CHARACTERIZING THE RELATIONSHIP BETWEEN ASIAN TIGER MOSQUITO

ABUNDANCE AND HABITAT IN URBAN NEW JERSEY

by

CAROLIN FERWERDA

A Thesis submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Master of Science

Graduate Program in Geography

written under the direction of

Richard Lathrop

and approved by

____

____

____

____

New Brunswick, New Jersey

[October, 2009]

ABSTRACT OF THE THESIS

Characterizing the Relationship between Asian Tiger Mosquito Abundance and Habitat in Urban New Jersey

By CAROLIN FERWERDA

Thesis Director:
Richard Lathrop

Since its introduction to North America in 1987, the Asian tiger mosquito (*Aedes albopictus*) has spread rapidly. Due to its unique ecology and preference for container breeding sites, *Ae. albopictus* commonly inhabits urban/suburban areas and is often in close contact with humans. An aggressive pest, this mosquito species is a vector of multiple arboviruses. In order for mosquito control efforts to remain effective, control of this important vector must be guided by spatially explicit habitat models that aid in predicting mosquito outbreaks.

Using linear regression, I determined the relationship between adult *Ae. albopictus* abundance and climate, census, and land use factors in nine urban/suburban study sites in central New Jersey. Systematically collected adult counts (females and males) from July to October 2008, served as estimates of abundance. Fine-scale land use/land cover data were obtained from object-oriented classifications of 2007 CIR orthophotos in Definiens eCognition. Mosquito abundance data were tested for spatial autocorrelation via Moran's I, semivariograms, and hotspot analysis in order to reveal

consistent patterns in abundance.

Spatial pattern analysis produced little evidence of consistent spatial autocorrelation, though several sites exhibited recurring hotspots, especially in areas near residential housing and vegetation. Stepwise multiple regression was able to explain 20-25 percent of variation in *Ae. albopictus* abundance at the 'backyard' or cell level and 72-78 percent of variation in abundance at the 'neighborhood' or study site level. Meteorological variables (temperature on the trap date and precipitation), census variables (vacant housing units and population density), and more detailed land use/land cover classes (deciduous woody vegetation, rights-of-way and vacant lots) were frequently selected in all eight models, though many other independent variables were included in the individual models. The results of the spatial statistics suggest that clustering may occur at a broader extent, while the superior predictive ability of the site level models over the finer grain cell level models supports this conclusion. Future work should focus on validating these models with 2009 field data and testing whether finer grain weather and census data enhance the models' predictive ability. Given the major differences between individual county models, future studies should further explore variations in *Ae. albopictus* habitat preferences in different geographic locations.

Acknowledgements

I would like to thank all of the people who have helped and inspired me during my graduate study. I especially want to thank my advisor, Prof. Richard Lathrop, for his invaluable guidance, encouragement, and useful suggestions throughout my research and study at Rutgers University. This work would have been impossible without the help of Mercer County and Monmouth County mosquito control. I am grateful for their data collection efforts, their expert advice on all things mosquito, and their enthusiasm for collaboration. I would also like to thank Dr. Dina Fonseca, Dr. Randy Gaugler, and all of the folks at the Center for Vector Biology for their help and input, which shaped and improved my approach to this topic. Many thanks also to Dr. David Tulloch and Dr. Laura Schneider for their constructive comments on this thesis.

My deepest gratitude goes to my husband and my family for their unflagging love and encouragement. Finally, thanks be to God for my life through all tests in the past two and a half years. May your name be honored and glorified through this work.

Table of Contents

List of Tables

List of Illustrations

INTRODUCTION

In response to outbreaks of vector-borne diseases such as malaria and dengue, scientists and public health organizations have sought to develop accurate and precise models to predict mosquito abundance and disease risk. In the face of a changing global climate and new patterns of human land use, predicting vector and disease outbreaks has become increasingly complex as models must incorporate climate factors, habitat variables, and human behavior patterns. Place-based decisions are needed to optimize the geographic targeting and management of vector populations in order to control disease outbreaks. Furthermore, models must integrate species-specific, biologically relevant information as mosquito species exhibit diverse habitat preferences and, thus, their distributions may be affected by different climatic and topographic factors (Gillett 1972; Spielman and D'Antonio 2001; Kolivras 2006; Peterson 2006). Targeted control requires information that is tailored to an area's unique landscape, while also taking into consideration more general biological information about a species.

Myriad models have been developed for vectors such as *Anopheles spp.*, the primary carriers of malaria, while fewer models exist for other important, yet perhaps lesser-known, vectors such as *Aedes albopictus*. An aggressive pest, the Asian tiger mosquito (*Ae. albopictus*; ATM) is a public health concern because of its ability to transmit at least 22 arboviruses (Francy et al. 1990; Rai 1991; Moore and Mitchell 1997). Due to its unique ecology and preference for container breeding sites, this species has spread rapidly from its native range in Asia and has become a concern for mosquito control and public health officials as the introduction and presence of potential disease vectors creates health risks for humans and animals. While *Ae. albopictus* has inhabited the southern United States for over twenty years, only within the past decade has it begun to spread into the northeastern U.S., and little is known about the species' ecology or potential range within this region (CDC 2001).

For this reason, a study combining climate, land use, and census factors to model *Ae. albopictus* habitat preferences in the northeastern U.S. is timely and necessary for developing effective control plans. As a part of a pilot project for the Climate and Health Initiative at Rutgers University, I used remote sensing, spatial statistics, and linear regression to investigate the relationship between climate and land use/land cover and *Ae. albopictus* populations in urban and suburban areas of New Jersey. This project was undertaken in collaboration with scientists at the Rutgers University Center for Vector Biology. New Jersey has the highest population density in the country as well as one of the longest histories of mosquito control (Spielman and D'Antonio 2001). Given the large percentage of developed land in the state and the ATM's preference for container breeding sites in residential areas, New Jersey offers prime habitat for the species. In order for mosquito abatement efforts to remain effective, mosquito control and public health agencies require realistic models that provide predictions of the geographic locations where mosquito outbreaks are most likely to occur at fine enough spatial scales (i.e., down to the level of a city block) to effectively manage vector populations.

BACKGROUND

Since its introduction to North America in 1987, *Ae. albopictus* has spread through the southern and eastern United States, reaching New Jersey around 1995 (Crans et al. 1996; Moore and Mitchell 1997). The northern limit of this species' range is speculated to be the -5°C January isotherm based on the ability of eggs to overwinter colder temperatures (Nawrocki and Hawley 1987). As with all mosquito species, favorable climate and environmental conditions are essential at all stages of its lifecycle. *Ae. albopictus* lays its desiccation-resistant eggs in water-filled natural and artificial containers. The adults are most active during the early morning and late afternoon (Hawley 1988). The microhabitat of this species is described as urban and suburban because these areas often have a wide variety of suitable containers (e.g., gutters, bird baths, flower pots). Indeed, *Ae. albopictus* can complete its lifecycle in a container holding less than a quarter inch of water (Ibid.). In New Jersey, individuals of this species usually emerge sometime in late spring, pass through four larval stages and a pupal stage, and finally become adults around late April or early May. This cycle is estimated to be 10-14 days in the spring and 5-10 days in the summer in New Jersey (Teng and Apperson 2000; D. Fonseca, personal communication, 13 July 2009). Three to four days after emergence, female mosquitoes seek out their first blood meal and the cycle begins again.

Despite the long history of mosquito control in New Jersey, the compound effects of climate, land use, and socioeconomic factors on mosquito populations have not been quantitatively assessed. In a study of *Culex* abundance data in four New Jersey counties over the past 30 years, DeGaetano (2005) found that less than half of mosquito abundance could be explained by meteorological factors during summer months, and the most predictive meteorological factor differed between months. However, this study did not include landscape factors such as vegetation, human population density, or land use, and the temporal and spatial scales were relatively coarse (monthly and county-wide, respectively). While temperature and precipitation influence mosquito survival, habitat characteristics such as vegetation type and amount determine whether the area is suitable for protecting desiccation-vulnerable adults and providing food (nectar) for adults (Richards et al. 2006). Many studies (e.g., Linthicum et al. 1987; Shaman et al. 2002; Kolivras 2006; López-Cárdenas et al. 2005; Porphyre et al. 2005; Grieco et al. 2006; Mushinzimana et al. 2006; Benedict et al. 2007; Vanwambeke et al. 2007; Brown et al. 2008) have shown that land cover, elevation, and other habitat characteristics should be incorporated into models of mosquito distribution. Similarly, human-related variables such as land use, population density, and housing structure type greatly influence habitat characteristics and mosquito abundance and are undoubtedly related to disease risk (Rogers et al. 2002; Ostfeld et al. 2005; Kalluri et al. 2007). Therefore, including climate, land use/land cover, and census data into models of mosquito distribution and abundance is vital to generating accurate predictions of outbreaks and a species' potential range.

In addition to predictor variables, scale is also an important consideration when developing models of *Ae. albopictus* habitat (Rey et al. 2006; Brown et al. 2008). In relation to identifying habitat in the landscape, scale is characterized by grain and extent (Turner et al. 2001). Grain refers to the finest spatial resolution possible within a given dataset (i.e., minimum mapping unit), while extent refers to the size of the study area (Ibid.). This species' preference for container breeding sites requires fine grain models because coarser models may indicate correlations with residential areas (e.g., Barker et al.

2003; Braks et al. 2003) without providing information about where to target control within residential areas. Some quantitative studies of ATM have coarse grains in order to predict its potential range due to its status as an aggressive invasive species (e.g., Benedict et al. 2006), while others may have been limited by data availability. So far, there have been few fine grain spatially explicit models for this species. Richards et al. (2006) tested the relationship between general land cover classes (e.g., roads, buildings, wooded, and open areas) with a 4 m^2 minimum mapping unit and weekly ATM oviposition in suburban neighborhoods in North Carolina using linear regression, though this was not the primary focus of their study. Rey et al. (2006) used principal component analysis to correlate fine-scaled ATM larval abundance with 17 classes encompassing ground versus canopy vegetation, paved versus unpaved lots and roads, and different types of water bodies derived from aerial photos in Florida, though the temporal scale of the study was coarse (1 month intervals). Similarly, Kobayashi et al. (2002) analyzed fine grain oviposition data relative to climate factors in Japan. However, none of these studies accounted for climate, land use/land cover, and human abundance and housing factors at the same time, and most focused on oviposition or larval abundance rather than adult abundance.

This project is in collaboration with a larger project at the Rutgers University Center for Vector Biology. The purpose of the principal project is to compare the effectiveness of different control strategies to suppress *Ae. albopictus* (95% decrease in abundance) in urban areas. The 2008 field data collection was a precursor and benchmark for 2009, when ATM populations in three sites in each county will be compared based on differences in control efforts at each site (education only, education and chemical control, no intervention). Education will be implemented through local schools and public information sessions, announcements, and brochures in two of the three final sites in each county, while chemical control will include the application of larvicides and adulticides within one of the three final sites within each county. One site in each county will serve as the experimental control with no mosquito abatement or education efforts. While assessing the effectiveness of control methods, the project also aims to resolve unanswered questions about the life history and management of the ATM. Within the broader context of this project, my work will reveal variations in ATM abundance across the landscape in relation to habitat heterogeneity. This study will provide information about useful variables and methods for future work as well as help to create guidelines for how to refine sampling of ATM populations in the 2009 field season. As a multidisciplinary collaboration, this project incorporates the expertise of entomologists on the species' biology and ecology with the spatial perspective of a geographer and landscape ecologist. Since developing the most effective plan for reducing mosquito populations is one of the primary objectives of the larger project, my work will also contribute by generating spatially explicit prediction maps and hot spot maps that show where to focus control and trapping efforts. The mosquito abundance data and the expertise of the entomologists and mosquito control officials involved in the larger project have been generously provided in aid of this project.

OBJECTIVE STATEMENT AND RESEARCH QUESTIONS

The primary objective of this study is to characterize the relationship between adult ATM density and habitat characteristics in urban areas of New Jersey. I used adult abundance data (dependent variable), remote sensing methods, and multiple sources of secondary data to answer the following interrelated questions:

- What are the spatial patterns of adult ATM density in the study sites?
    - Are there specific hot spots or spatial trends? How do patterns vary over time?
- What are the relationships between adult ATM density and habitat characteristics?
    - What climate and landscape variables account for variations in ATM populations?
    - Do variables representing human presence, abundance, and dwelling type affect the distribution and abundance of ATM?
    - How applicable are models between counties?
- Is object-oriented image classification a useful technique for describing urban ATM habitat?
    - Which of the hierarchical object scales is most relevant to the adult ATM?

METHODS

Study Area and Species Abundance Data

The study areas are located in Mercer (four sites) and Monmouth Counties (five sites) (Figure 1). Each urban/suburban site contains approximately 1000 housing units and is further divided into cells of approximately 6-8 lots (Mercer) or 8-10 lots (Monmouth) (Figure 2). Monmouth County sites range in area from 100 to 180 hectares, while Mercer County sites range from 30 to 60 hectares. Sites were selected by county mosquito control agencies and the Center for Vector Biology to have minimal variability in socioeconomic aspects in order to increase comparability of the areas during the evaluation of control methods in 2009. Each week from early July to late October 2008, mosquito control randomly placed nine (in Mercer) or eleven (in Monmouth) BG-Sentinel traps (Biogents GmbH, Regensburg, Germany) per study area (81 traps per week total). Each trap was placed in a cell, which had been randomly determined with replacement allowed, and the exact location within the preselected cell was at the field operators' discretion depending on permission from homeowners. The traps attract adult day-biting mosquitoes using a non-toxic lure and convection currents generated by a small electric fan. Traps were retrieved 24 hours after placement at pre-selected locations, which were documented via photographs and GPS points. Trapped mosquitoes were then taken back to the counties' labs for identification by species and sex. Post-processing of the counts involved geocoding misplaced points (due to equipment or human error) using tax parcel data, though less than a third of the observations required this treatment. Due to concerns about the accuracy of the recorded GPS locations, cells served as the basic unit of analysis rather than individual points. This ensured that spatial error was not introduced into the analysis and it resulted in the loss of only some resolution since the cells are quite small. In cases where multiple points fell within a cell, the median served as the value for the cell. Prior to analysis, counts of total, females only, and males only within each cell were log-transformed ($\ln(x+5)$ for total and females; $\ln(x+1.5)$ for males) to achieve normality. To calculate counts for each site, cell counts were summed by site for each trapping week and then square-root transformed to achieve normality for regression analysis.

Land Use/Land Cover Classification

In the context of vector-borne diseases, land use indicates human activity on the landscape and the presence of humans near vector habitats (Vanwambeke 2007), while land cover describes the vegetation and structures that cover the landscape. Though land use and land cover map different aspects of the landscape, these layers may be combined to describe human activity and potential mosquito habitat in the landscape. I developed a hierarchical land use/land cover classification of the study sites using object-oriented methods in eCognition v. 5 (Figure 3; Definiens Imaging 2006). Classes consisted of variables that may be important to *Ae. albopictus* habitat, such as cemeteries and tree canopy.

The overall approach of the classification was to segment the image at several object levels, select multiple spectral classes for each information class, and then classify each object level using a rule-based system with rules based on image characteristics and the class hierarchy (Navulur 2007). First, the image was segmented into multiple levels using spectral and spatial homogeneity criteria. Each level of segmentation is comprised of objects, which are the entities that the user classifies (Geneletti and Gorte 2003).

eCognition employs a region-based segmentation that begins at the pixel-level and then iteratively groups pixels based on user-defined homogeneity criteria (Definiens Imaging 2006; Haralick and Shapiro 1985). This method creates hierarchical network of objects by grouping lower-level objects into larger, higher-lever objects based on spectral similarity, contrast with neighbor objects, and shape characteristics (Yan et al. 2006). The size of the objects at each segmentation level is controlled by the scale parameter, with a larger scale resulting in larger objects as greater heterogeneity becomes more acceptable (Benz et al. 2004). For each level, color (spectral information from each image band) and shape (smoothness and compactness) are weighted against one another to determine how pixels are grouped into objects.

Object-oriented classification allows hierarchical classes, which represent the landscape at multiple scales. I created a multi-level representation of land use/land cover in the study areas, from the 'backyard' (sub-object) level to broader classes at higher (object and super-object) levels (Figure 5). Specifically, super-objects (level 1) classified the landscape into basic cover types such as pavement, vegetation, and water. Object-level classes (level 2) further divided super-object classes into different types of land use and cover such as residential buildings and grass. Sub-object classes (level 3) provided the smallest grain of information, such as lawns versus parks or swimming pools versus ponds. Not all classes described in the classification scheme were present in the study sites, though all classes should be present at the county level. All object levels were created using the same weights on shape/color and compactness/smoothness (0.1 and 0.5, respectively). Spectral information was weighted more heavily than shape to create more spatially homogeneous objects. Of the shape parameters, compactness was given equal weight to smoothness in order to extract individual objects with smooth borders (Yan et al. 2006). Segmentation scales for each level varied by image subset since subsets' extents varied. A classification rule set was developed for each county and then modified slightly for each image subset to achieve the best possible classification. Different rule sets were used for each county due to variations in spectral histograms and predominant cover and the availability of ancillary data such as building footprints and edge of pavement polygons. Ancillary data including edge of pavement and building footprint shapefiles from the counties supplemented the imagery. County shapefiles listing the locations of public buildings served to identify public buildings, while commercial and residential buildings were separated based on size and context (e.g., presence of driveways or parking lots, lawns). Some manual correction of level 2 and 3 classes was necessary since the CIR imagery had difficulty picking up the differences between certain classes, such as buildings and pavement or the different grass classes. Shape and texture based rules proved especially important at levels 2 and 3 for distinguishing spectrally similar objects that belonged in different classes.

I classified 2007 leaf-off CIR orthophotos (1 foot resolution, captured March-May), though 2008 leaf-on true color orthophotos (1 meter resolution, captured in August) were used to identify tree/shrub canopy. To merge the two image data sets, woody vegetation objects were extracted from the 2008 imagery and then combined with the 2007 objects in eCognition. Coniferous versus deciduous woody vegetation was distinguished based on the 2007 imagery since leaf off imagery provides the best means to pick out coniferous vegetation. Mosaics were created in ERDAS Imagine for each county to normalize histogram values between photo tiles via band-by-band histogram matching performed by the Mosaic tool. After trial and error adjustment of color balancing and image dodging parameters in the ERDAS Mosaic tool, the best mosaic was selected for each county. The mosaics were then divided into tiles comprising the rectangular extent of a 200 meter buffer around each site to facilitate processing in eCognition, which has difficulty segmenting and classifying images larger than 5000x5000 pixels (Definiens Imaging 2006). A 200 meter buffer was created around each site in order to take into account the effects of adjacent cover types on mosquito populations in the regression models. The minimum mapping unit (MMU) of the classification was 4 ft^2, or four pixels, which is the smallest possible MMU given the resolution of the orthophotos. While other data in the models have a coarser grain, a small MMU was chosen to detect small features such as bushes and trees that may be vital to the mosquitoes' survival. The final land use/land cover information tested in the models included the percentage of each cover type in a cell and the percentage of each cover type within 200 meters of a cell to account for the effects of adjacency. I selected a 200 meter buffer since the species has a relatively short flight range (Kitron et al. 1998). Lastly, the percentages of land cover were log-transformed (ln(x+1)) to achieve normality and reduce multicollinearity for regression analysis.

Post-classification evaluation involved 1,289 accuracy assessment points with manually defined reference classes. The reference data were created using stratified random sampling and classes were assigned to each point at all three classification levels by visually interpreting the 2007 orthophotos and cross-checking on the ground where necessary. Pictometry from 2006 also served to confirm reference class assignments (Pictometry International Corp, 2004). This sampling structure ensured that points were randomly selected and that each class received at least 50 points, while more common classes received more reference points than rarer classes (Congalton 1991). However, this process did not prevent points from falling at or near the 'boundary' between two land cover types, so some mismatches between the classification and reference classes may have been a result of this issue.

Other Factors

In addition to land use/land cover, environmental variables including climate, soil drainage, distance to the ocean, property value, and census data were included in the models (Table 1). Climate information consisted of precipitation, temperature (mean, minimum, maximum), and wind direction and speed (mean and maximum) obtained from hourly and daily measurements collected at weather stations at the Trenton Airport in Mercer County and four stations in Monmouth County (NJWXNET 2008; Figure 6). The spatial resolution of the climate data is coarser than the resolution of all other variables since only one or two reliable weather stations were available near the sites in each county. Thus, the spatial resolution of meteorological data was at the county level, where all sites in a county were given the same weather values. Each meteorological variable was averaged (temperature, wind speed) or summed (precipitation) for 30 days and 14 days before a trap was set as well as the dates that the trap was in the field. The 30 and 14 day time periods roughly correspond to the ATM's lifecycle (egg to adult) length, and are thus true predictors of mosquito catch, while meteorological conditions on the catch date explain the influence of weather conditions on mosquito catch (DeGaetano 2005). Temperature and precipitation variables were also centered to reduce multicollinearity.

Likewise, soil drainage may be important because it is related to the prevalence of standing water, the water table, and humidity, which can influence mosquito survival (Day and Shaman 2008). Soil drainage data were derived from the Soil Survey Geographic (SSURGO) database. Monmouth County sites exhibited greater variation in soil type than Mercer County sites, which are relatively uniform (Figure 7). Elevation was initially considered as an additional variable, but elevation in the study sites is nearly invariant and would not provide new information, though it may be useful in state- or county-wide studies. Distance to the coast also was included in Monmouth County models because, as the ATM is a freshwater species, it may be less abundant near the coast (Hawley 1988). Distance to coast was calculated in ArcInfo 9.3 via the Near tool (ESRI, Redlands, CA, 1999-2009).

Census and housing structure variables included population density (number of persons per hectare), single parent households, renter occupied housing units, vacant housing units, and total property value. Property value was derived from 2002 tax assessor records and all other variables were derived from 2000 U.S. census block records. Total property value (land value plus improvement value) was calculated by adding the total value of all tax parcels within a cell, minus duplicates. This data was available at a finer grain than the census data because study cells were delineated based on tax parcel boundaries. Census data were aggregated to the site level since census block boundaries do not coincide with study cell boundaries. The values of blocks partially overlapping the study areas were extrapolated by multiplying the census data by the proportion of the area of the block overlapped by the study area. Although this produced only approximations of census information, very few census blocks (28 out of 586) required this treatment.

Statistical Methods: Spatial Pattern Analysis

Before developing the habitat models, I analyzed the spatial patterns of adult ATM abundance in the study sites. Understanding the presence or absence of clustering in mosquito populations in the study sites will inform the comparison of the different control regimes in 2009 because the interpretation of the results will depend on whether or not spatial autocorrelation must be taken into account. Investigating the spatial structure of mosquito abundance also serves as an indicator of whether or not mosquitoes prefer certain areas over others within the sites, while analyzing spatial autocorrelation over time can help mosquito control commissions decide if abatement efforts need to be targeted to different areas depending on the time of season or if the same areas always have more mosquitoes than others.

Three different statistics were applied to test for different aspects of autocorrelation within the sites. Moran's I was used to determine the degree of global autocorrelation in the data, which describes the degree of correlation of a variable and itself as a function of spatial distance based on the idea that nearer things are more similar than distant things (Tobler 1970; Fortin and Dale 2005). Moran's I was chosen as a measure of global autocorrelation because it is easily interpreted and unambiguously summarizes clustering in each study site as a whole. Semivariograms were used to reveal the spatial extent of autocorrelation, confirm the results of Moran's I, and test the data for stationarity. Although semivariograms are primarily used for kriging and interpolation, they also provide information about the distances at which the data are most highly autocorrelated and the degree to which local random effects or measurement errors cause variability in the data, which provides more detail than Moran's I more binary indication of whether or not clustering is present. However, neither Moran's I nor semivariograms detect individual clusters, so hot spot analysis, a measure of local spatial autocorrelation, was employed to test for preferential clustering of ATM abundance in certain areas.

All spatial statistics were calculated for log-transformed counts of total adults, females only, and males only in each month, the high season (July-August 30), and the entire season (July-October 30) in each site. The high season was tested separately because populations reach a peak and begin declining by late August (Sota et al. 1992; Barker et al. 2003). Testing autocorrelation and developing models for this period avoids the effects of seasonality and represents the time period in which control is most important. Transformed values were used in autocorrelation testing to achieve normality, which helps to avoid bias and improve stability in spatial statistics such as Moran's I and semivariograms (Hohn 1998; Fortin and Dale 2005). Sites were inspected individually for autocorrelation since only two of the sites are spatially contiguous. Moran's I and hot spot analyses were conducted in ArcGIS 9.3 via the Spatial Statistics toolbox, while semivariograms were generated in SAS 9.2.

Moran's I tests for global autocorrelation by calculating an index of covariation between different points, varying between -1 to +1 for negatively and positively autocorrelated data, respectively. The Z score (standard deviation) indicates the significance of the I statistics, or whether the difference between the predicted and actual value of I is greater than would be expected by chance. Normalized Z-scores were interpreted to represent clusters ($Z \geq 1.65$ or $Z \leq -1.65$) for I values greater than zero or random distribution ($-1.65 < Z < 1.65$) at statistical significance $P < 0.1$. For example, an I value of 0.1 with a Z score of 1.65 would indicate clustering with a 90% likelihood that the pattern was actually random, while an I value of -0.1 and a Z score of -1.97 would indicate dispersion with a 95% likelihood that the pattern was actually random. Moran's I was calculated with a Euclidean distance method and an inverse distance conceptualization of spatial relationships. Row standardization helped to account for the effects of the sampling structure on measures of autocorrelation.

