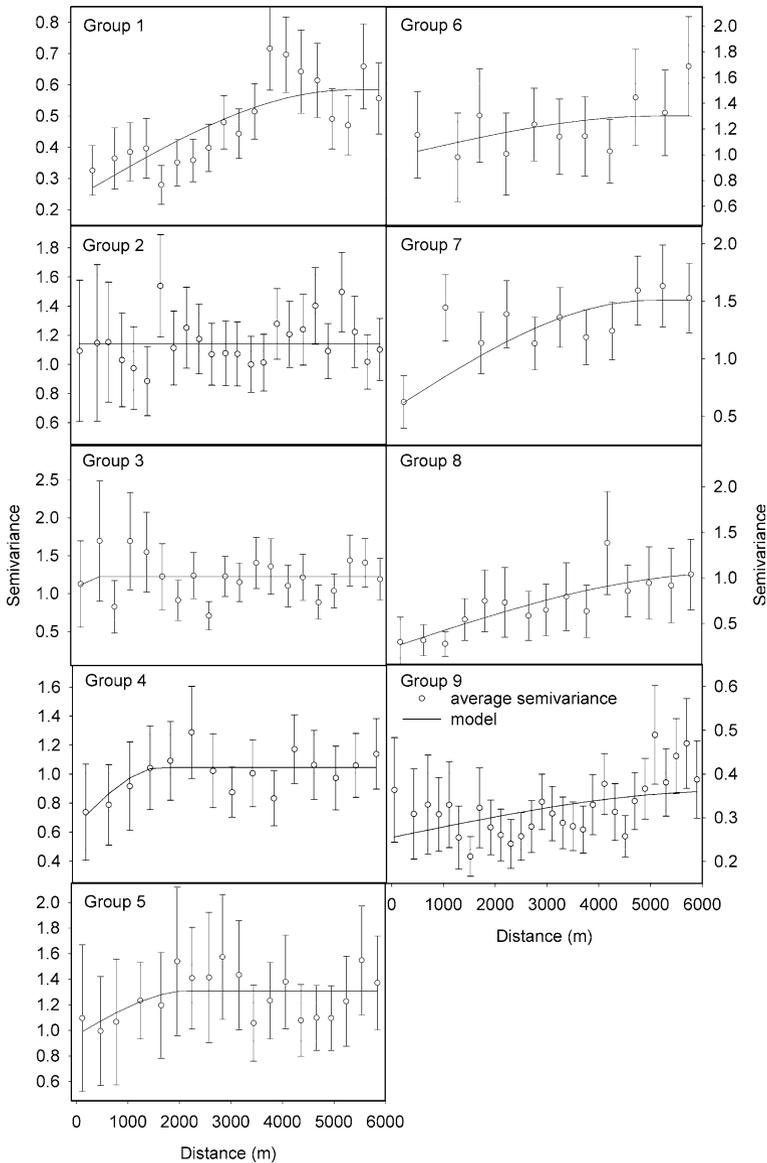


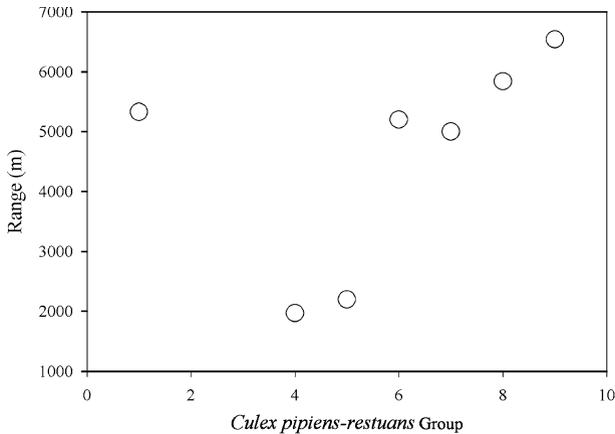
Fig. 3 Empirical and modeled semivariograms for *Culex pipiens-restuans*. Confidence intervals are given for empirical semivariograms

between 0.718 and 0.875. *Cx. pipiens-restuans* is the most important mosquito vector of WNV in the northeastern US and human cases are most likely to occur late in the mosquito season, with onset in late August and September. Because the best models of *Cx. pipiens-restuans* are the late season models, they may contribute significantly to control efforts when an outbreak occurs.