Semivariance measures the degree of spatial dependence between data points, and it is represented as a graph, or semivariogram, that shows the variance plotted against the distance between all pairs of data points (Bailey and Gatrell 1995; Figure 8). Semivariograms are often used for interpolating points via kriging, but for this project the semivariograms served to provide information about the spatial structure of the data, including the distance at which the data is the most highly autocorrelated, also known as the range. The graphs are in the geographic units of the dataset, and only the first two-thirds of the line should be interpreted (Fortin and Dale 2005). In addition to showing the distance up to which the spatial structure of the data varies, semivariograms express the degree of variability due to local random effects or measurement errors, also known as the nugget (see Figure 8 for an example). Interpretation of the graph can also reveal whether the sampling unit is appropriate for capturing spatial variability in the data (*Ibid.*). Most importantly, semivariograms can reveal the distances at which the data are clustered, which is important information for mosquito control commissions because they need to know how far from a point source of high abundance they need to focus treatment and control. That is, do they need to spray a large area to control a problem spot or can they simply focus on the point of concern.

Hot spot analysis consisted of calculating the Getis-Ord Gi* statistic, which detects concentrations of high values and low values (i.e., extreme spatial autocorrelation), where a Z score > 1.96 is 'very hot' and a Z score < -1.96 is 'very cold' at a significance level of $P < 0.05$. A random distribution is defined as $-1.96 \leq Z \leq 1.96$. First-order clusters represent groups of points that are closer together than the threshold distance and in which there is at least the minimum number of points specified by the user (Levine 2004). In this project, the threshold distance, or search radius, was chosen to maximize spatial autocorrelation and to ensure that at least eight points were considered when calculating the statistic. Visually assessing the similarities of areas with consistently more mosquitoes hints at whether the mosquitoes prefer certain habitats or whether they blanket the study areas evenly. Comparing different time periods (e.g., high season versus entire season) and study sites can reveal spatial or temporal trends in the data and possible population hot spots that are constant over time, which identify important areas to target abatement.

Statistical Methods: Multiple Regression

Using stepwise multiple regression (PROC REG, SAS Institute, Cary, NC, 2007), I tested the relationship between ATM abundance (dependent variable) and independent habitat variables. Multiple regression was selected for the analysis because the results are easily interpreted and past studies have demonstrated that regression analysis is a robust approach for explaining the strength and type of associations between environmental variables and mosquito abundance (Hay and Lennon 1999; Richards et al. 2002; Rogers et al. 2002; DeGaetano 2005; Ostfeld et al. 2005; Mushinzimana et al. 2006; Rey et al. 2006; Vanwambeke et al. 2007; Brown et al. 2008; Tran et al. 2008). Additionally, habitat data can be plugged into the regression models to generate prediction maps of ATM abundance, which could be a very valuable tool for mosquito control. In this study, habitat variables included census variables, soil drainage, property tax value, meteorological variables, and the percent of each land cover type within a cell and within 200m of a cell (Table 1). Stepwise regression first selected the most important independent variables (where $P \leq 0.15$ was the threshold for variable inclusion), and then additional manual backward elimination removed redundant or 'noise' variables to produce the final models. Manual backward elimination consisted of iteratively removing non-significant ($P > 0.05$) independent variables and observing the effect on the error rate of the model. Variables contributing significantly to the regression models were determined by significant *t*-tests for β coefficients where $P \leq 0.05$. Variables were excluded from the final model if the *t*-test was non-significant and if the standard error of the estimate did not significantly increase when the variable was removed from the model. Stepwise methods will not necessarily produce the best model if there are redundant predictors. However, after examining correlation matrices for the independent variables and removing extraneous variables (correlated at 0.70 or greater), I further ensured no redundancy by only accepting models with condition indices less than 30 and variables with variance inflation factors less than four, which are common multicollinearity tests (Belsley et al. 1980; Allison 1999).

Three sets of models were developed (see also Table 2 for a summary):

- Model 1 included ATM counts in each cell from both counties during each week of the high season (July 15-August 30). The ATM count in each cell during each week of the season was treated as an independent observation (497 observations total). Independent variables included land cover, weather, census, and soil data, as well as a dummy variable for county. These models will be referred to as ‘cell level’ models throughout the text.

- Model 2 was comprised of unique models for each county using ATM counts in each cell during each week of the high season (July 15-August 30) as the dependent variable. The ATM count in each cell during each week of the season was treated as an independent observation (218 observations in Mercer, 279 observations in Monmouth). Independent variables included land cover, weather, and census data, plus soil data and distance to coast in the Monmouth model. These models will be referred to as ‘county cell level’ models throughout the text.
- Model 3 consisted of aggregated ATM counts in each site from both counties during each week of the entire season (July 15-October 30). The ATM count in each site during each week of the season was treated as an independent observation (149 observations total). Independent variables included census and weather data only, as well as a dummy variable for county. These models will be referred to as ‘site level’ models throughout the text.

For cell level analyses (models 1 and 2), counts of mosquitoes in each cell during the high season, July 15 to August 30, were regressed against land cover, tax, census, soil, and meteorological data. A limited time frame was selected because the primary focus of this research was to determine the ATM’s habitat preference when abundance is highest and control is the most important, which is usually between June and late August (Sota et al. 1992; Barker et al. 2003). Initial analysis of the mosquito field data showed a significant decrease in the number of mosquitoes caught in September and October versus July and August (Figures 3a and 3b). Using a limited time frame also prevented temporal autocorrelation from affecting the analysis. As discussed at the beginning of the methods section, all cell level ATM counts were log-transformed to achieve normality prior to analysis. Individual models were developed for total count (Model 1T), females only (Model 1F), and males only (Model 1M) using the combined data from both counties for a total of 497 observations, which produced more powerful models with a lower risk of overfitting than only creating models for each county. For exploratory purposes, individual models for each county using cell level total counts were also created (Model 2Mer for Mercer County and Model 2Mon for Monmouth County), though these models were not considered sources of definitive information about habitat preferences in each county due to a lower ratio of dependent observations to predictor variables. The county cell level models included slightly different sets of independent variables, which were adjusted based on presence of the variables in the sites. For example, the Monmouth County cell level model included soil drainage and distance to coast, while the Mercer County model did not since these features are relatively invariant in Mercer. Excluding some independent variables from the county cell level models ensured at least five dependent observations per independent variable, which is considered the minimum for exploratory research (Allison 1999).

Additionally, site level models were developed for total count (Model 3T), females only (Model 3F), and males only (Model 3M) in each site from both counties during each week of the entire field season (July 15-October 30) to determine the effect of scale on model performance. Square-root transformed counts per week per site served as the dependent variable, for a total of 149 observations (4 sites x 16 weeks plus 5 sites x 17 weeks). Independent variables included meteorological data (temperature, precipitation, and wind), census data (population density, single parent households, vacant households, and rented households), and a coded dummy variable to indicate the county in which the data were collected. Land use/land cover data were not included in these models because the sites within each county had relatively similar proportions of each class.

In this analysis, 10-fold cross validation tested the models' performance against an independent data set via resampling. In *K*-fold cross validation (where *K*=10 in this case), the data are partitioned into *K* randomly selected subsamples of approximately equal size and the model is tested *K* times, each time leaving out one of the subsets as validation data. The results of the *K* folds are then combined to estimate the error of the model. Similarity between the original and cross-validated residuals confirms whether or not that the regression results are distorted by too many (or irrelevant) predictors and whether the models can be applied successfully to an independent dataset.

RESULTS

Land Use/Land Cover Classification

Post-classification evaluation of the sub-object level map revealed that several classes were often confused with one another (Table 3). In comparison to the reference data, the classified data often confused deciduous and coniferous woody vegetation, which is not surprising given the spectral and shape similarities of these two classes and the prevalence of shadows in the imagery. Also, grass (especially lawn and other grassland) was often classified as bare earth, probably because the ground was very dry throughout both counties when the 2007 imagery were taken. The accuracy totals also reveal that deciduous woody vegetation was underestimated (low producer's accuracy) in both counties (Table 3). Producer's accuracy measures errors of omission, or the likelihood that a reference point has been correctly classified, while the user's accuracy measure the error of commission, or the probability that a sample from the classified data actually represents that category on the ground. Bare land in Monmouth and other grassland in Mercer have high producer's accuracies and low user's accuracies, which indicate that most of the bare land and other grassland were classified, but many objects assigned to these classes actually belonged to another class. Also, several classes (cemeteries, all water classes except swimming pools, and commercial and public buildings) have 'perfect' user's or producer's accuracy, but these high accuracies may exist because fewer ground control points were selected from these classes.

Despite these errors, the sub-object (level 3) classification was 83.3percent accurate in Mercer County and 82.01 percent accurate in Monmouth County (Table 3). The Kappa statistics, which take into account the off-diagonal elements and errors of commission and omission, confirm the accuracy totals. The overall Kappa values also indicate that the classifications are at least 81.89 and 80.7 percent better than would be expected by random chance in Mercer County and Monmouth County, respectively. This level of accuracy is sufficient for modeling Asian tiger mosquito abundance in this case. The accuracy results also indicate that object-oriented classification is a useful method for describing urban ATM habitat.

Compared to previous studies of ATM abundance, the spatial grain of this land characterization was very fine (4 ft^2) and the classification scheme was very detailed. Such a fine scale classification was created to portray microhabitat and to potentially serve as a proxy for microhabitat weather data, which was not available for this time period. The super object level (level 1) represented basic land cover information, while the object and sub-object levels characterized a mixture of detailed land cover and land use (Figure 5; see Figure 4 for classification scheme). The percentages of each class by site are shown in Table 4. When creating the models, I focused on levels 2 and 3 of the classification since I suspected that these classes would be the most relevant to the mosquitoes as these levels better correspond to the landscape that an individual experiences.

Spatial Pattern Analysis

Spatial pattern analysis results are shown in Tables 4a and 4b. For the most part, spatial autocorrelation was inconsistent. Only two sites (Site 2 and Union Beach) exhibited consistent positive autocorrelation as measured by Moran’s I statistic for total, females, and males during both the high season and the entire season, though Cliffwood Beach exhibited consistent positive autocorrelation for total and females during both time periods (Table 5a). Keyport exhibited clustering for total and females during the entire season and females during the high season, while North Middletown only exhibited clustering for females during the high season. All other sites had random distributions for the high season and the entire season.

Moreover, consistent autocorrelation (across total, males, females) was more common for the seasonal counts than the monthly counts. None of the monthly counts showed consistent patterns across time (i.e., a site's counts are autocorrelated in every month) or across different types of counts (i.e., total, females, and males are autocorrelated in the same month) (Table 5b). July in Monmouth County appeared to have more clustering than any other month. July and October in Mercer County could not be tested because there were less than thirty observations per site in each month.

Similar to Moran's I, semivariograms did not reveal consistent patterns across the sites or counties (Table 6; see Appendix 1 for complete catalog of semivariograms). All of the graphs can be classified as either random, non-random with a perceptible sill, or hole effect (see Figure 9 for examples). Two sites, 5 and 7 in Mercer, had random semivariograms for all periods, though Keansburg and North Middletown each had only one non-random semivariogram. Although Moran's I and semivariograms are both measures of autocorrelation, some sites' semivariograms did not match the Moran's I results (Table 6). Major conflicts between Moran's I and the semivariograms occurred for Sites 2 and 3 in Mercer County, and Keyport, North Middletown, and Union Beach in Monmouth County.

Specifically, Site 2 exhibited clustering during the entire season and the high season according to Moran's I statistics. However, semivariograms for these periods showed random effects (Table 6). Conversely, Moran's I indicated that all counts in Site 3 were randomly distributed during the entire season and the high season, while semivariograms for the high season showed clustering between distances of approximately 75 and 190 meters (Table 6). This mismatch is almost exactly opposite the results for Site 2, but the data were verified and these results are not due to error.

Likewise, the Moran's I statistics and semivariograms for North Middletown and Union Beach conflicted. North Middletown exhibited clustering in females only during the high season according to Moran's I, while the semivariogram for this period was random (Table 6). Conversely, all monthly Moran's I statistics were random, but the semivariogram for August revealed clustering up to 80 meters. Additionally, though both Moran's I and semivariograms indicated clustering in Union Beach for all counts during the high season and entire season, monthly results conflicted. According to Moran's I, all monthly counts except July were random, while the semivariogram for July showed a random pattern and semivariograms for August and September showed clustering up to 450 and 110 meters, respectively. These differences may have occurred because Moran's I and semivariograms test autocorrelation slightly differently and serve different purposes in spatial analysis. Moran's I is an index of covariation between points and determines dissimilarity based on deviations from the mean (Cressie 1993; Fortin and Dale 2005). Semivariograms determine the degree of autocorrelation as a function of distances between points (i.e., variance). Also, semivariograms describe how spatial patterns vary as a function of scale, while Moran's I, as a measure of global autocorrelation, evaluates the overall pattern at a single scale (Ibid.).

Aside from these major differences, semivariogram results mostly agreed with

Moran's I results (Table 6). Cliffwood Beach exhibited clustering of total abundance up to approximately 475 meters during the entire season, while clustering of total and females only counts during the high season ranged from approximately 400 to 500 meters. Total counts during the entire season in Keyport were clustered, with a range of approximately 390 meters. Semivariograms for total abundance in July and September clearly indicated clustering up to 75 meters and 290 meters, respectively, though Moran's I statistics for these periods indicated a random distribution. Otherwise, Moran's I results and semivariogram results matched (Table 6).

Monthly semivariograms showed little autocorrelation, and if any was present it was inconsistent. Several monthly graphs showed oscillation, also known as the hole effect, which reflects either extreme values, periodicity in the spatial variability of the errors, or preferential clustering in the data (Bailey and Gatrell 1995). A typical example of the hole effect are shown in Figure Y2b. These graphs were interpreted as random because ranges could not be determined from the graphs.

The spatial characterization of ATM adult abundance in selected sites during the high season is shown in Figure 10 (see Appendix 2 for complete catalog of hotspot maps). Hot spot analysis was performed on total, females only, and males only abundance during the high season, entire season, and each month, but the high season's results had the most hot spots. Hot spots were not consistent between months for any site. In several cases, there were no hot spots for the monthly counts, which also tended to have more cool spots than hot spots.

In regard to season wide analysis, hot spots occurred more frequently during the high season than the entire season, though the results of the two periods never contradicted each other (Appendix 2). Mercer County sites had fewer hot spots than Monmouth County sites. Based on visual inspection, hot spots usually occurred in more residential areas or near the borders of residential areas and patches of vegetation, and tended not to occur in commercial areas, especially in Monmouth. Most of the sites in both counties consistently had hot spots in one portion of the site. Sometimes this area was a small corner or a very specific area (e.g., the park in Union Beach, the Henry Hudson Trail in Keyport, or the northern corner of Site 2), while in other cases it was more general (e.g., the north side of Cliffwood Beach, the western portion of Site 3, the northern and southern sides of Keansburg) (Figure 10). At the same time, in some sites the lack of hot spots in certain areas was more conspicuous than the presence of hot spots. For example, North Middletown seemed to have hot spots consistently throughout the site except in the northern corner. Even during the high season, some sites consistently had very few hot spots, such as Sites 5 and 7.

Most hot spots in total abundance data were due to concurrent high counts of males and females. However, some hot spots in total abundance occurred due to an exceptionally high catch of either females or males, which was evident depending on whether the same hot spot existed for females or males. For the most part, total, females only, and males only hot spot maps shared most hot spots, though occasionally, females or males had more hot spots than the total data (see Appendix 2 for examples).

Regression

At both the site level and the cell level, linear regression analyses were carried out to determine whether selected habitat variables and *Ae. albopictus* abundance were related. Cell level models were created using combined counts from both counties during the high season (July 15-August 30) for total count (Model 1T), females only (Model 1F), and males only (Model 1M), while models for each county (Model 2Mer and Model 2Mon) were created using only total counts from the high season. Table 7 shows the results of the cell level regression. Site level models for total count (Model 3T), females only (Model 3F), and males only (Model 3M) were created using data from both counties during each week of the entire field season, July 15-October 30. Table 8 shows the results of the site level regression.

Approximately 20-25% of the variation in mosquito abundance at the cell level can be explained by a combination of meteorological, land use/land cover, and housing structure variables (Table 7). The regression equations (with standard errors shown in parenthesis):

$$t_{cell} = \underset{(0.154)}{2.543} + \underset{(0.234)}{1.694}\textit{lndecid} - \underset{(0.718)}{1.197}\textit{lngrass2} - \underset{(0.552)}{2.174}\textit{lnwet2} + \underset{(6.516)}{19.666}\textit{lnstrm2} + \underset{(12.688)}{29.67}\textit{lnothh2o2} - \underset{(3.321)}{8.993}\textit{lnothgr2} + \underset{(0.014)}{0.049}\textit{TC} + \underset{(0.0006)}{0.0004}\textit{vacant}$$

$$f_{cell} = \underset{(0.113)}{1.979} + \underset{(0.202)}{1.819}\textit{lndecid} - \underset{(0.465)}{1.489}\textit{lnwet2} + \underset{(5.528)}{18.195}\textit{lnstrm2} + \underset{(9.644)}{33.01}\textit{lnothh2o2} - \underset{(2.89)}{7.071}\textit{lnothgr2} - \underset{(0.028)}{0.067}\textit{pre14} + \underset{(0.011)}{0.028}\textit{tminc} + \underset{(0.319)}{1.204}\textit{pre} + \underset{(0.0004)}{0.001}\textit{wdir} + \underset{(0.0006)}{0.003}\textit{vacant}$$

$$m_{cell} = \underset{(0.196)}{0.758} + \underset{(0.043)}{0.105}\textit{draincl1} - \underset{(0.456)}{1.65}\textit{lnpave} - \underset{(1.045)}{2.515}\textit{lnbare} + \underset{(1.82)}{3.338}\textit{lnothgr} + \underset{(0.449)}{1.376}\textit{proprent} + \underset{(0.043)}{0.135}\textit{pre14} + \underset{(0.017)}{0.033}\textit{tminc} - \underset{(0.509)}{1.418}\textit{pre} + \underset{(0.001)}{0.004}\textit{vacant}$$

Only two variables, number of vacant housing units and temperature (average or minimum) on the catch date, were selected in all three cell level models. Although temperature on the day of the catch was a significant predictor in all three models, its contribution was relatively small in every model ($\beta_{1T} = .144$, $p=.0003$; $\beta_{1F} = .112$, $p =$

.0115; β_{1M} = .082, p = .0507). In contrast, vacant housing units were relatively more important than most of the other variables in the three models (β_{1T} = .288, p<.0001; β_{1F} = .275, p <.0001; β_{1M} = .198, p = .0001). Seven other important variables (lnDecid, lnWet2, lnStrm2, lnOthH2O2, LnOthGr2, Pre14, and Pre) were selected in two of the three cell level models (Table 7).

Indeed, models for females only and total were quite similar, sharing significant predictors including the percentage of deciduous woody vegetation in a cell (lnDecid), streams and rivers within 200m of a cell (lnStrm2), and other water within 200m of a cell (lnOthH2O2) with positive coefficients, and wetland and other grassland within 200m of a cell (lnWet2 and lnOthGr2) with negative coefficients (Table 7). However, Model 1T required fewer predictors than Models 1F or 1M to achieve similar explained variance. Additionally, the model for males only was quite different from the other two cell level models. Half of the predictor variables in Model 1M were unique to males only, including soil drainage and percentages of pavement, bare land, and other grassland in a cell. Model 1M also had a higher error rate ($RMSE_{1M}$ = 0.897) than the other cell level models ($RMSE_{1T}$ = 0.6405, $RMSE_{1F}$ = 0.5518). The similarity between Models 1T and 1F makes sense because females comprised most of the catch in each cell, with a few exceptions. Error rates may have been higher for Model 1M due to less predictable patterns in male abundance and the pronounced difference between males abundance in the two counties.

Individual Mercer and Monmouth County models had similar R-squared values (R^2_{Mer} = 0.2122, R^2_{Mon} = 0.2028) to the combined county cell level models (Table 7). The county models also shared several predictor variables with the other cell level models, including temperature on the day of the catch and the percentage of woody or deciduous vegetation in a cell, as shown in the regression equations:

$$t_{Mer} = \underset{(0.125)}{2.558} + \underset{(0.354)}{1.644}\textit{lndecid} - \underset{(7.01)}{15.441}\textit{lnothgr2} - \underset{(0.001)}{0.042}\textit{TminC} + \underset{(0.02)}{0.004}\textit{vacant}$$

$$t_{Mon} = \underset{(0.351)}{0.716} + \underset{(0.328)}{1.876}\text{woody} + \underset{(0.036)}{0.1}\textit{draincl1} + \underset{(0.648)}{1.65}\textit{wet} + \underset{(15.819)}{55.127}\textit{lnswim2} -$$
$$\underset{(0.018)}{2.515}\textit{tmaxc} + \underset{(0.0002)}{0.0009}\textit{popden}$$

As in the combined county models, the temperature on the trapping date was a significant predictor in both models, though its contribution was relatively small ($\beta_{Me\ r}$ = -.132, p= .0309; β_{Mon}= -.158, p = .0036). TminC, or the minimum temperature on the trapping date, was a negative predictor in Model 2Mer, but a positive predictor in the combined county cell level models. Similarly, the percentage of woody/deciduous vegetation within the cell was a positive predictor of total ATM in Models 2Mer and 2Mon, causing a 1.64% and 1.88% increase in ATM abundance in a cell for every one percent increase in woody/deciduous vegetation in Mercer and Monmouth, respectively (β_{Mer} = .293, p= < .0001; β_{Mon} = .322, p < .0001).

Although the county models shared several predictor variables with other cell level models, the two county models were quite different from each other. In the Monmouth County model, the percentage of wetland within a cell (lnWet) was a significant predictor (β = .142, p = .033), though this variable was not tested in the Mercer model since lnWet and lnWet2 were equal to zero for all Mercer dependent observations in Mercer County (Table 7). However, in this case the percentage of wetland within the cell was positively associated with total ATM abundance in Monmouth County cells, which contradicts the results for the combined cell level models where the percentage of wetland within 200 meters of a cell negatively affected the number of ATM in the cell. Soil drainage was also a significant predictor in Model 2Mon, but not Model

2Mer, which was expected since soil drainage is nearly uniform in the Mercer sites (Table 7). In the Monmouth County model, better drained areas were associated with higher numbers of total ATM ($\beta = .191$, $p = .0056$), though this somewhat contradicts the inclusion of wetlands in the model since wetlands are poorly drained. Despite this discrepancy, the percentage of swimming pools within 200m of a cell (lnSwim2) was also selected as a significant predictor in the Monmouth County model ($\beta = .198$, $p = .0006$), which agrees with the inclusion of lnOthH2O2 (the percentage of other water within 200 meters of a cell) in Models 1T and 1F because lnSwim2 is one of the level 3 classes belonging to lnOth2O2 (Figure 4). Swimming pools are more common in the Monmouth sites than in the Mercer sites as the backyards tend to be larger and the houses farther apart, which may explain why this variable was selected for the Monmouth model but not the Mercer model. For the Mercer County model, the percentage of other grassland within 200m of a cell (lnOthGr2) was selected as a negative predictor of ATM abundance ($\beta = -.144$, $p = .0287$), which agrees with the combined cell level models. Similarly, the number of vacant housing units in a site was also selected as a significant positive predictor of ATM in Mercer County ($\beta = .322$, $p < .0001$), which may be because the Mercer sites historically and currently have a greater number of vacant houses than the Monmouth sites. This variable was not selected in Model 2Mon, though population density (PopDen) was a positive predictor variable ($\beta = .222$, $p = .0001$). Although the county models were for exploratory purposes only, the differences between them suggest that ATM habitat preferences may vary somewhat depending on location, with a few more universally important variables such as temperature, vegetation, and vacant houses.

While cell level models were able to explain some of the variance in ATM abundance, the site level models performed much better than the cell level models, explaining 72-78% of the variation in mosquito abundance (Adjusted $R^2_{3T} = 0.7781$, $R^2_{3F} = 0.7608$, $R^2_{3M} = 0.7331$; Table 8). Land use/land cover was not included in the site level models by design, so census and meteorological variables were the only predictor variables used in the models, as is evident in the regression equations:

$$t_{site} = \underset{(0.61)}{9.971} + \underset{(0.005)}{0.03}\textit{vacant} + \underset{(0.164)}{1.074}\textit{tc30} + \underset{(0.102)}{0.337}\textit{tc} - \underset{(0.173)}{0.625}\textit{pre14}$$

$$f_{site} = \underset{(0.765)}{7.65} + \underset{(0.0004)}{0.001}\textit{popden} + \underset{(0.005)}{0.023}\textit{vacant} + \underset{(0.139)}{0.844}\textit{tc30} + \underset{(0.086)}{0.285}\textit{tc} - \underset{(0.147)}{0.634}\textit{pre14}$$

$$m_{site} = \underset{(0.835)}{7.022} - \underset{(0.0003)}{0.001}\textit{popden} - \underset{(0.126)}{0.418}\textit{tmaxc30} + \underset{(0.07)}{0.303}\textit{tminc} - \underset{(0.027)}{0.093}\mathrm{w}\textit{max}$$

All site level models included vacant housing units (Vacant) and the average temperature during the 30 days preceding the trap set date (TC30) as significant positive predictors. Out of all the predictor variables, TC30 had the greatest effect on the number of ATM adults per site in all models ($\beta = 0.53$-0.56, $p < .0001$).

As with the cell level models for total and females only, models 3T and 3F shared all but one predictor variable (population density in Model 3F; Table 8). The average temperature on the trapping date (TC) and the total precipitation in the 14 days preceding the trapping date (Pre14) were significant predictors in both models, though Pre14 was a negative predictor and TC was a positive predictor. However, as at the cell level, Model 3M had additional predictive variables that were not present in the other site level models. The minimum temperature on the trap date (TminC) positively affected male abundance ($\beta = .393$, $p < .0001$), which is similar to the effect of TC in models 3T and 3F, while the maximum wind speed on the trap date (Wmax) and the maximum temperature in the 30 days preceding the trap date (TmaxC30) negatively affected male abundance

(β_{WMax} = -.172, p = .0012; $\beta_{TmaxC30}$ = -.261, p = .0006).

Besides these similarities between site and cell level models, several variables were consistently selected as significant predictors ($P < 0.05$) in most of the models, including the number of vacant housing units in a site and the temperature (average, minimum, or maximum) on the days that the trap was set and collected (Tables 4 and 5). At the site level, the average temperature during the 30 days prior trapping (TC30) was also very important. In the cell level models, the percentage of woody or deciduous woody vegetation in a cell and the percentage of other grassland and other water within 200m of a cell were also selected by several models, though these variables were not selected as consistently as temperature or vacant housing units.

Assumptions of normality, independence, homoscedasticity, and linearity were met by all models according to the Shapiro-Wilk test, the Durbin-Watson test, White's test, and graphs of the variables, respectively. The assumption testing results are not reported here for conciseness, but will be made available upon request.

Cross-validation demonstrates that overfitting was only problematic for the county cell level models (Table 9). For the cell level models (Models 1T, 1F, and 1M) and the site level models (Models 3T, 3F, and 3M), the validated R-squared was slightly higher than the original models' R-square and RMSE values were lower, suggesting that the original models' statistics underestimated the predictive power of the models. The cross-validation results indicate that these models will perform as well on future data as on the current data, though there is some uncertainty in this statement for the cell level models (1T, 1F, and 1M) since they only predicted about 20% of the variation in ATM abundance. In contrast, Models 2Mer and 2Mon had lower cross-validated R-squares than the original R-squares and higher RMSE (Table 9), suggesting that the original models for each county were slightly over fitted to the data and, thus, these models may not perform as well on future data. For this reason, cell level models for each county should be considered exploratory only and not good representations of ATM habitat preferences within each county.

DISCUSSION

This study revealed two important points about ATM ecology:

1. Scale (extent and grain) affects the detection of spatial patterns in ATM abundance. The extent of the study areas in this project did not capture consistent patterns in ATM abundance, though patterns may be present at a broader extent or multiple scales.
2. Fine-grain land use/land cover information in combination with more coarse-grain housing structure and meteorological information can predict 20% to 80% of the variability in ATM abundance, depending on the grain of the abundance.

Before these points can be explored in depth, however, certain limitations of this study should be noted. The original field work was designed to normalize differences between study areas and limit autocorrelation in the data so as to determine the effects of different control regimes. As a result, the field data, though very detailed, were not optimal for comparing different habitats because there was not a large range of habitat quality within the study sites. A wider range of habitat and environmental conditions would have been preferable for this analysis to better illuminate how habitat quality impacts ATM abundance. Moreover, only one season of field data was available since the larger project is in the early stages, so the results of the models may be specific to 2008. Though cross-validation indicated that all models except county cell level models would perform well on independent data, the true test will be validating the models with the 2009 data. Also, the focus of these models is adult density and where females go to feed and mate, rather than sources of adults. Since there is not always a link between larval and adult ATM abundance and distribution, habitat characteristics important for breeding (i.e., the prevalence of water-filled containers) may or may not be relevant to where male and female adults were trapped (Richards et al. 2006). Due to time constraints and a lag between the field data collection and analysis, the number of containers in the cells was not tested in the models. Other variables such as shade and nectar or blood meal availability may be more important to characterizing adult distribution than number of containers. Though vegetation classes served as proxies for shade and nectar, the models were missing other potentially important variables such as blood meal availability and relative humidity.

In addition to potentially excluding important predictor variables, the models were comprised of independent variables with different spatial and temporal grains. Some of the data, such as land use/land cover, tax parcel data, and soil drainage, provided detailed information about each cell, while other data (weather and census, especially) could only provide information at the site or county grain. Having all of the independent variables at a more consistent grain might change the results, though previous studies such as Rey et al. (2006) also included both fine- and coarse-scaled variables in their final models. Meteorological data from weather stations may be more appropriate for site level analysis, which could explain why cell level models did not perform well. Weather is indisputably important to forecasting ATM abundance, but coarsely scaled weather data may not be useful for forecasting abundance at the backyard level. Optimally, the imagery for the land use/land cover classification also would have been more recent because land cover changes may have occurred in the year lag between the image capture date and field data collection. Fortunately, orthophotos were available to classify woody vegetation for the 2008 field dates, though the 2008 imagery had a coarser resolution that could not detect trees that were smaller than 4 m^2. Similarly, tax parcel and census data were significantly older than the field data, which introduced additional error into the models.

Lastly, the methods themselves have limitations that should be noted. Stepwise regression carries with it the temptation to let the automatic algorithm determine the questions we ask about the data (Allison 1999). In this case, though, the large number of predictor variables required a streamlined approach for finding the most significant predictors of ATM abundance. The 'best' models selected by the stepwise algorithm were not accepted at face value, but were further tested to remove noise variables, which can be problematic in stepwise regression (Allison 1999; Tabachnick and Fidell 2001). In terms of spatial pattern analysis and determining the suitability of the data for regression, I only tested for spatial and temporal autocorrelation. Perhaps checking for spatio-temporal autocorrelation would show patterns that could not be captured by separate tests (Fortin and Dale 2005).

1. Considerations of Scale

While these limitations were considered and accounted for as much as possible, the results of spatial pattern analysis indicate that the spatial extent of the study itself limited the detection of patterns in ATM abundance, though certain sites (e.g., Site 2 in Mercer or Union Beach in Monmouth) displayed consistent positive global autocorrelation during the high season (July 15-August 30) and the entire season (July 15-October 30; Table 10a). According to Moran's I and semivariograms, very little consistent clustering occurred within the sites, suggesting that the sites effectively contained random or homogenous distributions of mosquitoes (Table 10a). The prevalence of random distributions is not surprising given that the study was designed to minimize autocorrelation and differences between sites. On a positive note, this result indicates that the 2009 field study can consider the sites equivalent for comparing different control regimes. As for the broader implications, the overall lack of consistent clustering suggests that the extents of the sites were too small to include enough variability in habitat characteristics to cause clustering in mosquito abundance, which may be present at a broader scale.

In addition to a lack of global autocorrelation, the semivariograms did not reveal a consistent distance radius for controlling problem areas. The semivariograms' ranges, which can be thought of as the maximum distance at which autocorrelation occurs within a site, varied from 75 meters to 500 meters (Table 10a). Although a range of 450 meters, for example, implies that ATM abundance within 450 meters of any given point is correlated, this number does not provide general enough information to determine the best distance to focus control around a problem area unless the sites' ranges of clustering are relatively similar to one another. Similarly, the prevalence of completely random semivariograms means that mosquitoes are randomly distributed throughout many of the sites and so the number of adults at one location cannot be used to predict the number of adults at another location. In short, according to this data, control efforts may need to blanket a study site of this size in order to manage problem areas. Moreover, the ranges of the non-random semivariograms were approximately one fifth to one half of the sites' area, which implies that the sites should contain only a few clusters since the range is nearly as large as the site. This, too, points toward variation at a broader scale or extent.

Additionally, spatial statistics for monthly ATM counts revealed mostly random patterns, which implies that clustering in mosquito abundance does not change over time and it does not occur only during a specific time period. Many monthly semivariograms exhibited oscillation (i.e., hole effect), which can be caused by extreme values, periodicity in the spatial variability of the errors, or preferential clustering in the data (Bailey and Gatrell 1995). In regard to causes of this pattern, periodicity does not make sense with so few points and preferential clustering should not occur in randomly collected data, so this effect may be due to the presence of extreme values. Similarly, monthly hot spot analysis revealed few hot spots and no consistency in hot spot locations over time. The lack of consistency in these maps may have occurred because different cells were sampled each month, while the lack of hot spots may have occurred because often the minimum of 30 points for analysis was barely met. With either points at the same location over time or more points per monthly sample, the spatial analysis of monthly abundance might be more conclusive. Nonetheless, the lack of variation in local clusters over time means that mosquito control does not necessarily need to be targeted to different areas depending on the time of season.

Comparison of site characteristics (land cover, census, soil) did not reveal any particularly conspicuous differences in habitat variability between sites that exhibited clustering and sites that had random ATM distributions (Table 10a). However, within each county, sites with clustering showed some differences, especially in vegetation composition. In Mercer County, Site 2 had more deciduous woody vegetation, other grassland, parks, vacant housing units, and single parent families than the other three sites in the county. In contrast, Monmouth County sites with the most clustering (Union

Beach, Cliffwood Beach, and Keyport) had fewer vacant houses than the other two sites. Other than this difference, the three Monmouth sites with clustering did not display consistent dissimilarity from the other two sites. However, more Monmouth sites than Mercer sites exhibited global autocorrelation and local clustering. Interestingly, Monmouth County sites are also larger than Mercer County sites by 708,000 m^2 on average and they contain more heterogeneous land cover (i.e., a patchwork of wetlands, development, and forested areas in contrast to the more uniformly developed landscape of the Mercer County sites), which suggests that either a larger site extent or greater variability in habitat produces greater spatial heterogeneity in ATM abundance.

Yet even though autocorrelation was not consistent over time or within counties, subtle patterns of consistently ‘hotter’ areas were observed within several sites, such as the Henry Hudson trail in Keyport and the park in Union Beach (Figure 10). The areas that supported higher abundances were generally more residential and vegetated, which corresponds to the results of the regression models where woody vegetation was a significant positive predictor of abundance and previous studies that found ATM in predominantly residential areas (Sota et al. 1992; Barker et al. 2003; Richards et al. 2006). Therefore, even though global autocorrelation may not be present in the sites, more subtle local patterns can inform the priority areas for control.

Overall, the idea that scale is important to understanding populations and systems is not new. Indeed, a common tenet of landscape ecology states that “[t]he scale at which studies are conducted may profoundly influence the conclusions” (Turner 1989, 174). Applying this idea to an otherwise purely entomological or ecological study gives insight into the interactions between patterns and processes that may be affecting ATM populations. Within the broader context of geography and landscape ecology, the results of this study that indicate patterns in ATM abundance at broader extents compel a multiscale perspective to reveal the dynamics of the relationship between ATM populations and the landscape through varied grains and extents of spatial and temporal scales. Mosquito control decisions are often made at the county level about the types of control, the areas of greatest concern, and control policies. However, mosquito populations may vary at different scales. By applying a geographic perspective to mosquito control, we can better understand where the mosquitoes are, why they are where they are, and how they can be managed. Ultimately, this approach will reveal interactions between patterns and processes at a local site-specific scale and eventually at a larger regional scale.

2. Predicting ATM Abundance with Regression Models

Despite the study's limitations and the lack of strong spatial patterns in ATM abundance, regression analysis revealed significant relationships between adult ATM density and habitat characteristics. Components of the models can be divided into three categories:

- meteorological factors, especially temperature and precipitation;
- human abundance and housing characteristics, specifically vacant housing units and population density; and
- landscape composition, especially more detailed classes (levels 2 and 3).

Meteorological data was the most coarsely scaled independent variable, and temperature on the day the traps were set was consistently included as a positive predictor variable in the models, where higher temperatures resulted in greater abundance (Table

10b). Furthermore, average temperature during the 30 days preceding the trap date was also a significant positive predictor in all of the site level models. Alto and Juliano (2001) found that in a laboratory setting, higher temperatures and the absence of evaporation resulted in greater production of adults, though the temperature on the trapping dates probably affects the catch rate more than the population (DeGaetano 2005). Besides temperature, previous studies suggest that rainfall may be an important predictor of ATM populations (Hawley 1988; DeGaetano 2005; Tsuda et al. 2006). The total amount of rain in the 14 days preceding the catch (Pre14) was included in four of the six primary models (excluding exploratory county models) (Table 10b). However, the relationship was not uniform in the models as three of the models (Models 1F, 3T, and 3F) showed a negative relationship between Pre14 and ATM abundance, while one model (1M) showed a positive relationship between the two variables. Perhaps the inconsistency in relationship type occurred due to a behavioral difference between males and females, or it could be a fluke in the models caused by other missing independent variables. The predominantly negative relationship between precipitation and abundance contradicts previous findings of rainfall increasing adult populations (DeGaetano 2005), though it may be that many lighter rain events flush immature mosquitoes from containers or that rainy days reduce trapping efficiency (Reinert 1989; DeGaetano 2005). In order to predict ATM abundance for planning control, temperature data can be obtained from weather forecast information and precipitation can be determined from real time climate information. In this way, climate information could be a potential short-term predictive tool (DeGaetano 2005).

Several census variables were consistently selected in the models. The number of vacant houses in a site was selected as a significant predictor variable in all but one model (Table 10b). Standardized beta coefficients show that in two models, the variable Vacant had the greatest relative effect on ATM abundance, while it was still one of the more important variables in the other five models (Table 10b). Originally, vacant housing was tested in the models because empty houses often have more garbage, tire piles, and neglected vegetation than inhabited homes, providing more habitat and oviposition opportunities for ATM. However, the census data were eight years older than the field data, so it is possible that this result is noise. Also, according to the US Census Bureau (2009), a ‘vacant unit’ is any housing unit that is unoccupied at the time of the census, regardless of the reason for vacancy (e.g., for rent/sale, for seasonal use, or abandoned). A survey of current vacant or abandoned houses in the sites would definitively determine the validity of this result. Aside from vacant housing units, population density (PopDen) was also selected as a significant variable in Models 2Mon, 3F, and 3M, which may have occurred due to the greater availability of blood meals or a larger amount of garbage in areas where there are more people per unit area. Rather than truly socioeconomic measures, the census variables selected in the models describe the presence of humans on the landscape and the types of structures that they inhabit, much like land use indicates the type of human activity on the landscape.

Of the variables describing landscape composition, several land use/land cover classes were included in the final models (Table 10b). Soil drainage was only selected in two models (1M and 2Mon), and it may only be important when creating customized local models. The most relevant land use/land cover scales were levels 2 and 3, which represent the more detailed classes (Table 10b). Classes in level 1, the most general classification, were only selected in Model 1M as significant predictor variables, and when I tested models with only level 1 classes (plus weather and census data), the models performed much more poorly than when levels 2 and 3 were included. Furthermore, land use/land cover variables were removed from the stepwise cell level models, a much lower R-squared (approximately 0.1) was observed, indicating that land use/land cover information significantly improves the models. More detailed classes may capture the habitat that is most similar to what the ATM experiences and needs for survival, while coarser classes cannot capture microhabitat conditions.

Indeed, level 2 and 3 vegetation classes were included in all of the cell level models. Specifically, woody or deciduous vegetation appeared as a significant positive predictor variable in four of the five models, where a one percent increase in woody/deciduous vegetation resulted in a 1.6-1.88% increase in ATM abundance (Table 10b). In this case, woody and deciduous vegetation can be considered almost identical classes because coniferous woody vegetation comprised less than 10% of woody vegetation in both counties. Spatial analysis confirmed this result as borders between woody land and residential areas seemed to contain more hot spots (Figure 10), possibly because woody vegetation provides resting cover for adult mosquitoes and artificial containers used for oviposition are scattered throughout residential areas (Richards et al. 2006). Previous studies have found similar correlations between woody vegetation and mosquito abundance. Rey et al. (2006) found that *Ae. albopictus* abundance was positively associated with canopy and mixed vegetation coverage and negatively associated with urbanization-related variables such as building coverage, though their study focused on oviposition habitat rather than adult habitat. Similarly, Akram and Lee (2004) found that habitats that were exposed to the sun throughout the day were less preferred than habitats in the shade. Other studies (Sota et al. 1992; Barker et al. 2003) have found that densely wooded areas supported lower ATM populations than more open suburban areas, which corresponds with the results of this study because most of the woody vegetation was located in parks or lawns in suburban areas. Clearly, this variable is important to ATM microhabitat selection in urban and suburban areas.

Additionally, the percentage of other grassland (OthGr2) within 200 meters of a cell was a significant predictor in three of the five cell level models and other grassland within a cell (OthGr) was a significant predictor in Model 1M, though both variables were usually relatively less important predictor variables according to beta weights (Table 10b). However, OthGr2 had a negative relationship to ATM abundance, while OthGr elicited the opposite response. Although ATM inhabits highly urbanized areas, it is often found in suburban and rural areas where open spaces with vegetation are common (Tsuda et al. 2006). Since rights of way and vacant lots or brown fields often contain little woody vegetation and abut paved or commercial areas, there may be less shade available for adults to rest during the heat of the day, though litter in these areas may provide ample containers for oviposition.

Likewise, the percentage of wetland (Wet2) within 200 meters of a cell was a significant negative predictor in two of the five cell level models and wetland within a cell (Wet) was a significant positive predictor in Model 2Mon (Table 10b). While it could be argued that the ATM is not a wetland mosquito and thus should prefer suburban yards to wetland, it is also plausible that this result occurred because wetlands were not directly sampled, though field samples were taken nearby. The negative relationship between Wet2 and ATM abundance in Models 1T and 1F may have occurred due to differences in the two counties' land cover composition. For all Mercer County observations, lnWet and lnWet2 were equal to zero, so all ATM counts greater than zero would contribute to a negative correlation with wetland, thus causing the negative relationship in the model since Mercer observations comprised half of the dependent dataset. The positive relationship between ATM abundance and wetland in the Monmouth County model (2Mon) also points toward this explanation. Thus, this issue may have masked the true relationship between lnWet2 and ATM abundance, if any.

Similarly, the inclusion of wetland as a positive predictor within the Monmouth County cell level model (2Mon) contradicts the inclusion of soil drainage in the model ($\beta = .191$, $p = .0056$) since wetlands are poorly drained. This incongruity may have occurred because the presence of containers in wetlands is more relevant to ATM abundance than the poor drainage in wetlands. Garbage from storm drains often appears in wetlands when heavy rains flush debris from the sewers onto the floodplain, providing a wealth of oviposition sites.

Interestingly, the land cover classes most commonly included in models were those that described about the percentage of cover within 200 meters of a cell, which suggests that relevant habitat occurs over a wider area than the discrete trap location. For example, besides vegetation classes, several water classes (OthH2O2, Swim2, and Strm2) describing the amount of water within the 200 meter buffer were selected as significant positive predictor variables in three of the cell level models (Table 10b). Streams may be predictors because they provide cooler, moister microhabitats for adult ATM, though they probably are not sources of adults since larger, moving bodies of water contain predators (Hawley 1988). In contrast, other water, including swimming pools, bird baths, and temporary standing water, may be a significant predictor because it is a source of ATM adults. Besides these water classes, most of the other significant land cover variables (except woody/deciduous vegetation) described cover within 200 meters of a cell. This buffer may be significant because the ATM has a flight range of approximately 200 meters (Kitron et al. 1998), which could complicate precisely targeted control, especially if an individual's foraging and breeding habitats encompass a large range.

Model Improvement and Application

The extent of an individual's habitat is the most relevant scale for predicting abundance and planning control. Hence, exploring various grains of independent habitat variables may be important, especially as information about "the influence of intermediate scale (hundreds of meters) habitat characteristics on mosquito presence and abundance and data on microhabitat (less than a meter) discrimination…is scant" (Rey et al., 2006, 1135). The results of the linear regression suggest that fine grain data such as land use/land cover and soil drainage are relevant to the ATM, but that the extent of the area under consideration plays an equally important role in detecting patterns in abundance. Indeed, the extent may need to be larger than the size of the sites in this study in order to capture enough habitat heterogeneity to cause patterns in ATM abundance. While vegetation may have served as a proxy for microclimate variables in this analysis, finer scale weather information should be tested in future models as it may be more important than or rule out vegetation variables. Although the coarse grain meteorological data from weather stations are useful for predicting when outbreaks will occur, microclimate data may be more useful for fine grain predictions of abundance because it better simulates what an individual organism experiences in its environment. Indeed, the higher R-square values of the site level models may have been related to the correspondence in grain size between the sites and the weather and census data, and, thus, cell or 'backyard' level models might perform better given microhabitat weather and census data with finer grain size.

In terms of the value of these models for mosquito control, prediction maps can be generated from the models to display predicted abundance at a specific time. Based on the independent variables, including land use, weather, soil, and census data, the abundance of the ATM within individual cells or within entire sites can be estimated using the models developed in this study. For example, I created a map of estimated abundance in Union Beach during the first week of August 2008, by inputting the independent variables into Model 1T[1] to calculate abundance. Estimates of abundance were then joined to the spatial data layer showing the cells (Figure 11). A specific week was chosen because Model 1T includes the temperature on the day the traps were set, though this is the only time variant variable in the model. A period from 2008 was chosen so that the results could be verified with the field data, but this procedure could easily be applied to future time periods. Compared to the actual field counts for August 5-6, 2008, in Union Beach, the predicted values were -0.86 to 486.5 percent different, though most of the predicted values were within ±50 percent of the actual counts (Table 11). The RMSE is approximately ±15 mosquitoes, which is not surprising given that Model 1T could predict only 21 percent of the variation in ATM abundance. Despite these errors, the higher estimates of abundance in the southeastern corner of the study site somewhat coincide with the location of hot spots in this site during the high season (see Appendix

---

[1] Model 1T: $t_{cell} = 2.543 + 1.694\mathit{lndecid} - 1.197\mathit{lngrass2} - 2.174\mathit{lnwet2} + 19.666\mathit{lnstrm2} + 29.67\mathit{lnothh2o2} - 8.993\mathit{lnothgr2} + 0.049\mathit{TC} + 0.0004\mathit{vacant}$

2). This application of the models can produce fine grain maps of mosquito abundance for larger areas and, thus, provide another tool for achieving more precise targeted control.

CONCLUSION

As human populations continue to grow, conflict between mosquitoes and humans will increase, resulting in the need for carefully planned control efforts. Climate and land use change further complicate the relationship between humans and mosquitoes and increase the need for dynamic models that can use real-time data to predict mosquito outbreaks. In this project, I have attempted to determine the relationship between land use, climate, and mosquito distribution and abundance.

In this study, I found that grain size and extent play a vital role in determining the relationship between mosquito abundance and habitat factors. At the extent of this study, there was a lack of consistent autocorrelation in ATM abundance both within sites and over time, meaning that mosquitoes were randomly or homogenously distributed throughout the sites. Additionally, analysis of the ranges of autocorrelation in non-random semivariograms revealed that there was not a distance at which mosquito abundance was consistently autocorrelated, which implies that control efforts in these sites may need to occur throughout the sites rather than just in specific problem areas. Despite the lack of global autocorrelation, some consistent hot spot patterns were apparent in specific sites, especially in more residential and vegetated areas. These results corroborate the findings of previous studies (Rey et al. 2006; Rochlin et al. 2009). Based on the results of this study, I suspect that the spatial extent of the study sites (60-61 ha in Mercer, 99-140 ha in Monmouth) was too small to capture enough heterogeneity in ATM abundance in order to create accurate predictive models, though the grain (2017 m^2 cells in Mercer, 6831 m^2 cells in Monmouth) was a good size.

Overall, the results of the regression analysis provide some clues about ATM habitat preferences and the effect of scale on model results. Comparing 2009 field data to the models will provide a real test of the models' robustness, especially at the site level. Undoubtedly, meteorological conditions are very important to forecasting ATM abundance at all scales, and fine grain land use data may supplement coarsely scaled weather station data at the 'backyard' level. Regression analysis revealed that many variables are important to adult ATM, especially present temperature, vegetation (i.e., woody vegetation, other grassland, other water, and wetland), and the number of vacant housing units (Tables 4 and 5). These models provide a flexible means to elucidate the relationship between the ATM and landscape variables, while offering mosquito control experts the ability to develop more precise models depending on local conditions.

Future work should follow up on the site and cell level models with the 2009 field data. Given that exploratory analysis revealed some differences between counties, further development of county level models may provide more insight into ATM habitat preferences in subtly different landscapes. Testing fine scale weather data collected in 2009 will determine whether more detailed weather information enhances the predictive ability of the models. On a more extensive scale, continuing to fine-tune the models in these small areas will provide the foundation for county level and, eventually, state level predictive models. A potential extension of this project, which ties into other studies of the species' invasive capability (e.g., Levine et al. 2004; Benedict et al. 2007; Moffett et al. 2007), would be to gather presence data in New Jersey and create an ecological niche model to predict its range in the state, especially in areas where the ATM is currently not present but could spread.