Table 5 *Culex pipiens-restuans* semivariogram model parameters for temporal groups 1–9

Group	Model	C_0	$C_0 + C$	A	Residual SS	R^2	$C/C_0 + C$	N	Mean
1	Spherical	0.239	0.585	5,330	0.128	0.599	0.591	99	0.51
2	Nugget	1.154	1.154	–	–	–	–	114	1.38
3	Nugget	1.224	1.224	–	–	–	–	81	1.31
4	Spherical	0.657	1.048	1,971	0.173	0.453	0.373	78	1.37
5	Spherical	0.963	1.310	2,200	0.477	0.243	0.265	73	1.45
6	Spherical	0.980	1.303	5,200	0.351	0.212	0.248	64	1.66
7	Spherical	0.552	1.509	5,000	0.579	0.465	0.634	72	1.61
8	Spherical	0.219	0.995	5,840	0.395	0.686	0.780	43	0.97
9	Spherical	0.225	0.361	6,540	0.098	0.213	0.294	121	0.40

The nugget is represented by C_0 , the sill is represented by $C_0 + C$, the range is represented by A, and the spatially structured variance is represented by $C/C_0 + C$

**Fig. 4** Graph of change in range for *Culex pipiens-restuans*

Insect populations in general are spatially heterogeneous in their densities, which is important for devising optimal sampling procedures, understanding predator–prey relationships and intraspecific competition, and in development of rational pest management strategies (Liebhold et al. 1993; Ryan et al. 2004). Furthermore, the spatial heterogeneity found in ecosystems is functional—ecosystems would not operate properly without spatial continuity of habitats (Legendre 1993). The characteristics of the geostatistical models provide insight on the spatial patterns of the landscape. Legendre (1993) proposed that a nugget-to-sill ratio of 0.4 or less may indicate biological significance in spatial pattern. Much of the effort in characterizing spatial patterns of insect densities has emphasized descriptions of spatial patterns with dispersion indices such as patchiness indices. Such methods focus on frequency distributions of samples and ignore sample locations. The result

Table 6 Regression coefficients from cross-validation of interpolated mosquito abundances

Mosquito Group	Regression coefficient
<i>Culex pipien-restuans</i>	
Group 1	0.422
Group 4	0.709
Group 5	0.250
Group 6	0.541
Group 7	0.875
Group 8	0.718
Group 9	0.771
<i>Aedes vexans</i>	
Group 1	0.010
Group 3b	0.297
Group 7	0.206

is often a failure to differentiate spatial patterns and descriptions of spatial patterns that are highly dependent on the size of the sample units (Liebhold et al. 1993). Spatial analysis of mosquito and other insect populations are better served by the use of geostatistics, in which both value and location are taken into account. However, geostatistical techniques have been under-utilized in the analysis of disease vectors in general, and specifically for mosquito vectors of disease. Geostatistical analyses of tick vectors, primarily of Lyme disease, have been performed more frequently than analyses of mosquito vector populations (Nicholson and Mather 1996; Estrada-Peña 1998). Recently, a few studies have begun to explore spatial autocorrelation in mosquito population distributions. Ryan et al. (2004) used Moran's I indices and kriged maps to illustrate spatial dependence in three of four species of mosquitoes studied with the strongest spatial autocorrelation occurring at intervals between 0 and 1.5 km. Gleiser et al. (2000) found a range of 10 km with semivariogram analysis for *Aedes albifasciatus* in Central Argentina. No evidence of spatial autocorrelation was found for *Cx. pipiens*, *Cx. restuans*, and *Ae. vexans* at trap sites averaging 4.7 km apart in Fairfield and New Haven Counties, Connecticut (Nicholson and Mather 1996; Estrada-Peña 1998; Diuk-Wasser et al. 2006). Our results indicate that the range of spatial autocorrelation varies over time for *Cx. pipiens-restuans*, occurring up to a range of 6.5 km in the late season. Many mosquito control programs conducted on a state-wide or county-wide basis perform sampling at sparse locations that are many kilometers apart. The sample spacing of these studies is often too large to detect spatial autocorrelation in the sample (Diuk-Wasser et al. 2006). However, our results indicate that sparse mosquito trapping data that are collected over long periods of time can be successfully categorized into groups based on time and meteorological conditions and analyzed with geostatistical techniques. These results should encourage spatial analysis of abundance data by other researchers with mosquito sampling data that are sparse in an instantaneous sense through the appropriate integration of the data over time. This integration of mosquito abundance over time may be especially useful for spatial analysis of late season *Cx. pipien-restuans* populations. However,