While these results are just the beginning of a deeper understanding of ATM ecology, they will serve as a strong basis for future work. Limited spatial autocorrelation in the six sites selected for 2009 trapping (Cliffwood Beach, Union Beach, North Middletown, and Sites 3, 5, and 7) indicates that the study of different control treatments can be assured that variations between sites will not confound the comparison. In terms of ATM ecology, all results emphasize the importance of scale to studying an organism's habitat preferences, especially in reference to understanding spatial variability at multiple extents and grains. An additional challenge may be to determine the fundamental cues used by this species for selecting habitat and the most relevant extent and grain for predicting population fluctuations (Juliano et al. 2004; Rey et al. 2006). Spatial pattern analysis and geostatistical methods serve as useful alternatives or supplements to conventional statistics, though this approach has only recently been to be applied to studies of mosquito abatement (Rochlin et al. 2009). The inquiry into spatial patterns of ATM abundance in this study contributes to future control of this species in New Jersey by providing clues about where to focus control and trapping efforts and by suggesting the presence of patterns in ATM abundance at multiple scales.

Furthermore, this comparison of mosquito abundance to climate, land use/land cover, and human population and housing builds upon previous work that used similar modeling methods, while focusing on adult populations and incorporating a wider range of fine-scale information about habitat at a scale that is relevant to informing management decisions. Moreover, the regression models developed in this study may guide future efforts to create predictive models for larger areas. Since there are few published studies that attempt to explain the relationship between mosquito populations and landscape or climate characteristics in the northeastern United States (DeGaetano

2005, Brown et al. 2008), this study fills an important gap in research and will hopefully spur future efforts to understand the ecology of mosquitoes in this region. As the ATM is an important vector of many arboviruses and is regarded as one of the world's worst invasive species (Moore and Mitchell 1997; ISSG 2009), the empirical mosquito abundance models that are presented here have the potential to be a component of more extensive risk and invasion monitoring programs.

TABLES & FIGURES

Table 1. Independent variables assessed in the regression model.

	Independent Variable	Categorical/ Continuous	Resolution	Abbreviation
SSURGO	Soil Drainage	continuous	Cell	DrainCl
	Distance from Cell Centroid to Coast	continuous	Cell	DistOcn_Ft
LULC	(log) Percent LU class within 200m of cell		Cell	ln"LUCode"2*
	(log) Percent LU class in cell		Cell	ln"LUCode"*
Tax	Total property value of tax parcels in 2002	continuous	Cell	Propval
2000 Census	Population Density (persons/hectare)	continuous	Site	PopDen
	Proportion of Renter Occupied Housing Units	continuous	Site	PropRent
	No. Vacant Housing Units	continuous	Site	Vacant
	No. Single Parent Households	continuous	Site	SingleHH
Weather	Average Temperature (centered)	continuous	County by Date	TC, TC30
	Max Temperature (centered)	continuous	County by Date	TmaxC, TmaxC30
	Min Temperature (centered)	continuous	County by Date	TminC, TminC30
	Precipitation	continuous	County by Date	Pre, Pre14, Pre30
	Wind Speed	continuous	County by Date	W
	Wind Direction	continuous	County by Date	Wdir
	Max Wind Speed	continuous	County by Date	Wmax

* Land use (LU) codes for each class are given in Figure 4.

Table 2. Summary of regression models and their input parameters.

	Model Names	Grain	Indep. Variables	Other Notes
Model 1 'cell level'	1T – total ATM 1F – females only 1M – males only	Cell	land use/land cover, census, tax, soil drainage, weather	497 observations
Model 2 'county cell level'	2Mer – Mercer Cty 2Mon – Monmouth Cty	Cell	land use/land cover, census, tax, weather; distance to coast and soil drainage (Monmouth only)	• 218 obs. (Mer) • 279 obs. (Mon) • Total ATM count. • Exploratory only.
Model 3 'site level'	3T - total ATM 3F – females only 3M – males only	Site	weather, census	149 observations

Table 3. Level 3 (sub-object level) land use/land cover classification accuracy totals and Kappa (K^) statistics.

	MERCER COUNTY					MONMOUTH COUNTY				
Class	**Reference Totals**	**Classified Totals**	**Number Correct**	**Producers Accuracy**	**Users Accuracy**	**Reference Totals**	**Classified Totals**	**Number Correct**	**Producers Accuracy**	**Users Accuracy**
Pavement	49	42	34	69.39%	80.95%	70	71	59	84.29%	83.10%
Bare Land	30	44	25	83.33%	56.82%	63	64	46	73.02%	71.88%
Residential Bldg	48	44	42	87.50%	95.45%	72	60	58	80.56%	96.67%
Commercial Bldg	37	43	37	100.00%	86.05%	48	61	46	95.83%	75.41%
Public Buildings	19	14	14	73.68%	100.00%	40	30	29	72.50%	96.67%
Wetland	8	8	7	87.50%	87.50%	51	43	40	78.43%	93.02%
Stream/river	12	12	12	100.00%	100.00%	35	37	34	97.14%	91.89%
Pond or reservoir	11	13	11	100.00%	84.62%	24	31	24	100.00%	77.42%
Lawn	37	41	28	75.68%	68.29%	73	66	49	67.12%	74.24%
Cemetery	20	22	18	90.00%	81.82%	22	23	22	100.00%	95.65%
Park	23	18	17	73.91%	94.44%	33	33	30	90.91%	90.91%
Other Grass	40	40	33	82.50%	82.50%	42	61	39	92.86%	63.93%
Coniferous	33	42	32	96.97%	76.19%	63	63	50	79.37%	79.37%
Deciduous	55	41	37	67.27%	90.24%	79	71	53	67.09%	74.65%
Swimming Pool	38	38	37	97.37%	97.37%	70	64	60	85.71%	93.75%
Temporary Water	12	10	9	75.00%	90.00%	15	17	14	93.33%	82.35%
Bird Baths	0	0	0	0.00%	0.00%	17	22	17	100.00%	77.27%
Totals	472	472	393			817	817	670		

Overall Classification Accuracy (Level 3)	83.26%	**Overall Accuracy (L3)**	82.01%
Overall Kappa (Level 3)	0.8189	**Overall Kappa (L3)**	0.8283
Overall Classification Accuracy (Level 2)	85.59%	Overall Accuracy (L2)	84.94%
Overall Kappa (Level 2)	0.8318	Overall Kappa (L2)	0.807
Overall Classification Accuracy (Level 1)	90.47%	Overall Accuracy (L1)	90.70%
Overall Kappa (Level 1)	0.8654	Overall Kappa (L1)	0.8696

Table 4. Land use/ land cover composition of all sites.

	Site	S2	S3	S5	S7	Cliffwood	Keansburg	Keyport	Middletown	Union Beach
Level 3	Bare Land	3.64	1.22	2.21	2.21	4.35	2.90	4.21	3.66	5.04
	Bird Bath	0.00	0.00	0.00	0.00	1.83	0.00	0.00	1.40	0.00
	Commercial Bldg	4.53	5.82	6.48	3.31	2.79	3.17	1.90	2.55	1.51
	Coniferous	0.98	1.88	1.18	2.51	32.95	2.26	6.23	26.95	1.09
	Deciduous	22.80	18.58	10.94	19.44	19.21	25.19	3.37	22.77	18.04
	Lawn	12.81	15.51	10.52	18.89	1.60	23.45	23.22	0.81	25.13
	Other Grassland	0.44	0.15	0.01	0.27	1.63	0.35	12.22	1.08	1.06
	Park	3.14	0.09	0.00	1.90	22.99	0.45	0.30	22.59	0.79
	Pavement	34.32	38.06	45.03	37.42	0.00	27.02	1.25	0.07	26.19
	Public Bldg	0.05	0.00	0.83	2.81	0.06	0.89	34.82	0.06	0.11
	Residential Bldg	17.30	18.65	22.72	11.10	9.91	13.71	1.02	14.06	13.52
	Rivers/Streams	0.00	0.00	0.00	0.05	0.05	0.61	10.91	0.00	0.46
	Swimming Pool	0.00	0.05	0.09	0.10	0.46	0.00	0.42	0.78	0.79
	Wetland	0.00	0.00	0.00	0.00	2.16	0.00	0.12	3.22	6.27
Levels 1& 2	Building	21.87	24.47	30.03	17.22	12.76	17.76	37.74	16.67	15.13
	Other Water	0.00	0.05	0.09	0.10	2.29	0.00	0.42	2.17	0.79
	Water	0.00	0.05	0.09	0.15	2.33	0.62	11.33	2.17	1.25
	Vegetation	40.17	36.20	22.64	43.01	80.55	51.70	45.47	77.43	52.38
	Woody Veg.	23.77	20.45	12.11	21.95	52.16	27.45	9.61	49.72	19.14
	Grass	16.39	15.75	10.53	21.06	26.22	24.25	35.74	24.48	26.98
	Total area (m^2)	368,821	484,807	298,957	612,189	1,307,155	993,207	1,037,827	1,002,943	1,390,619

Table 5a. Moran's I results for seasonal counts. For I > 0, Cluster1 (p≤.05); Cluster2 (p<.01). For I ≈ 0, cell value is Random (p>.05). Bold text in grey cells represent a strong spatial pattern (Z>1.96) with less than 5% chance if being random. Gray cells represent spatial patterns (Z>1.65) with less than 10% chance of being random.

	Site	All Months, Total (T)	All Months, Females (F)	All Months, Males (M)	Active Season, T	Active Season, F	Active Season, M
Mercer County	2	**0.2 (3.59)**	**0.19 (3.37)**	**0.23 (4.06)**	0.16 (1.9)	0.09 (1.15)	0.14 (1.66)
	3	0 (0.14)	-0.01 (-.06)	0.03 (0.46)	-0.02 (-.04)	-0.02 (-.06)	-0.06 (-.32)
	5	-0.01 (-.02)	-0.02 (-.021)	-0.01 (-.14)	0 (0.14)	0.09 (1.04)	-0.19 (-1.7)
	7	-0.03 (-.033)	-0.06 (-.079)	0.05 (0.84)	0.04 (0.56)	-0.01 (0.05)	0.18 (2.08)
Monmouth County	Cliffwood	0.09 (1.74)	0.08 (1.61)	0.05 (0.84)	0.16 (1.92)	0.14 (1.68)	-0.02 (-.12)
	Keyport	**0.16 (2.91)**	**0.18 (3.24)**	0.09 (1.65)	0.13 (1.93)	**0.14 (2)**	0.03 (0.54)
	Keansburg	-0.05 (-.056)	-0.02 (-.12)	-0.05 (-.059)	0.13 (1.4)	0.14 (1.5)	0.15 (1.57)
	Middletown	0.04 (0.96)	0.05 (1.27)	0.02 (0.44)	0.09 (1.36)	**0.14 (2.03)**	0.03 (0.5)
	Union Beach	0.09 (1.88)	0.09 (1.73)	0.1 (1.91)	**0.28 (2.87)**	**0.28 (2.88)**	**0.21 (2.19)**

Table 5b. Moran's I results for monthly counts. Bold text in grey cells represent a strong spatial pattern (Z>1.96) with less than 5% chance if being random. Gray cells represent spatial patterns (Z>1.65) with less than 10% chance of being random. Z scores are shown in parenthesis under each I value.

	Site	July, T	July, F	July, M	Aug, T	Aug, F	Aug, M	Sept, T	Sept, F	Sept, M	Oct, T	Oct, F,	Oct, M
Mercer County	2	<30 obs.	<30 obs.	<30 obs.	0.09 (0.87)	-0.03 (-0.03)	0.07 (0.73)	0.14 (1.33)	0.17 (1.58)	0.1 (1.02)	<30 obs.	<30 obs.	<30 obs.
	3	<30 obs.	<30 obs.	<30 obs.	-0.06 (-0.36)	-0.03 (-0.08)	-0.1 (-0.87.)	0.02 (0.31)	-0.05 (-0.23)	0.06 (0.64)	<30 obs.	<30 obs.	<30 obs.
	5	<30 obs.	<30 obs.	<30 obs.	-0.13 (-0.61)	0.09 (0.72)	-0.31 (-1.68)	-0.21 (-1.64)	-0.25 (-1.96)	-0.06 (-0.29)	<30 obs.	<30 obs.	<30 obs.
	7	<30 obs.	<30 obs.	<30 obs.	-0.07 (-0.34)	-0.05 (-0.16)	0.1 (0.96)	-0.15 (-1.16)	-0.08 (-0.53)	-0.19 (-1.51)	<30 obs.	<30 obs.	<30 obs.
Monmouth County	Cliffwood	-0.08 (-0.5)	-0.06 (-0.33)	-0.14 (-0.95)	0.17 (1.3)	0.21 (1.55)	-0.13 (-0.7)	0.12 (1.07)	0.02 (0.3)	**0.32 (2.56)**	0.02 (0.3)	-0.02 (-0.07)	-0.1 (-0.54)
	Keyport	0.06 (0.86)	0.11 (1.38)	-0.06 (-0.37)	0.13 (0.98)	0.22 (1.57)	-0.15 (-0.76)	0.09 (0.97)	0.07 (0.74)	0.07 (0.75)	0.07 (0.77)	0.07 (0.8)	-0.12 (-0.88)
	Keansburg	0.22 (1.69)	0.18 (1.44)	0.19 (1.45)	0.06 (0.59)	-0.01 (0.09)	0.25 (1.85)	-0.19 (-1.19)	-0.2 (-1.26)	-0.17 (-1.04)	<30 obs.	<30 obs.	<30 obs.
	Middletown	-0.06 (-0.32)	-0.03 (-0.05)	-0.11 (-0.8)	0.1 (0.94)	0.12 (1.11)	0.03 (0.41)	-0.08 (-0.54)	-0.07 (-0.49)	-0.05 (-0.29)	-0.14 (-0.88)	-0.2 (-1.36)	-0.15 (-1.36)
	Union Beach	**0.21 (2.47)**	**0.16 (1.97)**	0.11 (1.38)	0.06 (0.63)	0.09 (0.82)	0.04 (0.48)	0.1 (1.23)	0.1 (1.3)	0.06 (0.87)	0.03 (0.51)	-0.02 (0.03)	-0.02 (0.1)

Table 6. Semivariogram results with Moran's I results. Bold text represents non-random semivariograms with ranges. Red cells show mismatches between the two statistics, while green cells show non-random matches. All October semivariograms and I values were random.

	All Months, T		Active Season, T		Active Season, F		Active Season, M		July, T		Aug, T		Sept, T	
Site	Semivar	Moran I	Semivar	Moran I	Semivar	Moran I	Semivar	Moran I	Semivar	Moran I	Semivar	Moran I	Semivar	Moran I
2	Random	Cluster2	Random	Cluster1	Random	Random	Random	Cluster1	Random	n/a	**90m**	Random	Random	Random
3	Random	Random	**190m**	Random	**120m**	Random	**75 m**	Random	Random	n/a	Random	Random	Hole Effect	Random
5	Hole Effect	Random	Random	Random	Random	Random	Random	Random	Hole effect	n/a	Random	Random	Hole Effect	Random
7	random	Random	Random	Random	Random	Random	Random	Random	Random	n/a	Hole Effect	Random	Random	Random
Cl	**475m**	Cluster1	**400m**	Cluster1	**500m**	Cluster1	Random	Random	Random	Random	Random	Random	Random	Random
Ky	**390m**	Cluster2	Random	Random	Random	Cluster2	Random	Random	**75m**	Random	Random	Random	**290m**	Random
Kn	Random	Random	Random	Random	Random	Random	Random	Random	**150m**	Cluster1	Random	Random	Random	Random
Md	Random	Random	Random	Random	Random	Cluster2	Random	Random	Hole effect	Random	**80m**	Random	Random	Random
Ub	**460m**	Cluster1	**330m**	Cluster2	**450m**	Cluster2	**330m**	Cluster2	Random	Cluster2	**450m**	Random	**110m**	Random

* Moran's I Coding: Cluster1 (1.65 < Z < 1.96); Cluster2 (Z > 1.96). See tables 4a and 4b for Moran's I statistics.
Site Abbreviations: Cl = Cliffwood Beach, Ky = Keyport, Kn = Keansburg, Md = North Middletown, Ub = Union Beach

Table 7. Multiple regression coefficients and results for cell level models, July 15-August 30, 2008.

Variable / Model	Model 1T	Model 1F	Model 1M	Model 2Mer	Model 2Mon
lnDecid (% deciduous woody vegetation in cell)	0.297**	0.297**		0.293**	
lnGrass2 (% Grass within 200m)	-0.0878				
lnWet2 (% Wetland within 200m)	-0.262**	-0.203**			
lnStrm2 (% Streams within 200m)	0.184**	0.192**			
lnOthH2O2 (% Other Water within 200m)	0.131**	0.164**			
lnOthGr2 (% Other Grassland within 200m)	-0.110**	-0.097**		-0.144**	
TC (Temp on trap date)	0.144**				
Vacant housing units in site	0.288**	0.275**	0.198**	0.323**	
Pre14 (Total precipitation 14 days before trap date)		-0.113**	0.144**		
TminC (Min Temp on trap date)		0.112**	0.082*	-0.132**	
Pre (Precipitation on trap date)		0.159**	-0.119**		
Wdir (Wind Direction)		0.125**			
DrainCl (Soil Drainage)			0.130**		0.191**
lnPave (% Pavement in cell)			-0.157**		
lnBare (% Bare land in cell)			-0.102**		
LnOthGr (% Other Grassland in cell)			0.078*		
PropRent (Proportion rented housing units)			0.157**		
lnWoody (% Woody vegetation in cell)					0.322**
lnWet (% Wetland in cell)					0.142**
lnSwim2 (% Swimming Pools within 200m)					0.198**
TmaxC (Max Temp on Trap Date)					-0.158**
PopDen (Population Density)					0.223**
R-squared	0.2212	0.2657	0.2199	0.2267	0.2200
Adjusted R-squared	0.2085	0.2519	0.2053	0.2122	0.2028
No. Observations	497	497	497	218	279

Reported values are standardized beta weights. All
* and ** indicate significance at the 90% and ≥ 95% level, respectively.

Table 8. Multiple regression coefficients and results for site level models, July 15-October 30, 2008.

Variable / Model	Model 3T	Model 3F	Model 3M
Vacant (Vacant housing units in site)	0.237	0.225	0.295
TC30 (Avg Temp 30 days before trap date)	0.559	0.539	0.565
TC (Temp on trap date)	0.288	0.299	
Pre14 (Total precipitation 14 days before trap date)	-0.143	-0.178	
PopDen (Population Density)		0.126	-0.116
TmaxC30 (Max temp 30 days before trap date)			-0.261
TminC (Min Temp on Trap Date)			0.393
Wmax (Max wind speed on trap date)			-0.172
R-squared	0.7839	0.7686	0.7331
Adjusted R-squared	0.7781	0.7608	0.7222
No. Observations	154	154	154

Reported values are standardized beta weights.
All variables were significant at the 95% level.

Table 9. Cross-validation results for all models.

	Model 1T	Model 1F	Model 1M	Model 2Mer	Model 2Mon	Model 3T	Model 3F	Model 3M
R-squared (developmental)	0.2212	0.2657	0.2199	0.2267	0.22	0.7839	0.7686	0.7331
Adjusted R-squared (developmental)	0.2085	0.2519	0.2053	0.2122	0.2028	0.7781	0.7608	0.7222
RMSE (developmental)	0.6405	0.5518	0.8970	0.6886	0.5944	3.1293	2.6496	2.2534
R-squared (validation)	0.2393	0.2721	0.2347	0.1552	0.1517	0.7823	0.8575	0.7547
Adjusted R-squared (validation)	0.2254	0.2554	0.2172	0.1376	0.1309	0.7758	0.8522	0.7436
RMSE (validation)	0.6184	0.5344	0.8751	0.7186	0.6264	3.1648	2.5633	2.4168

Table 10a. Summary of spatial pattern analysis for total counts in each site and physical differences between sites. In the grey cells, non-random Moran's I statistics are shown with Z scores in parenthesis and ranges for non-random semivariograms are given. See Tables 4 and 5 for complete record of spatial pattern analysis results, and Table 4 for more information about site land cover composition.

	Site	**2**	**3**	**5**	**7**	**Cl**	**Ky**	**Kn**	**Md**	**Ub**
Entire Season, Total (T)	Semivar.	CSR*	CSR	Hole Effect	CSR	475m	390m	CSR	CSR	460m
	Moran's I	0.2 (3.59)	CSR	CSR	CSR	0.09 (1.74)	0.16 (2.91)	CSR	CSR	0.09 (1.88)
Active Season, T	Semivar.	CSR	190m	CSR	CSR	400m	CSR	CSR	CSR	330m
	Moran's I	0.16 (1.9)	CSR	CSR	CSR	0.16 (1.92)	CSR	CSR	CSR	0.28 (2.87)
July, T	Semivar.	CSR	CSR	Hole effect	CSR	CSR	75m	150m	Hole effect	CSR
	Moran's I	n/a	n/a	n/a	n/a	CSR	CSR	0.22 (1.69)	CSR	0.21 (2.47)
Aug, T	Semivar.	90m	CSR	CSR	Hole Effect	CSR	CSR	CSR	80m	450m
	Moran's I	CSR	CSR	CSR	CSR	CSR	CSR	CSR	CSR	CSR
Sept, T	Semivar.	CSR	Hole Effect	Hole Effect	CSR	CSR	290m	CSR	CSR	110m
	Moran's I	CSR	CSR	CSR	CSR	CSR	CSR	CSR	CSR	CSR
Oct, T	Semivar.	CSR	CSR	CSR	CSR	CSR	CSR	n/a	CSR	CSR
	Moran's I	n/a	n/a	n/a	n/a	CSR	CSR	CSR	CSR	CSR
Site Characteristics	Deciduous Trees (%)	22.80	18.58	10.94	19.44	19.21	25.19	3.37	22.77	18.04
	Other Grass (%)	0.44	0.15	0.01	0.27	1.63	0.35	12.22	1.08	1.06
	Parks (%)	3.14	0.09	0.00	1.90	22.99	0.45	0.30	22.59	0.79
	Vacant HU	216	164	60.00	92	28	78	89	93	69
	Single Parent HH	165	219	111.00	147	97	150	139	149	143

* CSR = Complete Spatial Randomness
Site Abbreviations: Cl = Cliffwood Beach, Ky = Keyport, Kn = Keansburg, Md = North Middletown, Ub = Union Beach

Table 10b. Summary of the most notable results from the regression models. See Tables 6 and 7 for complete record of the results.

		Cell-level Models					Site-level Models		
	Variable / Model	1T	1F	1M	2Mer	2Mon	3T	3F	3M
Weather	TC (Temp on trap date)	0.144					0.288	0.299	
Weather	TminC (Min Temp on Trap Date)		0.112	0.082	-0.132				0.393
Weather	TC30 (Avg temp 30 days before trap date)						0.559	0.539	0.565
Weather	Pre14 (Total precipitation 14 days before trap date)		-0.113	0.144			-0.143	-0.178	
Census	Vacant (Vacant housing units in site)	0.288	0.275	0.198	0.323		0.237	0.225	0.295
Land Use/Land Cover	lnDecid (% deciduous woody vegetation in cell)	0.297	0.297		0.293				
Land Use/Land Cover	lnWet2 (% Wetland within 200m)	-0.262	-0.203						
Land Use/Land Cover	lnStrm2 (% Streams within 200m)	0.184	0.192						
Land Use/Land Cover	lnOthH2O2 (% Other Water within 200m)	0.131	0.164						
Land Use/Land Cover	LnOthGr2 (% Other Grassland within 200m)	-0.110	-0.097		-0.144				
Land Use/Land Cover	lnPave (% Pavement in cell)			-0.157					
Land Use/Land Cover	lnBare (% Bare land in cell)			-0.102					
Land Use/Land Cover	LnOthGr (% Other Grassland in cell)			0.078					
Land Use/Land Cover	lnWet (% Wetland in cell)					0.142			
Land Use/Land Cover	lnSwim2 (% Swimming Pools within 200m)					0.198			
	Adjusted R-squared	0.2085	0.2519	0.2053	0.2122	0.2028	0.7781	0.7608	0.7222
	RMSE	0.6405	0.5518	0.8970	0.6886	0.5944	3.1293	2.6496	2.2534
	No. Observations	497	497	497	218	279	154	154	154

Reported values are standardized beta weights.

Table 11. Validation results for map of predicted ATM abundance in Union Beach during the first week of August, 2008, based on model 1T. See also Figure 11.