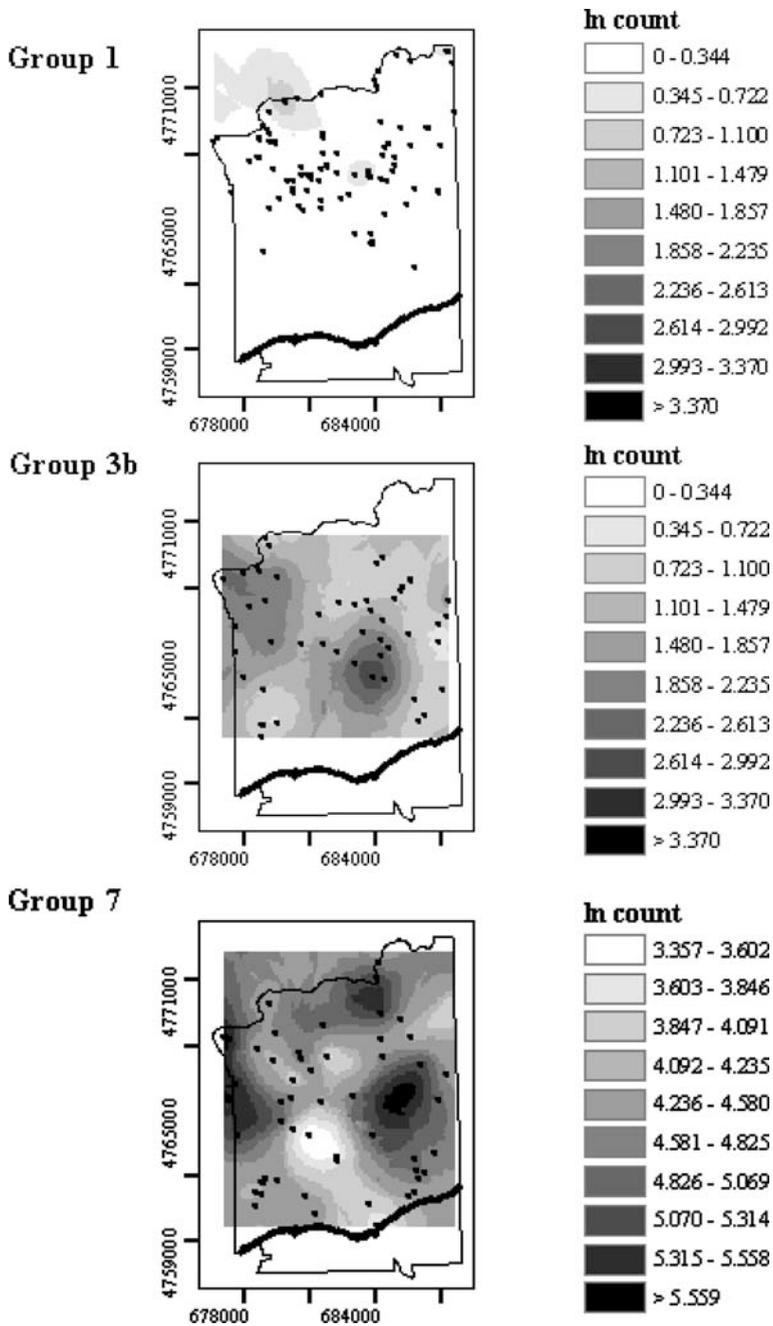


Fig. 5 Kriged maps of *Aedes vexans* groups 1, 3b, and 7

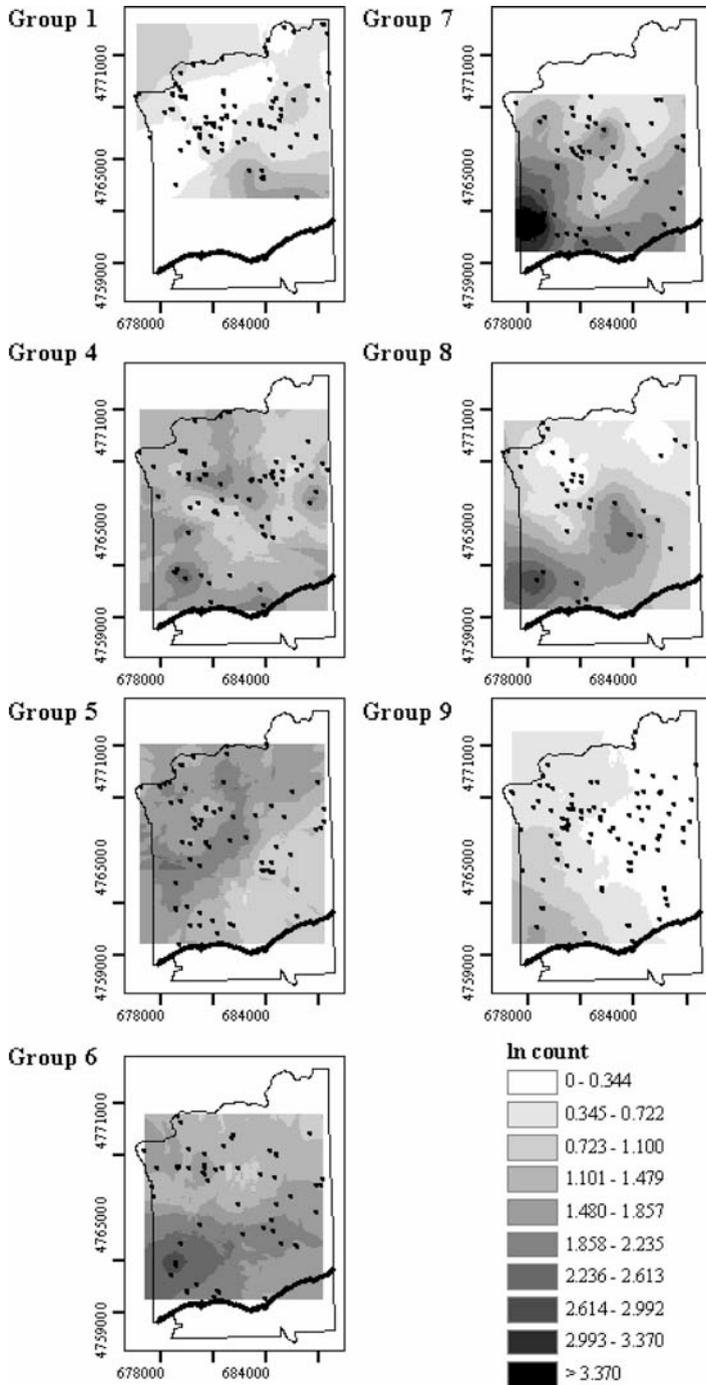


Fig. 6 Kriged maps of *Culex pipies-restuans* groups 1, 4, 5, 6, 7, 8, and 9

these bins may need to be re-examined; as the amount of data available for time series analysis of *Cx. pipiens-restuans* and *Ae. vexans* increases, drift in the bins of the vectors may occur between years.

The spatial relationships elucidated in this research allow for recommendations in future mosquito trapping designs. The confidence intervals of the empirical semivariogram values are relatively large, especially at the smaller lag intervals, and impeded identification of valid semivariogram models for nine mosquito groups. Further information on the spatial dependence of vector mosquitoes may be derived from more sampling at smaller lag intervals between 200 and 2,000 m. Systematic sampling at these distances will allow for more robust descriptions of spatial dependence. Systematic sampling would remove the bias in the sample data that occurs because traps are set in areas suspected to be problematic. This bias may actually lead to over-estimation of mosquito abundance in areas that are not sampled. Also, sampling over an area greater than the Town of Amherst would facilitate the modeling of spatial dependence, especially in *Cx. pipiens-restuans* populations, some of which have a range of spatial dependence estimated at over 6,000 m. Increasing the width of the study area will allow better estimations of the range and the sill in these populations.

A best mosquito sampling scheme might include a combination of targeted areas of cyclic sampling with random sampling throughout the study area. Unrestricted access to all areas in an urban environment is unreasonable and a comparison of cyclic, random, and uniform sampling designs to determine the optimal sampling scheme for extracting spatial information from the abundance data is not feasible. However, studies have indicated that cyclic sampling designs increase sampling efficiency, so that samples are distributed more efficiently for geostatistical analyses and that regular grid spacing requires many more points for spatial analysis than either random or cyclic sampling. It is harder to achieve narrow confidence intervals with random sampling (Burrows et al. 2002; Adelman et al. 2008; Loranty et al. 2008). Inaccessibility to certain areas, land use of the area designated to be sampled, and a limit to the available resources for trapping may hinder mosquito trapping efforts designed for maximizing information on spatial distribution.

Although geostatistics provide better estimates for unsampled locations than techniques which ignore space, several problems with kriging exist. Kriging is useful for interpolation within the study area but not for extrapolating outside the study area. Moreover, kriging must be used with caution on a coarse spatial scale. Kriging is ideal for continuous spatial processes, but vector abundance data is usually count data (Das et al. 2002). Also, kriging does not use the relationships of the organism to its environment in its estimations but only uses the idea that locations close together seem to be more similar (Kitron 1998). Furthermore, kriging techniques for mosquito populations are good for static data collected in one time period but do not include any reference to the dynamic state of population data. Analytical techniques that combine spatial statistics with methods for analyzing changes over time should be developed to provide a comprehensive analysis of the landscape epidemiology of vector-borne diseases such as West Nile virus.

Despite many calls to incorporate spatial analysis in the study of infectious disease and the vectors that transmit such diseases to humans, analysis of space has largely

been neglected in such studies (Liebhold et al. 1993; Graham 2004; Ostfeld et al. 2005). Especially as it pertains to mosquito data, spatial analysis based on trap collection data can assist in determining base population levels of disease vectors. Global climate change will likely change the distributions and abundances of such vectors, but without baseline abundance data these changes can not be documented (Epstein 2001). Furthermore, increased global transportation of people and goods potentially will allow more infectious agents to establish in non-endemic environments, in locales where the population is especially susceptible (Petersen and Marfin 2005). The use of spatial models to generate risk maps can assist in the formulation of plans to control and manage the ever-changing distributions of infectious diseases.

5 Conclusions

The finer spatial resolution of sample points achieved through appropriate grouping of the data used for this study allowed for detection of spatial dependence in *Ae. vexans* and *Cx. pipiens-restuans* populations. Low spatial resolution often hinders geostatistical analysis of mosquito control data but our results show that other studies may successfully apply a similar grouping to mosquito abundance data and enable spatial analysis of such data. Spatial dependence was conditional mainly for *Ae. vexans* data, in which only 30% of groupings exhibited evidence of spatial dependence as compared with almost 80% of *Cx. pipiens-restuans* groups. The best models of spatial variation were developed for late-season *Cx. pipiens-restuans* and may contribute significantly to control efforts of this important vector species. A limitation of this study is the lack of consideration for the known breeding sites and habitats of these vector species. Future spatial modeling efforts should include the preferred breeding sites and adult habitats, especially including human population density in the modeling of *Cx. pipiens-restuans* abundance.