Cell ID	Field Count	Predicted Count	Percent Difference
Ub - 6	13	17.78	36.77
Ub - 25	11	15.98	45.26
Ub - 36	5	18.31	266.18
Ub - 47	21	20.82	-0.86
Ub - 65	56	25.44	-54.57
Ub - 76	10	25.45	154.49
Ub - 98	4	23.46	486.53
Ub - 108	22	11.17	-49.23
Ub - 119	14	22.24	58.89
Ub - 160	37	17.49	-52.73
Ub - 164	7	5.87	-16.16
RMSE	15.303	Lower CI (95%)	23.946
MAE	11.676	Upper CI (95%)	25.836

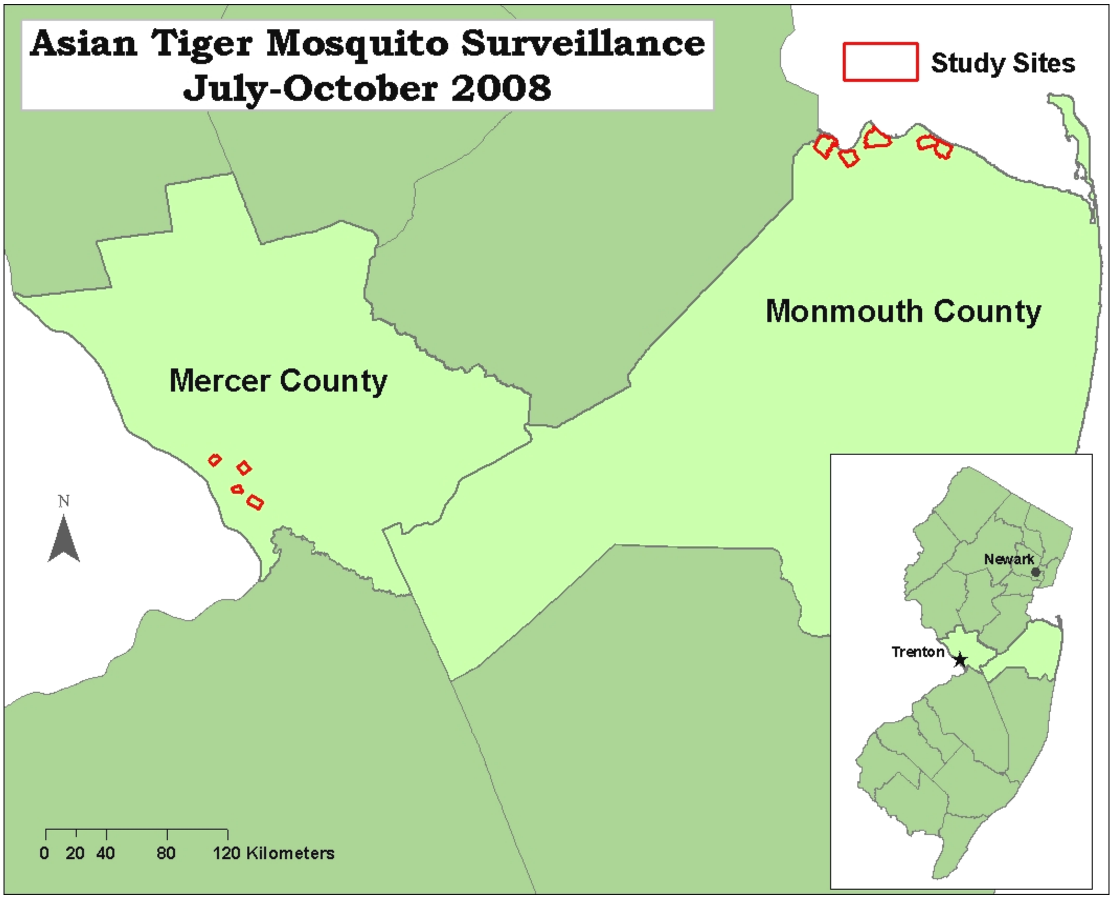


Figure 1. ATM trapping sites in Mercer and Monmouth Counties, New Jersey. Each site was further subdivided into 102-167 cells containing 6-8 property parcels (Mercer) or 8-10 property parcels (Monmouth).

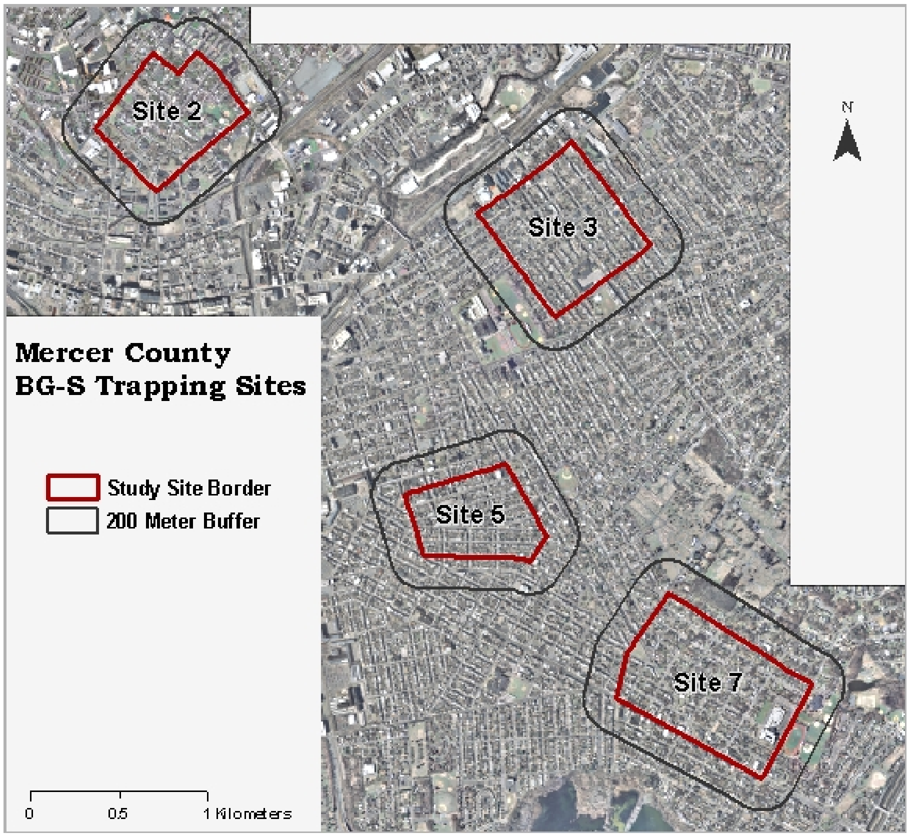


Figure 2a. Mercer County ATM trapping sites. Sites to be used in 2009 are Site 3, Site 5, and Site 7.

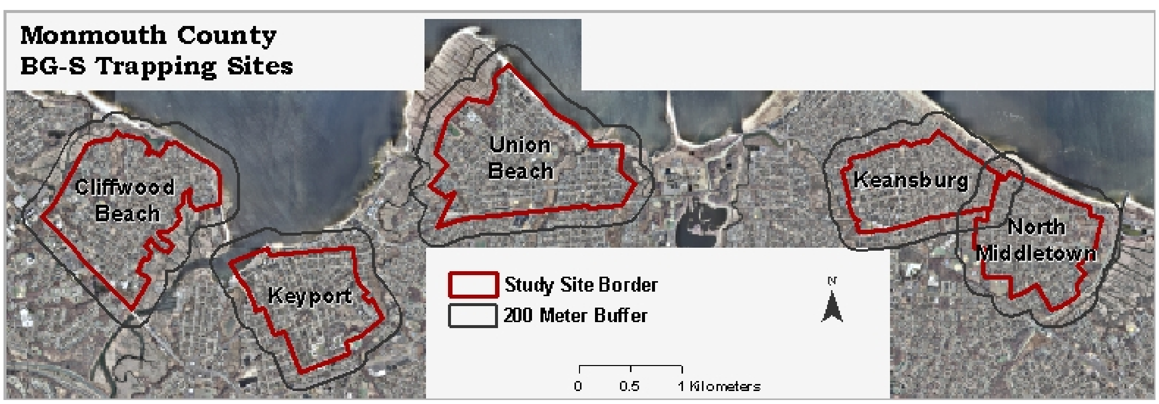


Figure 2b. Monmouth County ATM trapping sites. Sites to be used in 2009 are Cliffwood, Union Beach, and North Middletown.

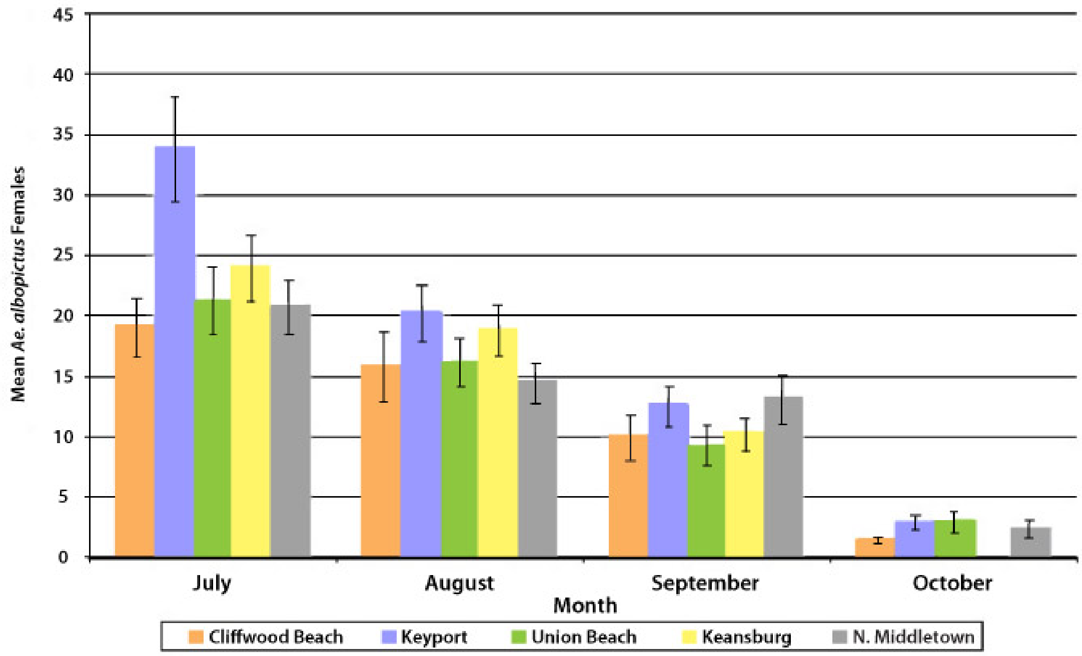


Figure 3a. Monmouth County, mean *Ae. albopictus* per trap per site per month.

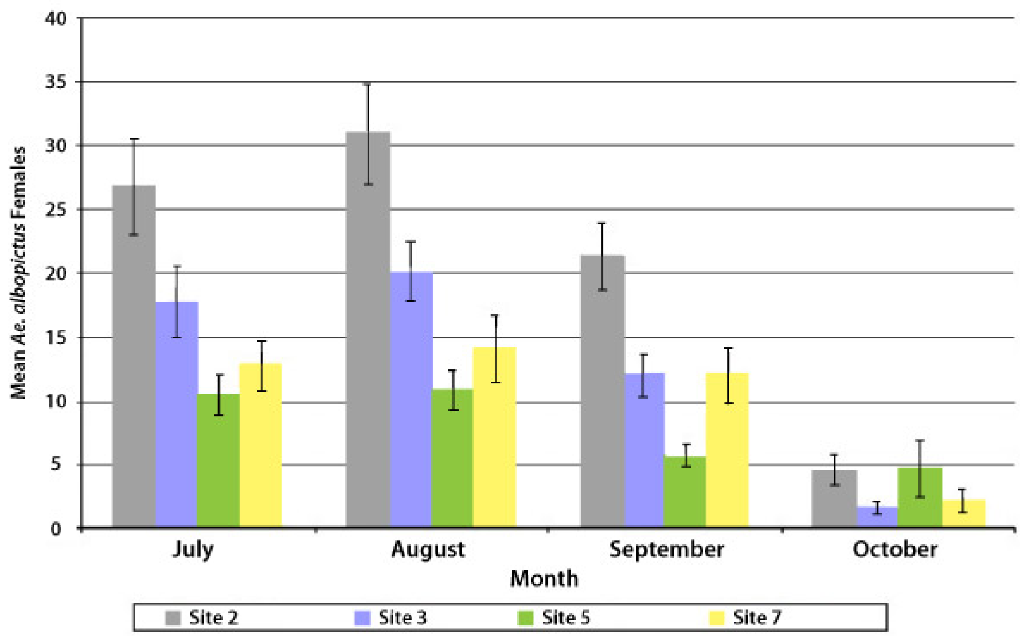


Figure 3b. Mercer County, mean *Ae. albopictus* per trap per site per month.

Inheritance Hierarchy: Super Objects (Level 1) $\rightarrow$ Objects (Level 1.1) $\rightarrow$ Sub-Objects (Level 1.1.1)

Bldg	1. Buildings		
Res		1.1. Residential Building	
Comm		1.2. Commercial Building	
Pub		1.3. Public Building (library, school, government, etc)	
Pave	2. Pavement (roads, driveways)		
Veg	3. Vegetation		
Ag		3.1. Agriculture	
Crop			3.1.1. Cropland
Hort			3.1.2. Orchards, Nurseries, and Other Horticultural
Grass		3.2. Grassland	
Lawn			3.2.1. Lawn (residential and commercial)
Cem			3.2.2. Cemetery
Park			3.2.3. Parks and Recreation
OthGr			3.2.4. Other (rights of ways, vacant lots)
Woody		3.3. Woody (trees and shrubs)	
Conif			3.3.1. Coniferous
Decid			3.3.2. Deciduous
Wet		3.4. Wetland	
Water	4. Water		
Strm		4.1. Stream/river	
Pond		4.2. Pond/lake/reservoir	
OthH2O		4.3. Other	
Swim			4.3.1. Swimming Pool
Temp			4.3.2. Temporary or Construction
BirdB			4.3.3. Bird Baths
Bare	5. Bare Land (construction sites, beach, bare earth, etc)		

Figure 4. Hierarchical, fine-scale land use/land cover classification based on object-oriented image segmentation. Classes shown in grey were not present in any of the study areas.

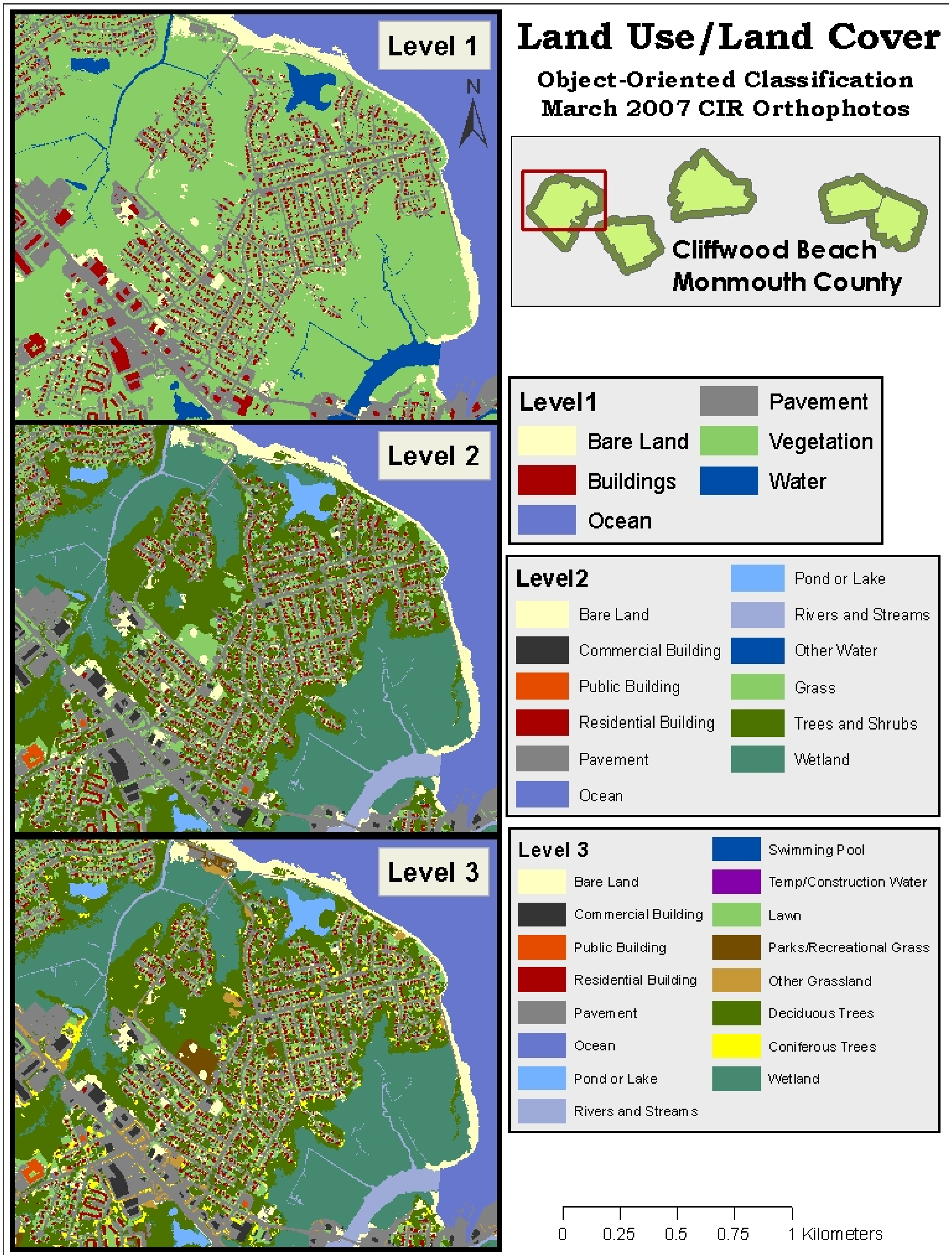


Figure 5. Hierarchical land use/land cover classification for part of Cliffwood Beach.

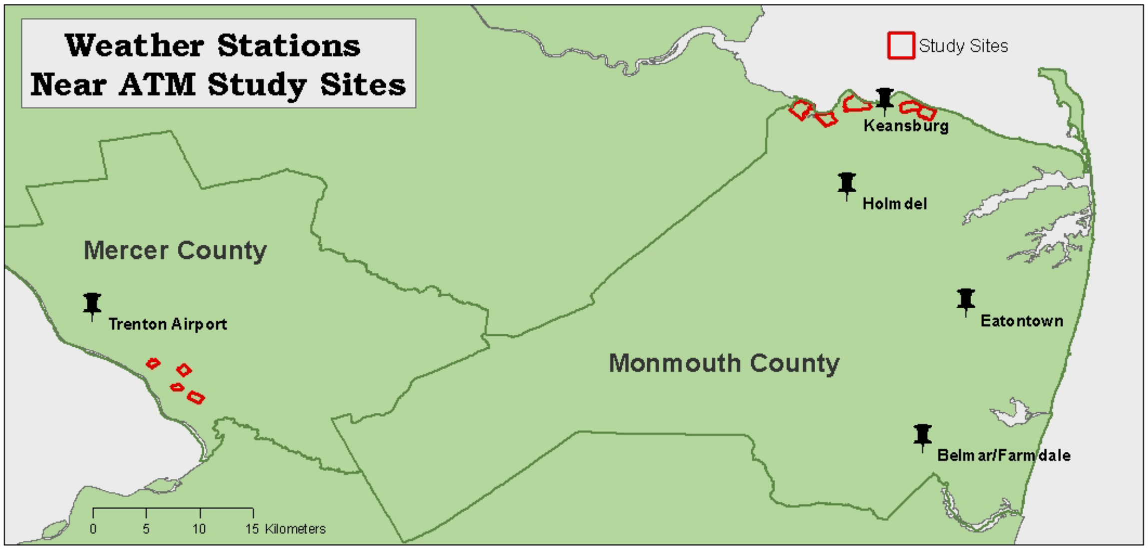


Figure 6. Weather stations used to determine meteorological conditions in each site. The Trenton Airport station was selected as the primary station in Mercer. Keansburg and Holmdel served as the primary stations for Monmouth, though other stations including Eatontown and Belmar/Farmdale filled in data when Keansburg and Holmdel data were unavailable.

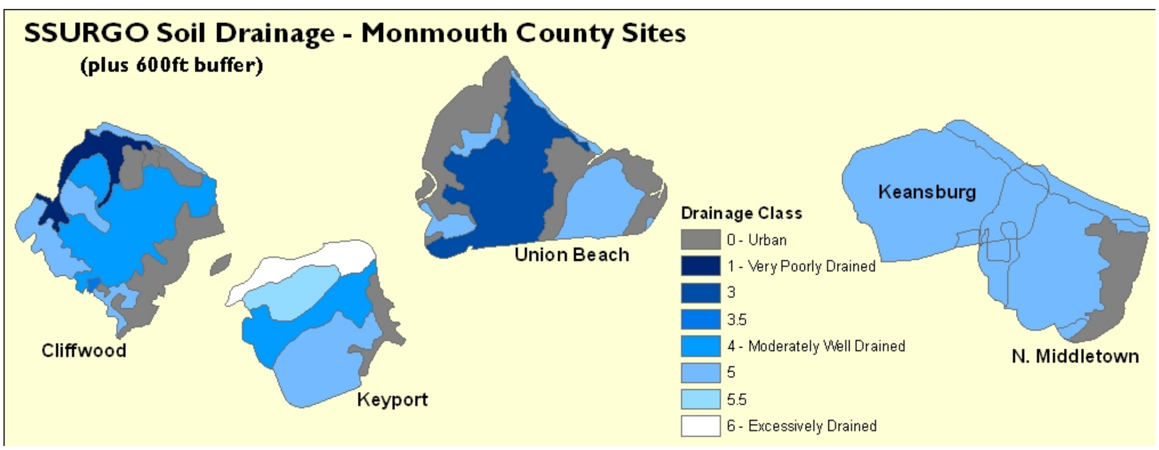


Figure 7. SSURGO soil drainage in Monmouth County sites. Soils were rated based on composition and drainage quality described in the SSURGO database.

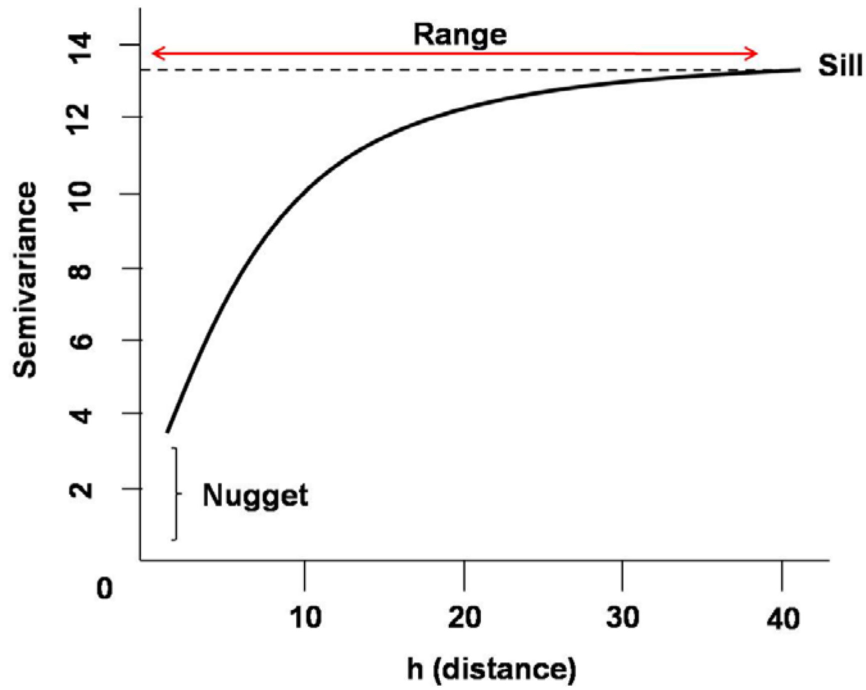


Figure 8. Idealized semivariogram showing the range and nugget. The range, which occurs where the graph levels off (i.e., sill), reveals the greatest distance at which the data are autocorrelated. The nugget, or the distance from the origin of the graph to the beginning of the data line, expresses the degree of variability due to local random effects or measurement errors.

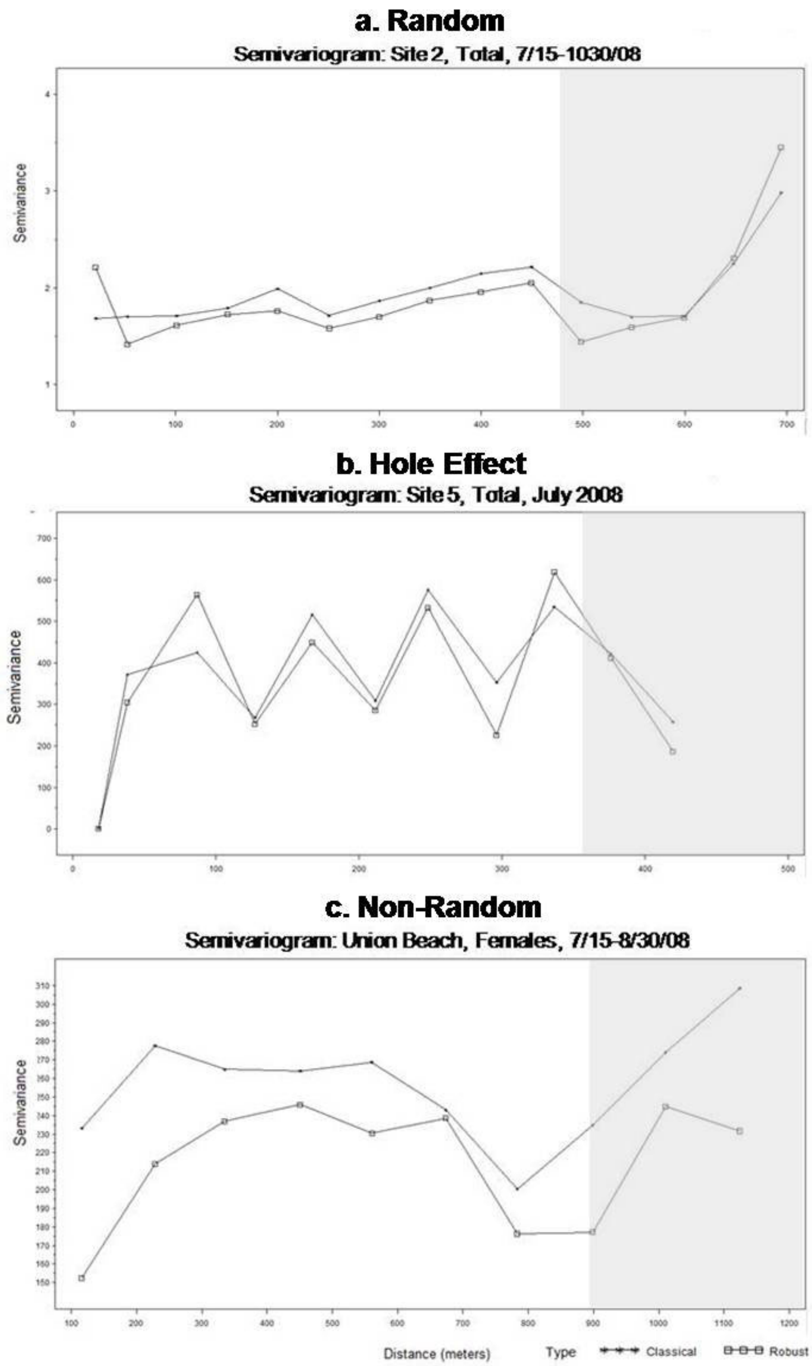


Figure 9. Examples of the three main patterns observed in the semivariograms (see Appendix 1 for complete catalog of semivariograms). Random graphs (a) showed no perceptible pattern or sill. The ‘hole effect’ (b), indicated by oscillation in the graphs, is present in several monthly semivariograms. This pattern reflects either extreme values, periodicity in the spatial variability of the errors, or preferential clustering in the data (Bailey and Gatrell 1995). Non-random semivariograms leveled off into a sill in the first two-thirds of the graph, which indicates the range of autocorrelation.

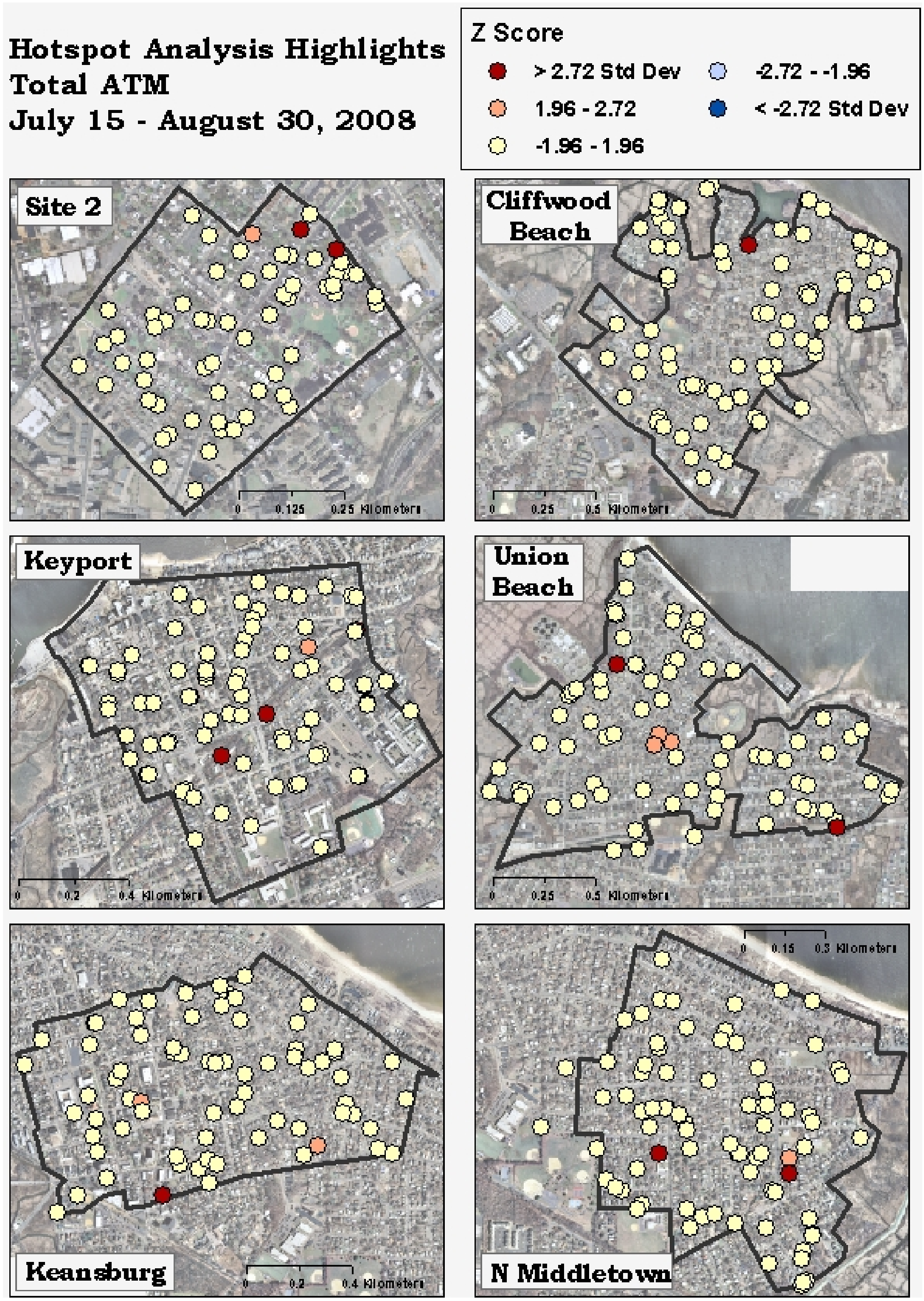


Figure 10. Hot spot analysis results for total ATM counts during the high season (July 15-August 30). Sites 3, 5 and 7 in Mercer County are not shown since few hot spots were present in each site (see Appendix 2 for complete catalog of hot spot maps).

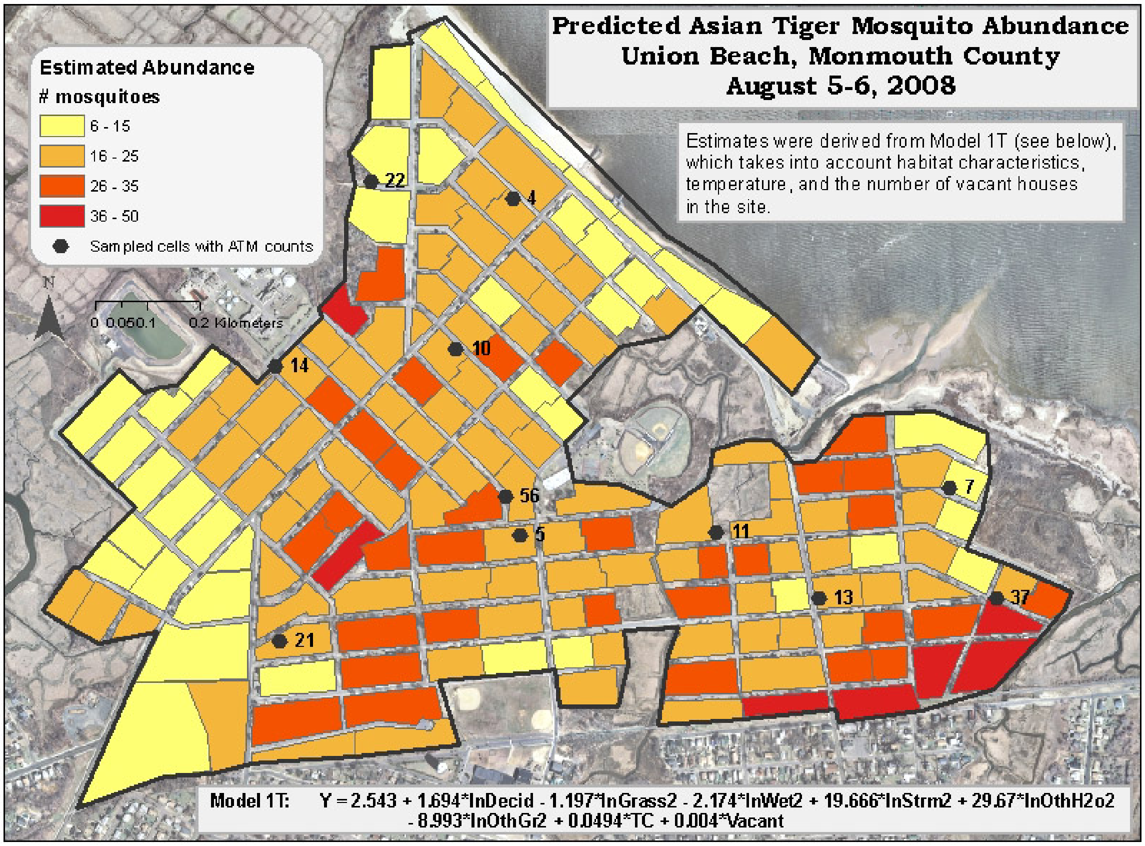


Figure 11. Sample predictive map of ATM abundance in Union Beach, Monmouth County, for the first week of August 2008. The map was created based on Model 1T, which is shown in the map.

BIBLIOGRAPHY

Allison, P.D. 1999. *Multiple regression*. Thousand Oaks, CA: Pine Forge Press.

Alto, B.W. and S.A. Juliano. 2001. Precipitation and Temperature Effects on Populations of *Aedes albopictus* (Diptera: Culicidae): Implications for Range Expansion. *Journal of Medical Entomology* 38(5):646-656.

Barker, C.M., S.L. Paulson, S. Cantrell, B.S. Davis. 2003. Habitat preferences and phenology of *Ochlerotatus triseriatus* and *Aedes albopictus* (Diptera: Culicidae) in Southwestern Virginia. *Journal of Medical Entomology* 40(4):403-410.

Bailey, T.C., and A. Gatrell. 1995. *Interactive spatial data analysis*. Prentice Hall.

Belsley, D.A., E. Kuh and R.E. Welsch (1980). *Regression Diagnostics: Identification Influential Data and Sources of Collinearity,* New York: John Wiley & Sons, Inc.

Benedict, M.Q., R.S. Levine, W.A. Hawley, L.P. Lounibos. 2007. Spread of the tiger: Global risk of invasion by the mosquito *Aedes albopictus*. *Vector-Borne and Zoonotic Diseases* 7(1):76-85.

Benz, U.C., P. Hofmann, G. Willhauck, I. Lingenfelder, and M. Heynen. 2004. Multi resolution object oriented fuzzy analysis of remote sensing data for GIS ready information. *ISPRS Journal of Photogrammetry and Remote Sensing* 58:239-258.

Braks, M.A.H., N.A. Honorio, R. Lourenço-De-Oliveira, S.A. Juliano, L.P. Lounibos. 2003. Convergent habitat segregation of *Aedes aegypti* and *Aedes albopictus* (Diptera: Culicidae) in Southeastern Brazil and Florida. *Journal of Medical Entomology* 40(6):785-794.

Brown, H.E., M.A. Diuk-Wasser, Y. Guan, S. Caskey, D. Fish. 2008. Comparison of three satellite sensors at three spatial scales to predict larval mosquito presence in Connecticut wetlands. *Remote Sensing of Environment* 112: 2301–2308.

Centers for Disease Control (CDC). 2001. Information on *Aedes albopictus*, *Arboviral Encephalitides.* URL (accessed 15 June 2009) <http://www.cdc.gov/ncidod/dvbid/arbor/albopic_new.htm>.

Congalton, R. 1991. A review assessing the accuracy of classification of remotely sensed data. *Remote Sensing of the Environment* 37:35-46.

Crans, W.J., M.S. Chomsky, D. Guthrie, and A. Acquaviva. 1996. First record of Aedes albopictus from New Jersey. *Journal of the American Mosquito Control Association* 12(2):307-309.

Cressie, N.A.C., 1993. *Statistics for Spatial Data*. Wiley: New York.

Day, .J.F. and J. Shaman. 2008. Using hydrologic conditions to forecast the risk of focal and epidemic arboviral transmission in peninsular Florida. Journal of medical entomology 45(3): 458-465.

Definiens Imaging. 2006. *eCognition User Guide*, version 5.0.6.1. Definiens AG: München, Germany. URL (accessed 1 December 2009) <http://www.pcigeomatics.com/products/ecog_applications/ecog5_UserGuide.pdf>.

DeGaetano, A.T. 2005. Meteorological effects on adult mosquito (*Culex*) populations in metropolitan New Jersey. *International Journal of Biometeorology* 49:345-353.

Fortin, M. and M.R.T. Dale. 2005. *Spatial Analysis: A Guide for Ecologists*. University Press: Cambridge.

Francy, D.B., N. Karabatsos, D.M. Wesson, C.G. Moore Jr, J.S. Lazuick, M.L. Niebylski, T.F. Tsai, and G.B. Craig Jr. 1990. A new arbovirus from Aedes albopictus, an Asian mosquito established in the United States. *Science* 250(4988):1738-1740.

Geneletti, D., and B.G.H. Gorte. 2003. A method for object-oriented land cover classification combining Landsat TM data and aerial photographs. *International Journal of Remote Sensing* 24(6):1273-1286.

Gillett, J.D. 1972. *The Mosquito: Its Life, Activities, and Impact on Human Affairs*. Doubleday: New York.

Grieco, J.P., S. Johnson, N.L. Achee, P. Masuoka, K. Pope, E. Rejmankova, E. Vanzie, R. Andre, and D. Roberts. 2006. Distribution of *Anopheles albimanus*, *Anopheles vestitipennis*, and *Anopheles crucians* associated with land use in Northern Belize. *Journal of Medical Entomology* 43(3): 614-622.

Haralick, R.M. and L.G. Shapiro. 1985. Image segmentation techniques. *Computer Vision Graphics and Image Processing* 29:100-132.

Hawley, W.A. 1988. The biology of Aedes albopictus. *Journal of American Mosquito Control Association* 4(Suppl.):1-40.

Hay, S.I. and J.J. Lennon. 1999. Deriving meteorological variables across Africa for the study and control of vector-borne disease: a comparison of remote sensing and spatial interpolation of climate. *Tropical Medicine and International Health* 4(1):58-71.

Hohn, M.E. 1998. *Geostatistics and Petroleum Geology.* Longman Scientific & Technical: Harlow.

Invasive Species Specialist Group (ISSG). 2009. Global Invasive Species Database: 100 of the World's Worst Invasive Alien Species. IUCN Species Survival Commission <http://www.issg.org/database/species/search.asp?st=100ss&fr=1&str=&lang=EN> Accessed 1 Aug 2009.

Kalluri, S., P. Gilruth, D. Rogers, M. Szczur. 2007. Surveillance of arthropod vector-borne infectious diseases using remote sensing techniques: A review. *PLoS Pathogens* 3(10):1361-1371.

Kitron, U., J. Swanson, M. Crandell, P.J. Sullivnan, J. Anderson, R. Garro, L.D. Haramis, P.R. Grimstad. 1998. Introduction of Aedes albopictus into a La Crosse Virus-Enzoontic site in Illinois. *Emerging Infectious Diseases* 4(4):627-630.

Kobayashi, M, N. Nihei, T. Kurihara. 2002. Analysis of northern distribution of *Aedes albopictus* (Diptera: Culicidae) in Japan by geographical information system. *Journal of Medical Entomology* 39:4–11.

Kolivras, K.N. 2006. Mosquito habitat and dengue risk potential in Hawaii: A conceptual framework and GIS application. *The Professional Geographer* 58(2):139–154.

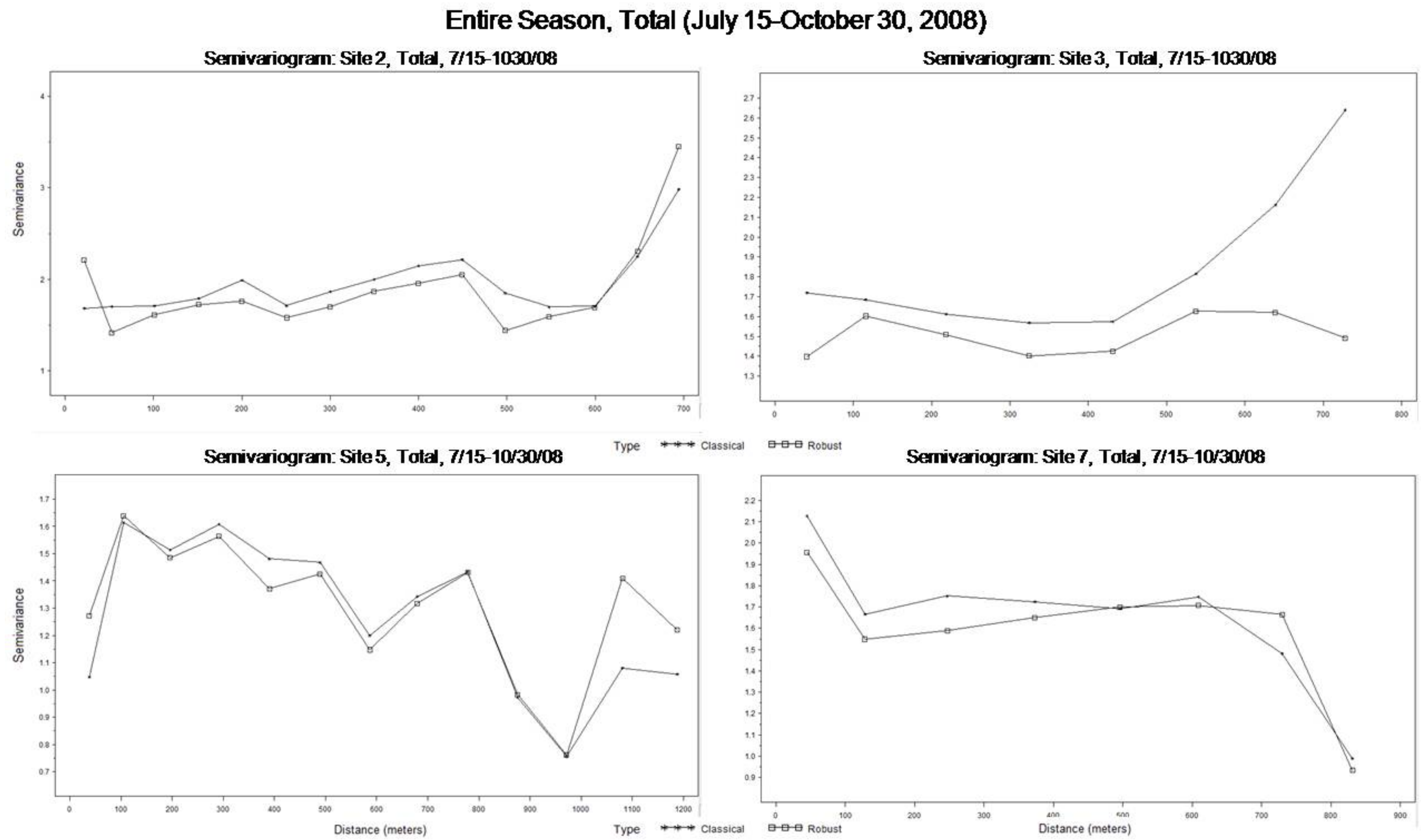
Levine, N. 2004. *CrimeStat III: A Spatial Statistics Program for the Analysis of Crime Incident Locations*. Ned Levine & Associates, Houston, TX, and the National Institute of Justice, Washington, DC.

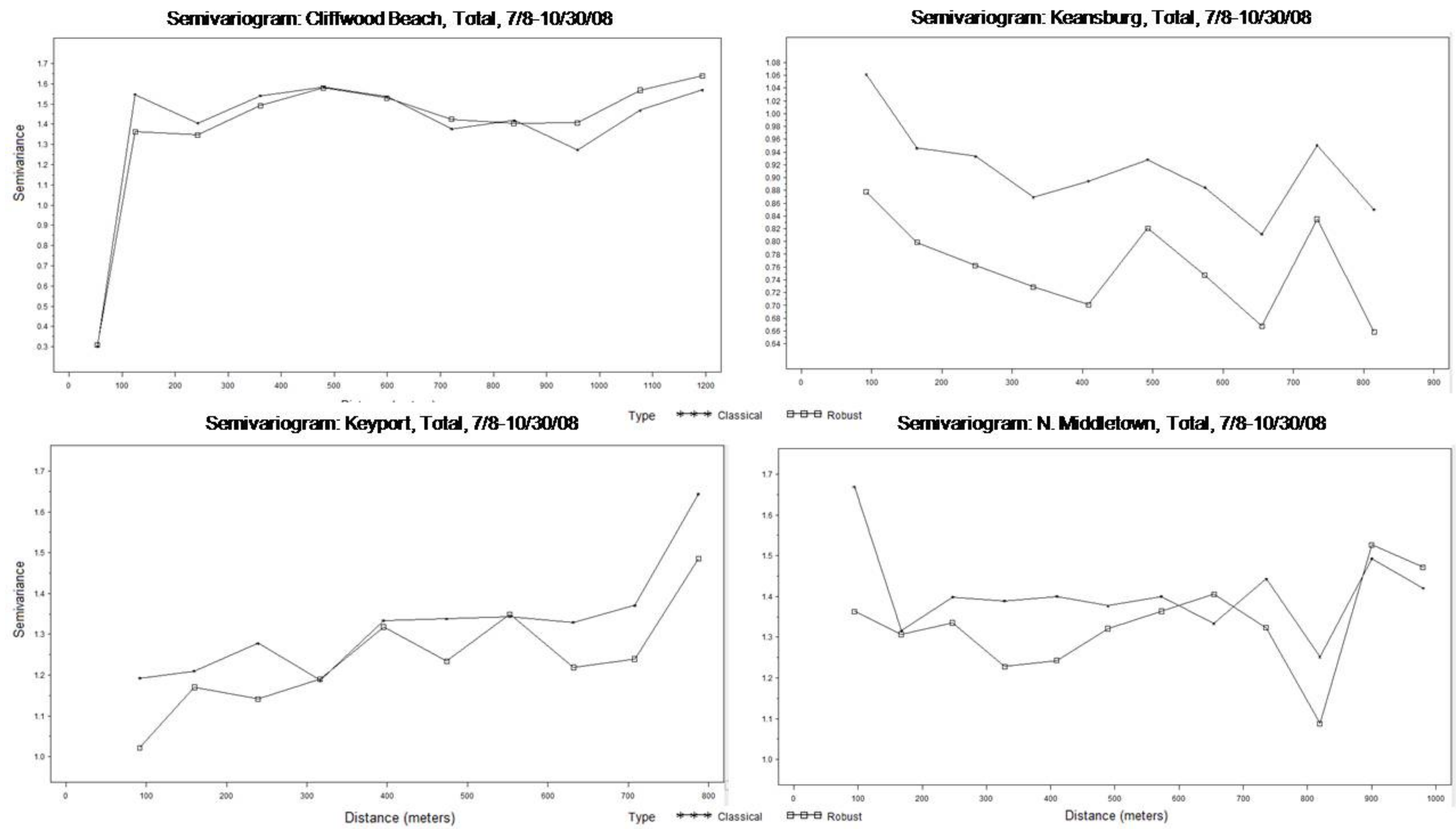
Levine, R.S., A.T. Peterson, M.Q. Benedict. 2004. Geographic and ecological distributions of the *Anopheles gambiae* complex members predicted using a genetic algorithm. *American Journal of Tropical Medicine and Hygiene* 70:105–

109.
Linthicum, K.J., C.L. Bailey, F.G. Davies, C.J. Tucker. 1987. Detection of Rift Valley Fever viral activity in Kenya by satellite remote sensing imagery. *Science* 235(4796):1656-1659.
López-Cárdenas, J., F.E.G. Bravo, P.M.S. Schettino, J.C.G. Solorzano, E.R. Barba, J.M. Mendex, V. Sanchez-Cordero, A.T. Peterson, J.M. Ramsey. 2006. Fine-scale predictions of distribution of Chagas Disease vectors in the State of Guanajuato, Mexico. *Journal of Medical Entomology* 42(6):1068-1081.
Moffett, A., N. Shackelford, and S. Sarkar. 2007. Malaria in Africa: Vector species' niche models and relative risk maps. *PLoS ONE* 2(9): e824.
Moore, C.G. and C.J. Mitchell. 1997. Aedes albopictus in the United States: Ten-year presence and public health implications. *Emerging Infectious Diseases* 3(3):329-334.
Mushinzimana, E., S. Munga, N. Minakawa, L. Li, C. Feng, L. Bian, U. Kitron, C. Schmidt, L. Beck, G. Zhou, A.K. Githeko, G. Yan. 2006. Landscape determinants and remote sensing of anopheline mosquito larval habitats in the western Kenya highlands. *Malaria Journal* 5(13).
Navulur, K. 2007. *Multispectral Image Analysis Using the Object-Oriented Paradigm.* CRC Press: Florida.
Nawrocki, S.J., and W.A. Hawley. 1987. Estimation of the northern limits of distribution of *Aedes albopictus* in North America. *Journal of the American Mosquito Control Association* 3:314–317.
New Jersey Weather and Climate Network (NJWXNET). 2008. Homepage. New Jersey Agricultural Experiment Station, Rutgers University. URL (accessed 27 October): <http://climate.rutgers.edu/njwxnet/>.
Ostfeld, R.S., G.E. Glass, F. Keesing. 2005. Spatial epidemiology: an emerging (or re-emerging) discipline. *Trends in Ecology and Evolution* 20(6):328-336.
Peterson A.T. 2006. Ecological niche modelling and spatial patterns of disease transmission. *Emerging Infectious Diseases* 12: 1822–1826.
Porphyre, T., D.J. Bicout, P. Sabatier. 2005. Modelling the abundance of mosquito vectors versus flooding dynamics. *Ecological Modelling* 183:173-181.
Rai, K.S. 1991. *Aedes albopictus* in the Americas. *Annual Review of Entomology* 36:459-484.
Reinert, W.C. 1989. The New Jersey light trap: an old standard for most mosquito control programs. *Proceedings of the seventy-sixth annual meeting of the New Jersey Mosquito Control Association* 17-25.
Rey, J.R., N. Nishimura, B. Wagner, M.A.H. Braks, S.M. O'Connell, L.P. Lounibos. 2006. Habitat segregation of mosquito arbovirus vectors in South Florida. *Journal of Medical Entomology* 43(6):1134-1141.
Richards, S.L., C.S. Apperson, S.K. Ghosh, H.M. Cheshire, B.S. Zeichner. 2006. Spatial analysis of *Aedes albopictus* (Diptera: Culicidae) oviposition in suburban neighborhoods of a Piedmont community in North Carolina. *Journal of Medical Entomology* 43(5):976-989.
Rochlin, I., T. Iwanejko, M.E. Dempsey, and D.V. Ninivaggi. 2009. Geostatistical evaluation of integrated marsh management impact on mosquito vectors using before-after-control-impact (BACI) design. *International Journal of Health*