Acknowledgments The authors would like to thank Paula Thompson and Dave Muck from the Town of Amherst Highway Department, Mosquito Control Program, and Peter Tripi, Glenn Robert, and John Eiss of the Erie County Health Department, Vector Control Program, for providing the mosquito trapping data. We also thank Dr. Peter Rogerson, Dr. Ling Bian, and the anonymous reviewers for helpful comments.

References

- Adelman JD, Ewers BE, Mackay DS (2008) Use of temporal patterns in vapor pressure deficit to explain spatial autocorrelation dynamics in tree transpiration. *Tree Physiol* 28:647–658
- Andreadis TG, Anderson JF, Vossbrinck CR, Main AJ (2004) Epidemiology of West Nile virus in Connecticut: a five-year analysis of mosquito data 1999–2003. *Vector Borne Zoonotic Dis* 4(4):360–378. doi:10.1089/vbz.2004.4.360
- Bailey TC, Gatrell AC (1995) *Interactive spatial data analysis*. Longman, Harlow
- Burrows SN, Gower ST, Clayton MK, Mackay DS, Ahl DE, Norman JM, Diak G (2002) Application of geostatistics to characterize leaf area index (LAI) from flux tower to landscape scales using a cyclic sampling design. *Ecosystems* 5:667–679
- Camp NJ, Slattery ML (2002) Classification tree analysis: a statistical tool to investigate risk factor interactions with an example for colon cancer (United States). *Cancer Causes Cont* 13(9):813–823. doi:10.1023/A:1020611416907

- Campbell GL, Marfin AA, Lanciotti RS, Gubler DJ (2002) West Nile virus. *Lancet Infect Dis* 2:519–529. doi:10.1016/S1473-3099(02)00368-7
- CDC (2003) Epidemic/epizootic West Nile virus in the United States: Guidelines for surveillance, prevention, and control, centers for disease control, US Department of Health and Human Services
- CDC (2008) Statistics, surveillance, and control. Retrieved 23 January 2008, from <http://www.cdc.gov/ncidod/dvbid/westnile/surv&control.htm>
- Census Bureau US (2006) American Community Survey. Retrieved 17 June 2008, from http://factfinder.census.gov/servlet/ACSSAFFacts?_event=Search&geo_id=&_geoContext=&_street=&_county=Amherst&_cityTown=Amherst&_state=04000US36&_zip=&_lang=en&_sse=on&pctxt=fph&pgsl=010
- Cressie NAC (1993) Statistics for spatial data. Wiley, New York
- Das A, Lele SR, Glass GE, Shields TM, Patz J (2002) Modelling a discrete spatial response using generalized linear mixed models: application to Lyme disease vectors. *Int J Geogr Inf Sci* 16:151–166. doi:10.1080/13658810110099134
- De'ath G (2002) Multivariate regression trees: a new technique for modeling species–environment relationships. *Ecology* 83(4):1105–1117
- De'ath G, Fabricius KE (2000) Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology* 81(11):3178–3192
- DeGaetano AT (2005) Meteorological effects on adult mosquito (*Culex*) populations in metropolitan New Jersey. *Int J Biometeorol* 49(5):345–353. doi:10.1007/s00484-004-0242-2
- Diuk-Wasser MA, Brown HE, Andreadis TG, Fish D (2006) Modeling the spatial distribution of mosquito vectors for West Nile virus in Connecticut, USA. *Vector Borne Zoonotic Dis* 6(3):283–295. doi:10.1089/vbz.2006.6.283
- Ebert CHV (1988) Disasters: violence of nature and threats by man. Dubuque, Kendall/Hunt
- ECDOH (2005) West Nile virus and other arbovirus five year surveillance report: 2000–2004. Erie County Department of Health, Erie County New York
- Epstein PR (2001) Climate change and emerging infectious diseases. *Microbes Infect* 3(9):747–754. doi:10.1016/S1286-4579(01)01429-0
- Estrada-Peña A (1998) Geostatistics and remote sensing as predictive tools of tick distribution: a cokriging system to estimate *Ixodes scapularis* (Acari: Ixodidae) habitat suitability in the United States and Canada from Advanced Very High Resolution Radiometer satellite imagery. *J Med Entomol* 35(6):989–995
- Gleiser RM, Gorla DE, Schelotto G (2000) Population dynamics of *Aedes albifasciatus* (Diptera: Culicidae) South of Mar Chiquita Lake, Central Argentina. *J Med Entomol* 37(1):21–26
- Graham AJ (2004) Spatial analysis for epidemiology. *Acta Trop* 91:219–225. doi:10.1016/j.actatropica.2004.05.001
- Hayes RO, Maxwell EL, Mitchell CJ, Woodzick TL (1985) Detection, identification, and classification of mosquito larval habitats using remote sensing scanners in earth-orbiting satellites. *Bull World Health Organ* 63:361–374
- Hubalek Z, Halouzka J (2000) West Nile fever—a reemerging mosquito-borne viral disease in Europe. *Emerg Infect Dis* 5:643–650
- Isaaks EH, Srivastava RM (1989) Applied geostatistics. Oxford University Press, New York
- Kilpatrick AM, Kramer LD, Campbell SR, Alleyne EO, Dobson AP, Daszak P (2005) West Nile virus risk assessment and the bridge vector paradigm. *Emerg Infect Dis* 11(3):425–429
- Kitron U (1998) Landscape ecology and epidemiology of vector-borne diseases: tools for spatial analysis. *J Med Entomol* 35(4):435–445
- Kramer LD, Li J, Pei-Yong S (2007) West Nile virus. *Lancet Neurol* 6:171–181. doi:10.1016/S1474-4422(07)70030-3
- Legendre P (1993) Spatial autocorrelation: trouble or new paradigm? *Ecology* 74(6):1659–1673. doi:10.2307/1939924
- Liebold AM, Rossi RE, Kemp WP (1993) Geostatistics and geographic information systems in applied insect ecology. *Annu Rev Entomol* 38(1):303–327. doi:10.1146/annurev.en.38.010193.001511
- Loranty MM, Mackay DS, Ewers BE, Adelman JD, Kruger EL (2008) Environmental drivers of spatial variation in whole-tree transpiration in an aspen-dominated upland-to-wetland forest gradient. *Water Resour Res* 44
- Mostashari F, Bunning ML, Kitsutani PT, Singer DA, Nash D, Cooper MJ, Katz N, Liljebjelke KA, Biggerstaff BJ, Fine AD (2001) Epidemic West Nile encephalitis, New York, 1999: results of a