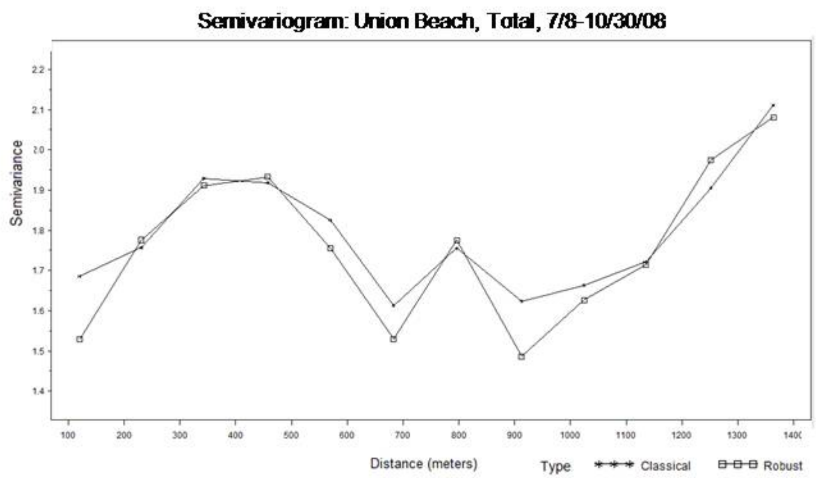
Geographics 8:35.
Rogers, D.J., S.E. Randolph, R.W. Snow, S.I. Hay. 2002. Satellite imagery in the study and forecast of malaria. *Nature* 415:710-715.
Shaman, J., M. Stieglitz, C. Stark, S. Le Blancq, M. Cane. 2002. Using a dynamic hydrology model to predict mosquito abundances in flood and swamp water. *Emerging Infectious Diseases* 8(1):6-13.
Sota, T., M. Mogi, AND E. Hayamizu Seasonal Distribution and Habitat Selection by *Aedes albopictus* and *Ae. riversi* (Diptera: Culicidae) in Northern Kyushu, Japan. *Journal of Medical Entomology* 29(2):296-304.
Spielman, A., and M. D'Antonio. 2001. *Mosquito: A Natural History of Our Most Persistent and Deadly Foe*. Hyperion: New York.
Tabachnick, B.G. and L.S. Fidell. 2001. *Using Multivariate Statistics*. Fourth Edition. Boston: Allyn and Bacon.
Teng, H-J. and C. S. Apperson. 2000. Development and survival of immature *Aedes albopictus* and *Aedes triseriatus* (Diptera: Culicidae) in the laboratory effects of density, food, and competition in response to temperature. *Journal of Medical Entomology* 37:40-52.
Tobler W. 1970. A computer movie simulating urban growth in the Detroit region. *Economic Geography* 46(2): 234-240.
Tran, A., N. Ponçon, C. Toty, C. Linard, H. Guis, J. Ferré5, D.L. Seen, F. Roger, S. de la Rocque, D. Fontenille, and T. Baldet. 2008. Using remote sensing to map larval and adult populations of *Anopheles hyrcanus* (Diptera: Culicidae) a potential malaria vector in Southern France. *International Journal of Health Geographics* 7:9.
Tsuda, Y., W. Suwonkerd, S. Chawprom, S. Prajakwong, and M. Takagi. 2006. Different spatial distribution of *Aedes aegeypti* and *Aedes albopictus* along an urban-rural gradient and the relating environmental factors examined in three villages in northern Thailand. Journal of the American Mosquito Control Association 22(2):222-228.
Turner, M.G. 1989. Landscape ecology: the effect of pattern on process. *Annual Review Ecology and Systematics* 20:171-197.
———, Gardner, R.H., O'Neill, R.V. 2001. *Landscape Ecology in Theory and Practice: Pattern and Process*. New York: Springer.
United States (US) Census Bureau. 2009. American Factfinder Glossary. URL (accessed 15 September 2009) <http://factfinder.census.gov/home/en/epss/glossary_v.html >.
Vanwambeke, S.O., E.F. Lambin, M.P. Eichhorn, S.P. Flasse, R.E. Harbach, L. Oskam, P. Somboon, S. van Beers, B.H.B. van Benthem, C. Walton, and R.K. Butlin. 2007. Impact of land-use change on dengue and malaria in Northern Thailand. *EcoHealth* 4: 37-51.
Yan, G., J.F. Mas, B.H.P. Maathuis, Z. Xiangmin, and P.M. Van Dijk. 2006. Comparison and pixel-based and object-oriented image classification approaches—a case study in a coal fire area, Wuda, Inner Mongolia, China. *International Journal of Remote Sensing* 27(18):4039-4055.

APPENDIX 1: SEMIVARIOGRAMS

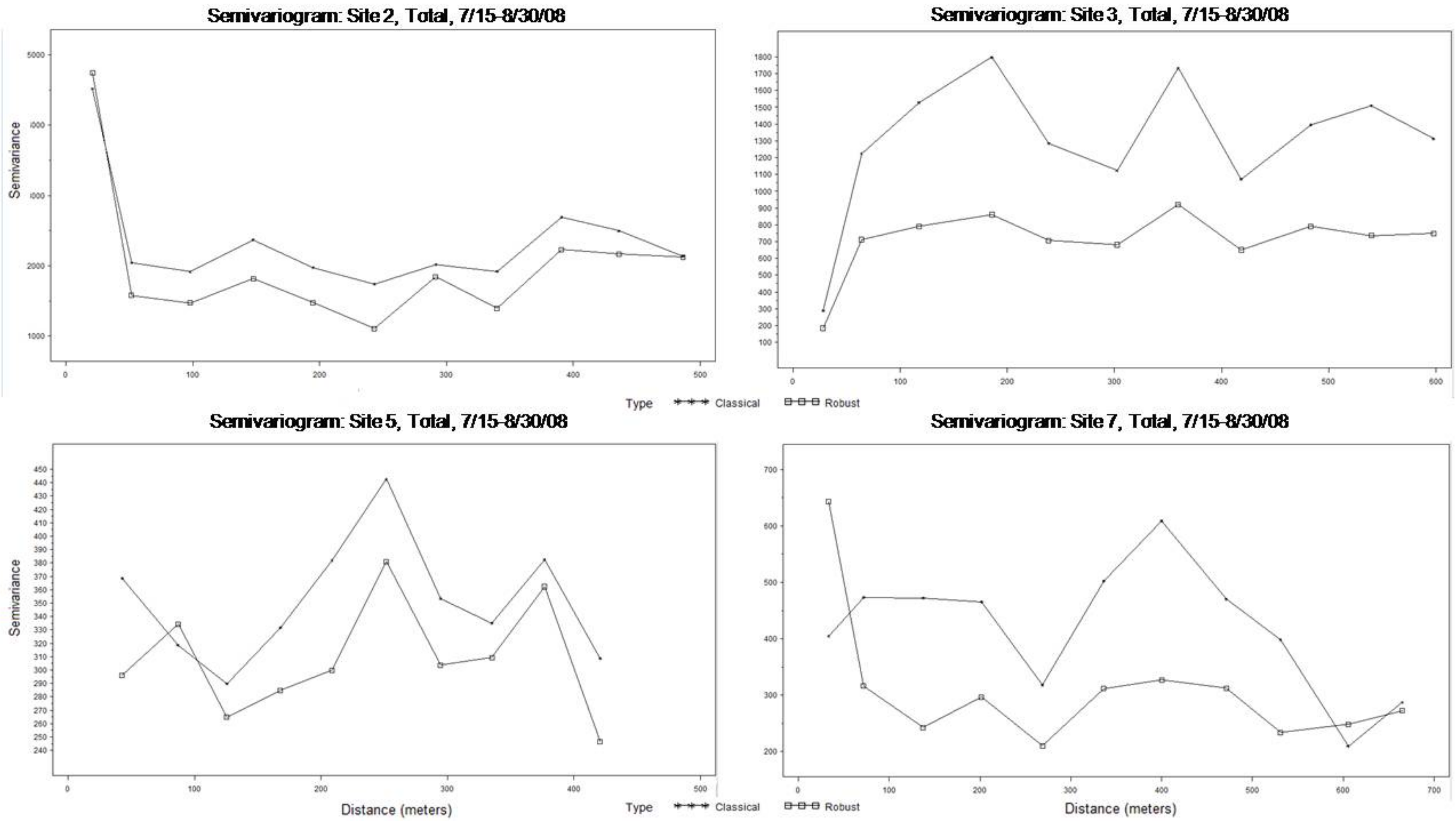


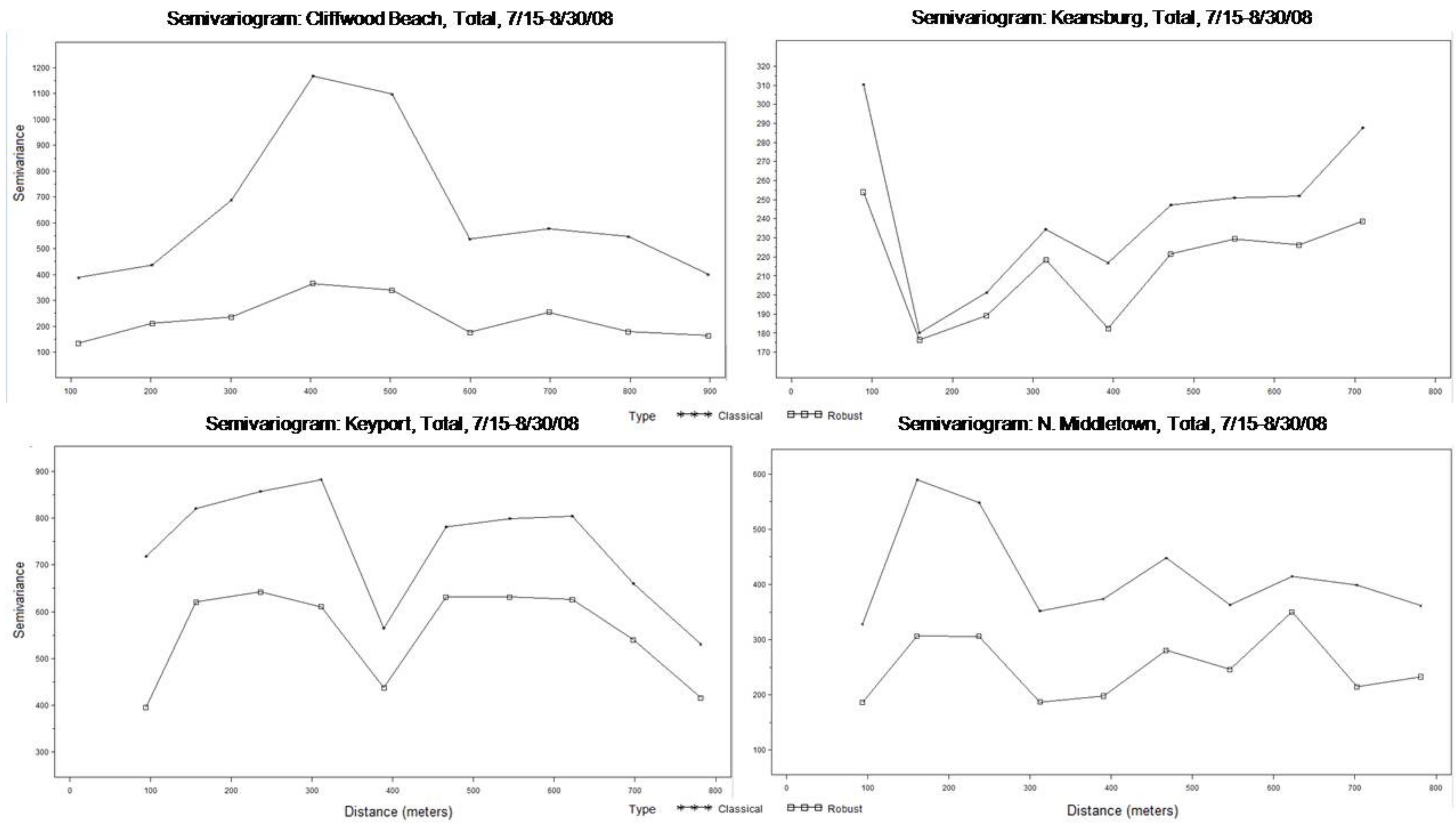


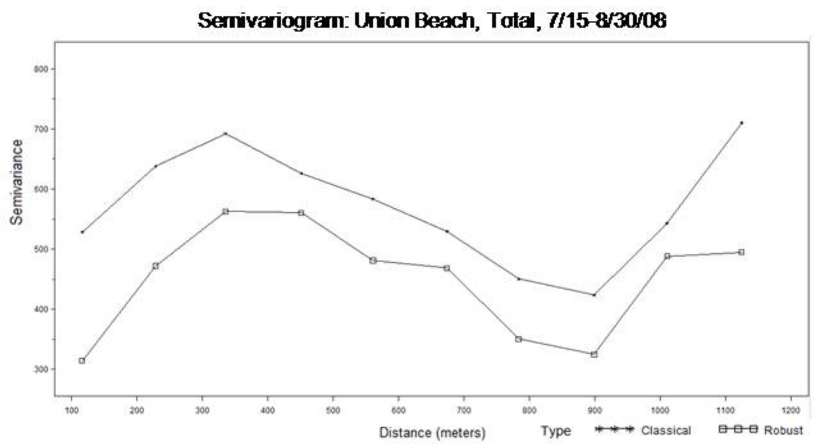
Semivariogram: Cliffwood Beach, Total, 7/8-10/30/08
Semivariogram: Keansburg, Total, 7/8-10/30/08
Semivariogram: Keyport, Total, 7/8-10/30/08
Semivariogram: N. Middletown, Total, 7/8-10/30/08
Semivariance
Distance (meters)
Type
Classical
Robust


Semivariogram: Union Beach, Total, 7/8-10/30/08
Semivariance
Distance (meters)
Type
Classical
Robust

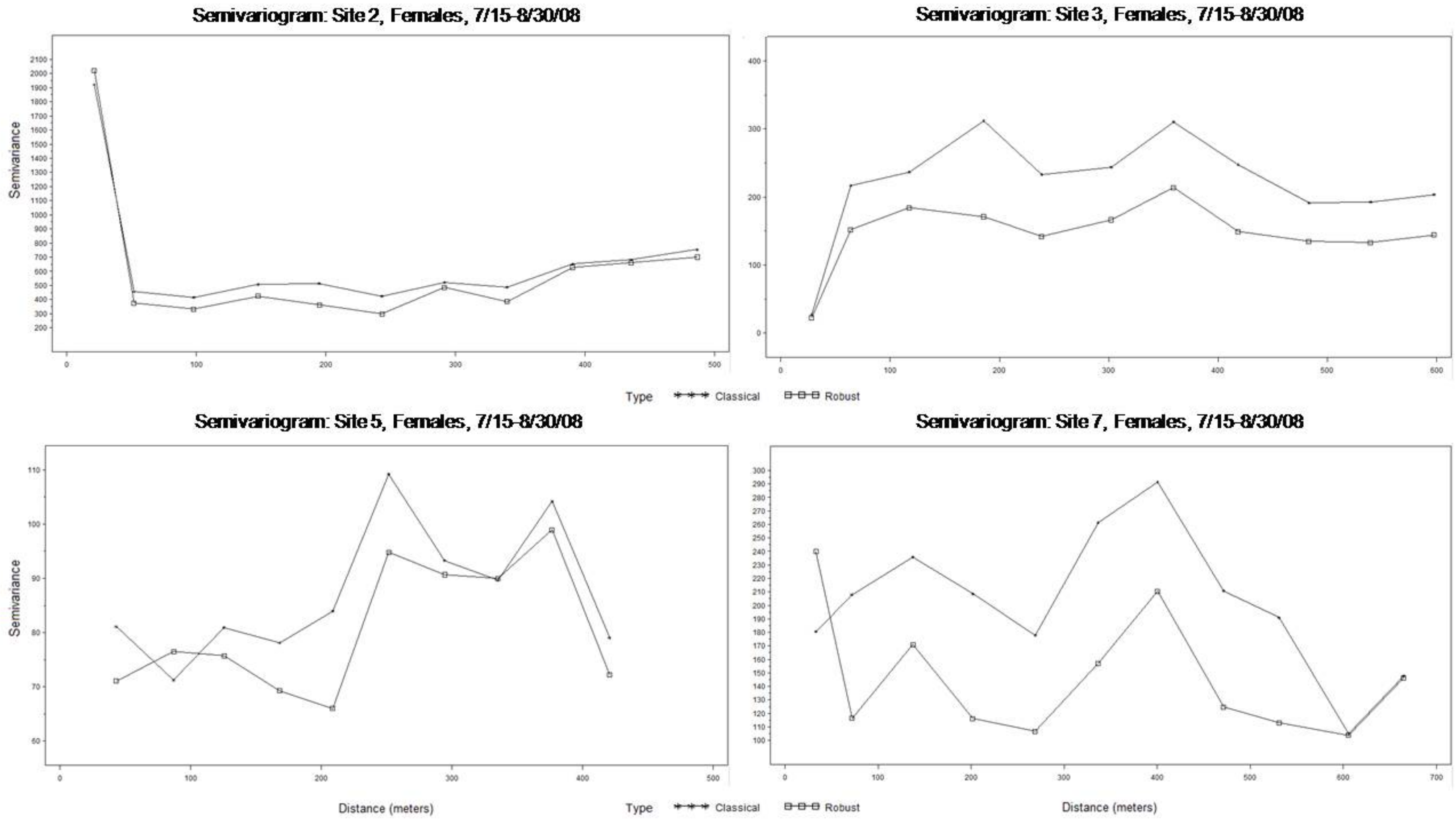
High Season, Total (July 15-August 30, 2008)

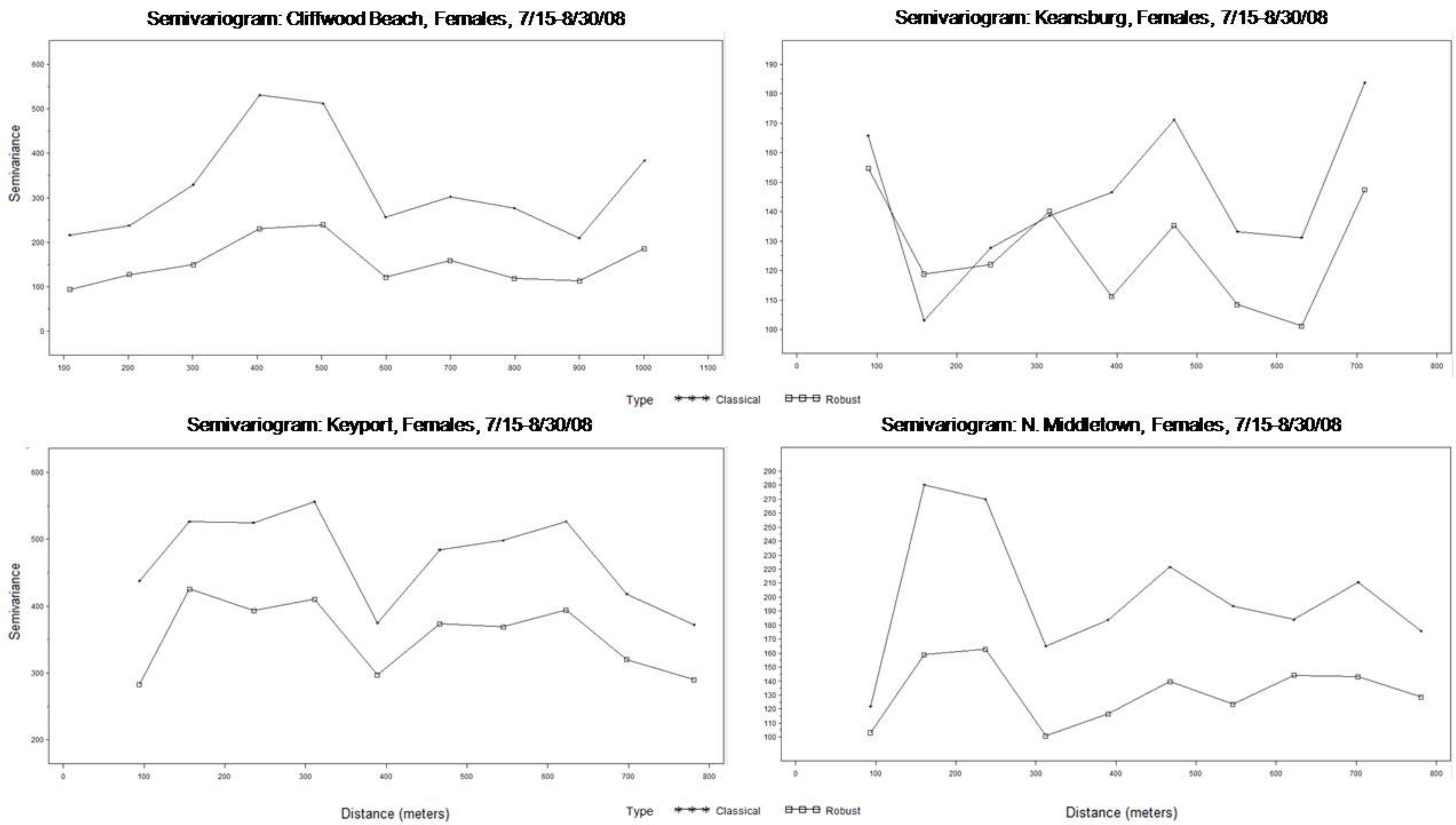


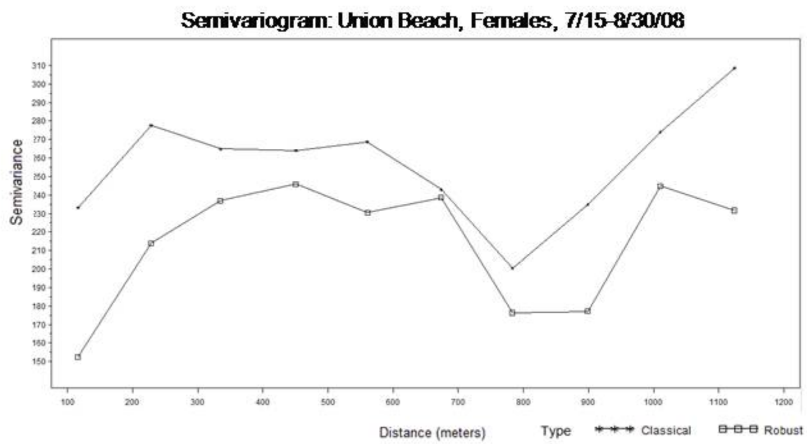

Semivariogram: Cliffwood Beach, Total, 7/15-8/30/08
Semivariogram: Keansburg, Total, 7/15-8/30/08
Semivariogram: Keyport, Total, 7/15-8/30/08
Semivariogram: N. Middletown, Total, 7/15-8/30/08
Semivariance
Distance (meters)
Type
Classical
Robust


Semivariogram: Union Beach, Total, 7/15-8/30/08
Semivariance
Distance (meters)
Type
Classical
Robust