- household-based seroepidemiological survey. *Lancet* 358(9278):261–264. doi:[10.1016/S0140-6736\(01\)05480-0](https://doi.org/10.1016/S0140-6736(01)05480-0)
- Nicholson MC, Mather TN (1996) Evaluating Lyme disease risks using geographic information systems and geospatial analytical methods. *J Med Entomol* 33(5):711–720
- O’Leary DR, Marfin AA, Montgomery SP, Kipp AM, Lehman JA, Biggerstaff BJ, Elko VL, Collins PD, Jones JE, Campbell GL (2004) The epidemic of West Nile virus in the United States, 2002. *Vector Borne Zoonotic Dis* 4(1):61–70. doi:[10.1089/153036604773083004](https://doi.org/10.1089/153036604773083004)
- Ostfeld RS, Glass GE, Keesing F (2005) Spatial epidemiology: an emerging (or re-emerging) discipline. *Trends Ecol Evol* 20(6):328–336. doi:[10.1016/j.tree.2005.03.009](https://doi.org/10.1016/j.tree.2005.03.009)
- Petersen LR, Marfin AA (2005) Shifting epidemiology of Flaviviridae. *J Travel Med* 12(suppl 1):S3–S11
- Piegorsch WW, Bailer AJ (2005) Analyzing environmental data. Wiley, Chichester
- Platonov AE, Shipulin GA, Shipulina OY, Tyutyunnik EN, Frolochkina TI, Lanciotti RS, Yazyshina S, Platonova OV, Obukhov IL, Zhukov AN, Vengerov YY, Pokrovskii VI (2001) Outbreak of West Nile virus infection, Volgograd Region, Russia, 1999. *Emerg Infect Dis* 7:128–132
- Pratt HD, Moore CG (1993) Mosquitoes of public health importance and their control. Atlanta, Centers for Disease Control and Prevention, United States Department of Health and Human Services
- Rappole JH, Derrickson SR, Hubalek Z (2000) Migratory birds and the spread of West Nile virus in the western hemisphere. *Emerg Infect Dis* 6:319–328
- Ryan PA, Lyons SA, Alsemeest D, Thomas P, Kay BH (2004) Spatial statistical analysis of adult mosquito (Diptera: Culicidae) counts: an example using light trap data, in Redland Shire, Southeastern Queensland, Australia. *J Med Entomol* 41(6):1143–1156
- Spielman A, Pollack RJ, Kiszewski AE, Telford SR (2001) Issues in public health entomology. *Vector Borne Zoonotic Dis* 1(1):3–19. doi:[10.1089/153036601750137606](https://doi.org/10.1089/153036601750137606)
- Srividya A (2002) A geostatistical analysis of the geographic distribution of lymphatic filariasis prevalence in southern India. *Am J Trop Med Hyg* 67(5):480–489
- Tobler WR (1970) A computer movie simulating urban growth in the Detroit region. *Econ Geogr* 46:234–240. doi:[10.2307/143141](https://doi.org/10.2307/143141)
- Trawinski PR, Mackay DS (2008) Meteorologically conditioned time-series predictions of West Nile virus vector mosquitoes. *Vector Borne Zoonotic Dis* 8(4):505–522. doi:[10.1089/vbz.2007.0202](https://doi.org/10.1089/vbz.2007.0202)
- Turell MJ, Dohm DJ, Sardelis MR, O’guinn ML, Andreadis TG, Blow JA (2005) An update on the potential of North American mosquitoes (Diptera: Culicidae) to transmit West Nile Virus. *J Med Entomol* 42(1):57–62. doi:[10.1603/0022-2585\(2005\)042\[0057:AUOTPO\]2.0.CO;2](https://doi.org/10.1603/0022-2585(2005)042[0057:AUOTPO]2.0.CO;2)
- Turell MJ, O’Guinn M, Oliver J (2000) Potential for New York mosquitoes to transmit West Nile virus. *Am J Trop Med Hyg* 62(3):413–414
- USDA (1986) Soil survey of Erie County, New York. Soil Conservation Service, United States Department of Agriculture
- Weinberger M, Pitlik SD, Gandacu D, Lang R, Nassar F, David DB, Rubenstein E, Izhaki A, Mishal J, Kitzes R, Siegman-Igra Y, Giladi M, Pick N, Mendelson E, Bin H, Shohat T, Chowers MY (2001) West Nile fever outbreak, Israel, 2000: epidemiologic aspects. *Emerg Infect Dis* 7:686–691
- White BJ, Andrew DR, Mans NZ, Ohajuruka OA, Garvin MC (2006) West Nile Virus in Mosquitoes of Northern Ohio, 2003. *Am J Trop Med Hyg* 75(2):346–349. Copyright © 2006 by The American Society of Tropical Medicine and Hygiene