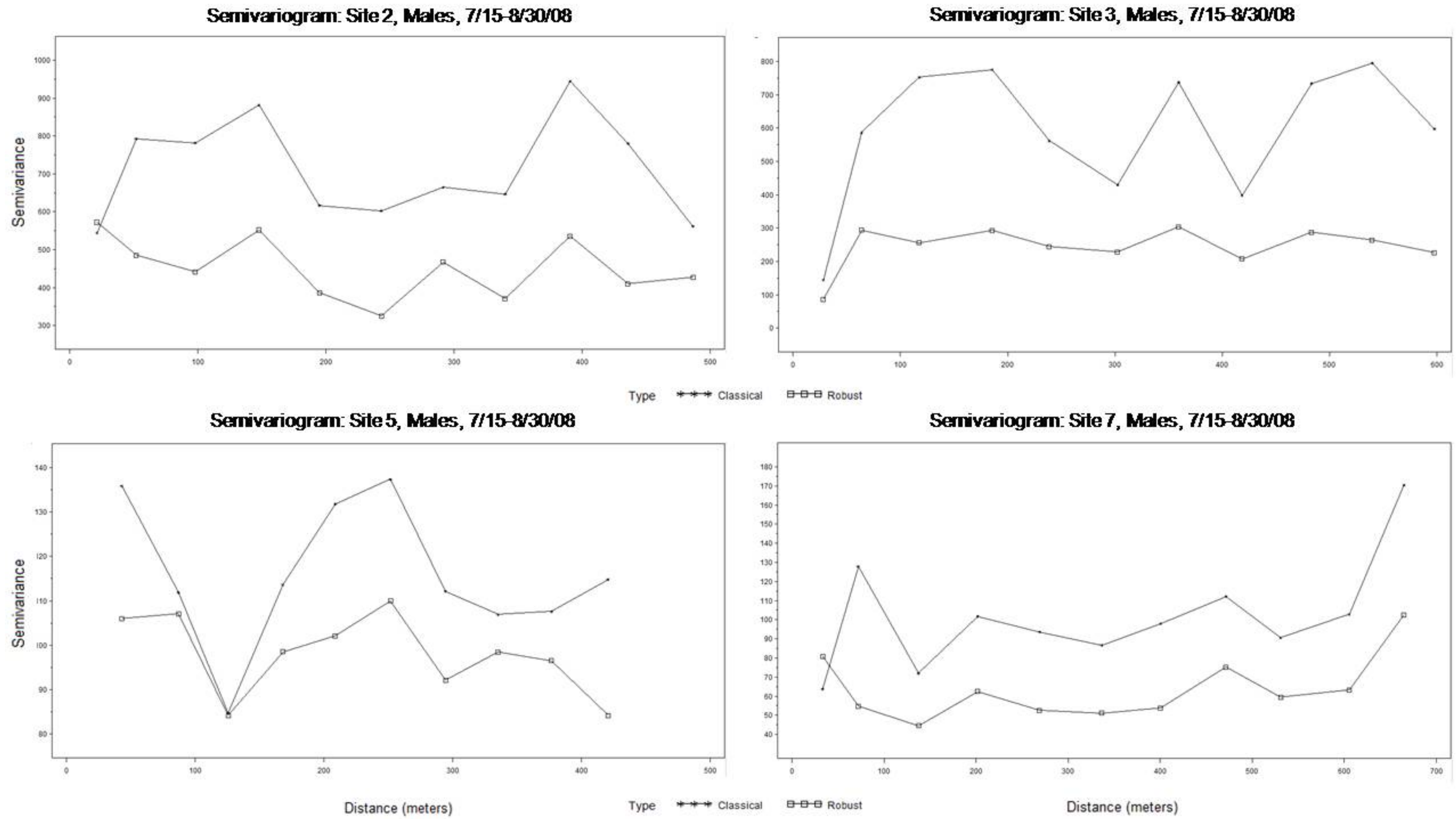
High Season, Females (July 15-August 30, 2008)

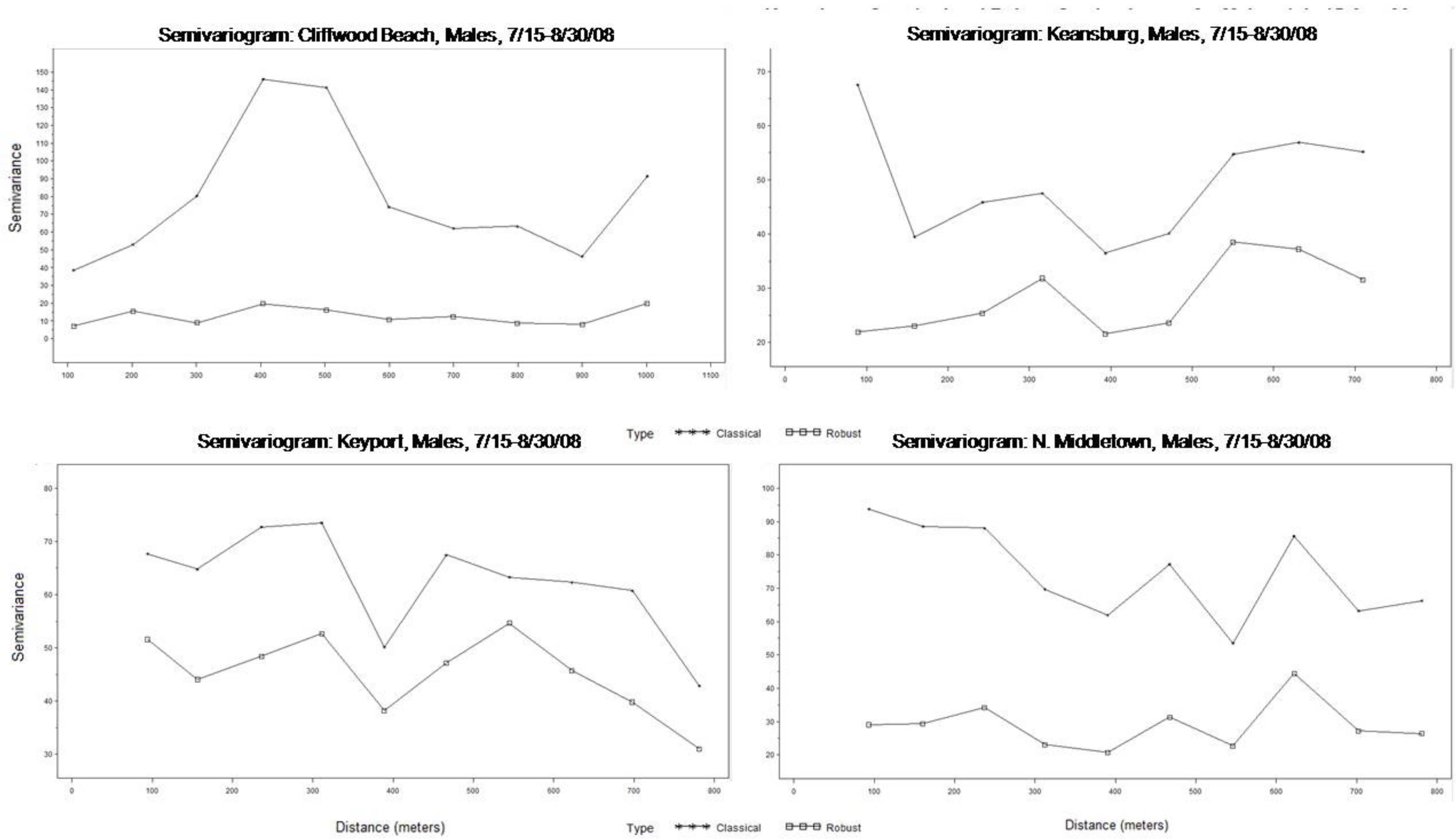


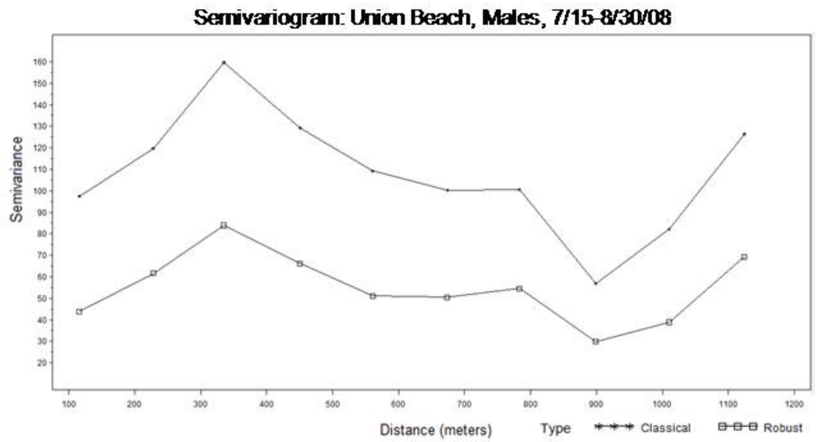

Semivariogram: Cliffwood Beach, Females, 7/15-8/30/08
Semivariogram: Keansburg, Females, 7/15-8/30/08
Semivariance
Type
Classical
Robust
Semivariogram: Keyport, Females, 7/15-8/30/08
Semivariogram: N. Middletown, Females, 7/15-8/30/08
Semivariance
Distance (meters)
Type
Classical
Robust
Distance (meters)


Semivariogram: Union Beach, Females, 7/15-8/30/08
Semivariance
Distance (meters)
Type
Classical
Robust

High Season, Males (July 15-August 30, 2008)

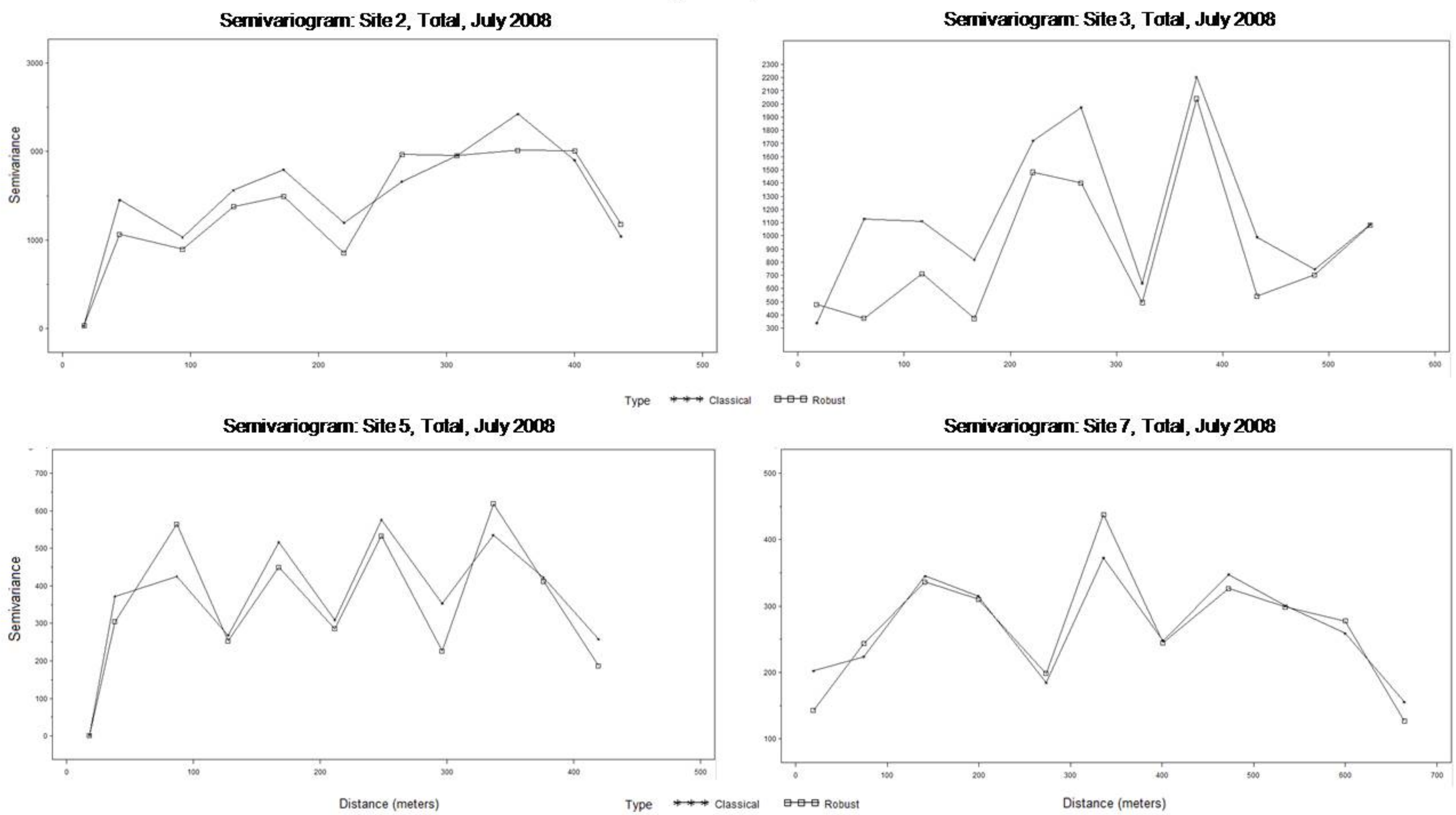


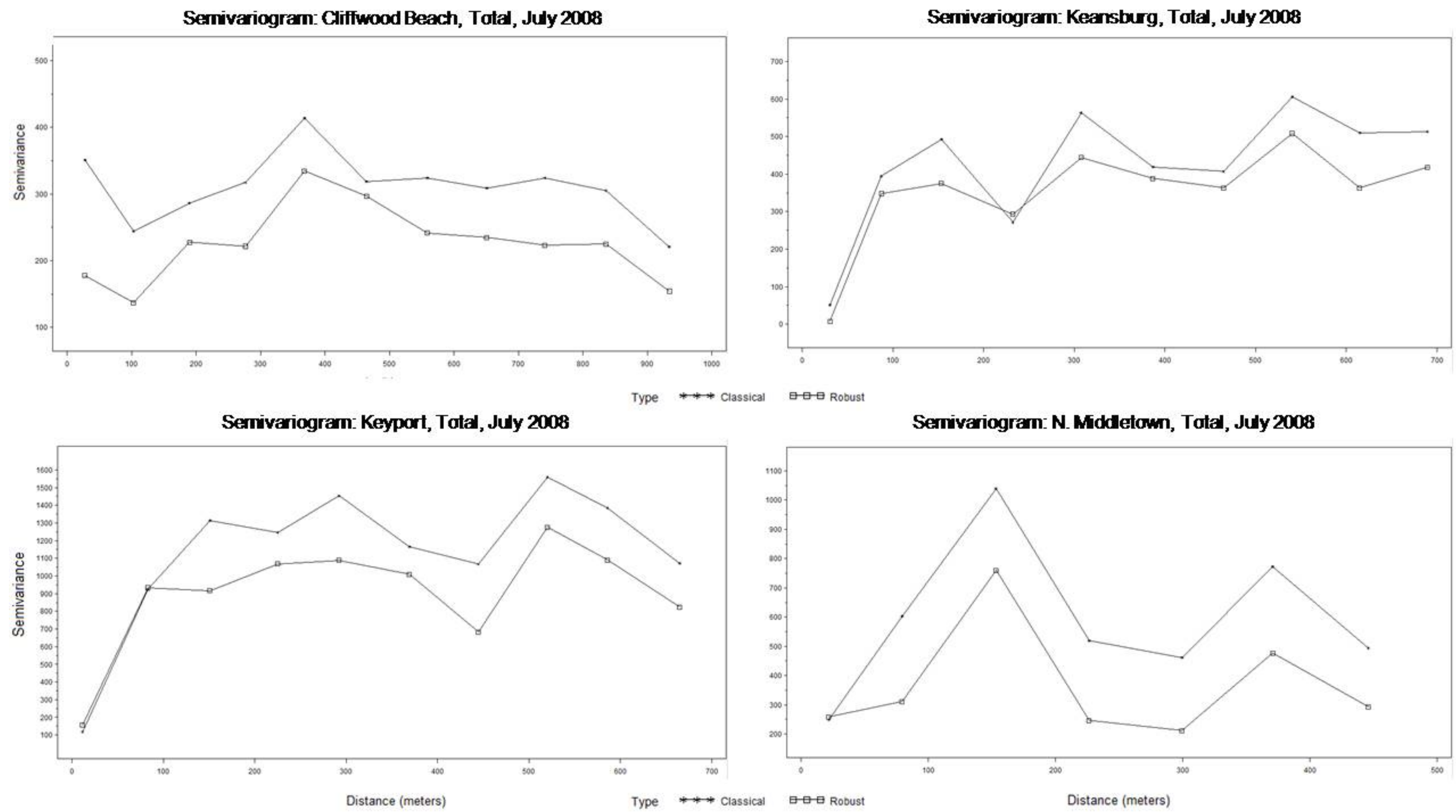

Semivariogram: Cliffwood Beach, Males, 7/15-8/30/08
Semivariogram: Keansburg, Males, 7/15-8/30/08
Semivariogram: Keyport, Males, 7/15-8/30/08
Semivariogram: N. Middletown, Males, 7/15-8/30/08
Semivariance
Distance (meters)
Type
Classical
Robust

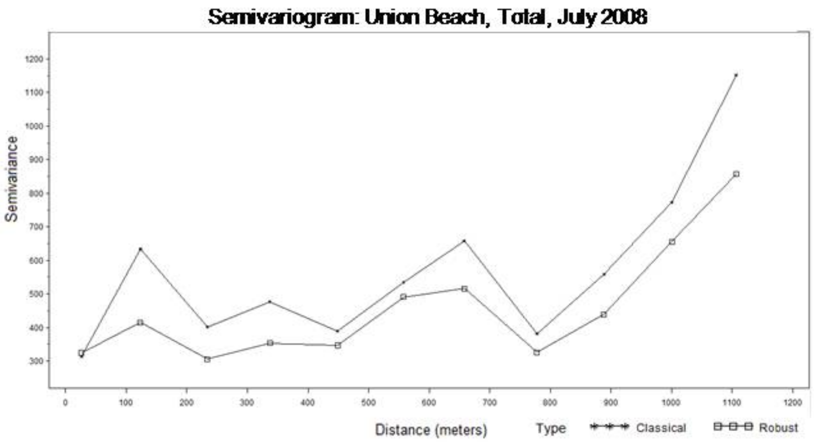


Semivariogram: Union Beach, Males, 7/15-8/30/08
Semivariance
Distance (meters)
Type
Classical
Robust

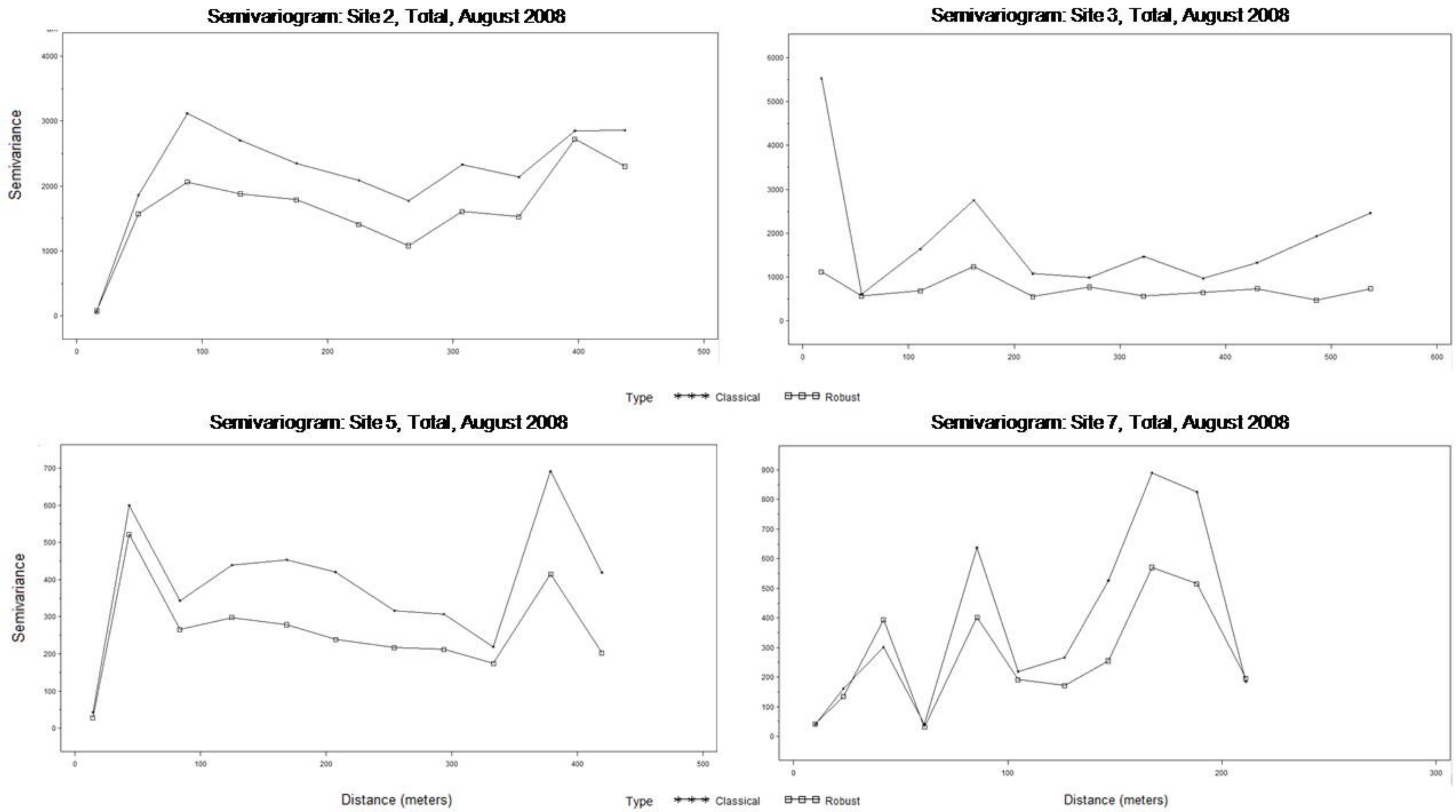
July 2008, Total

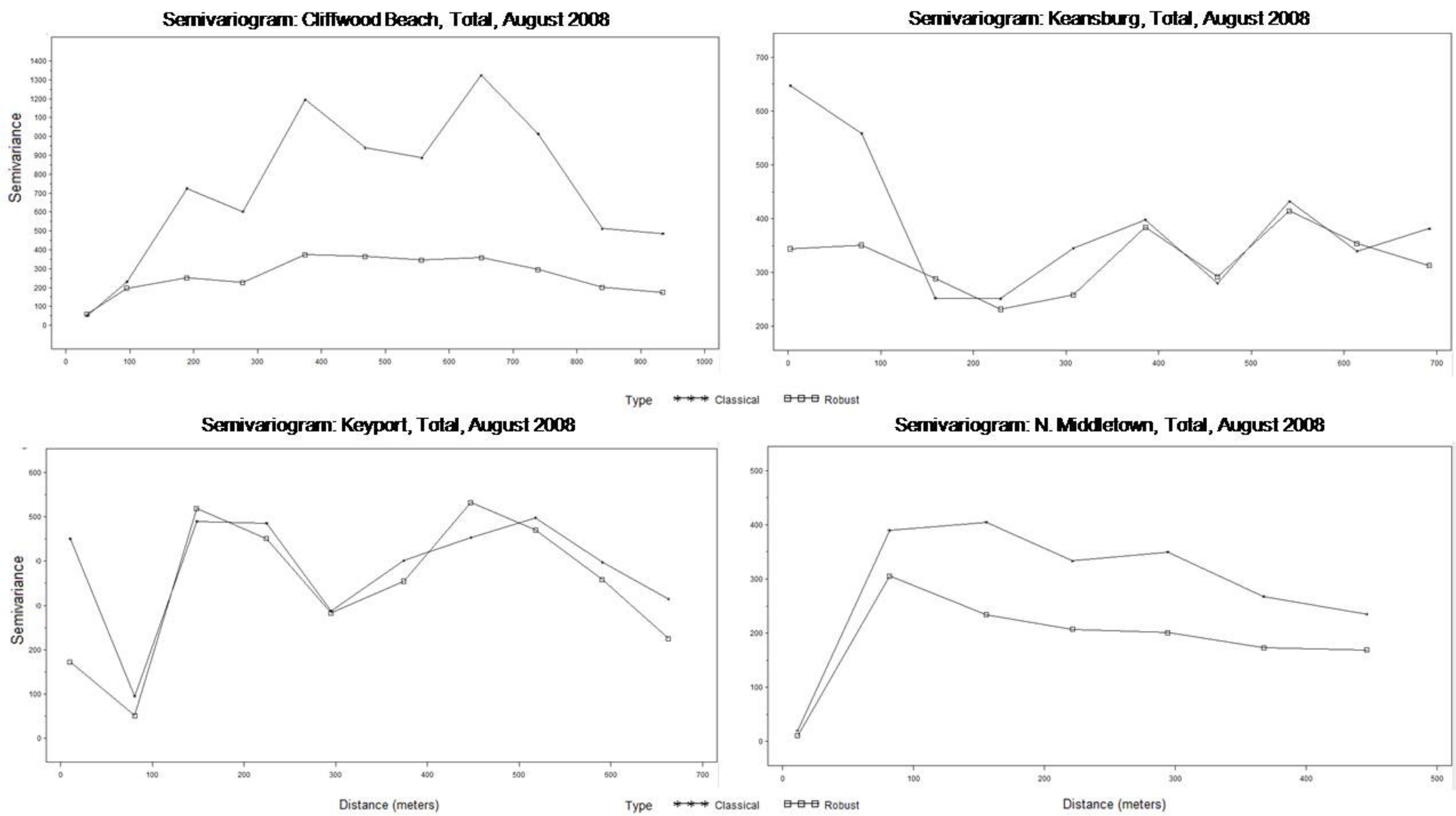


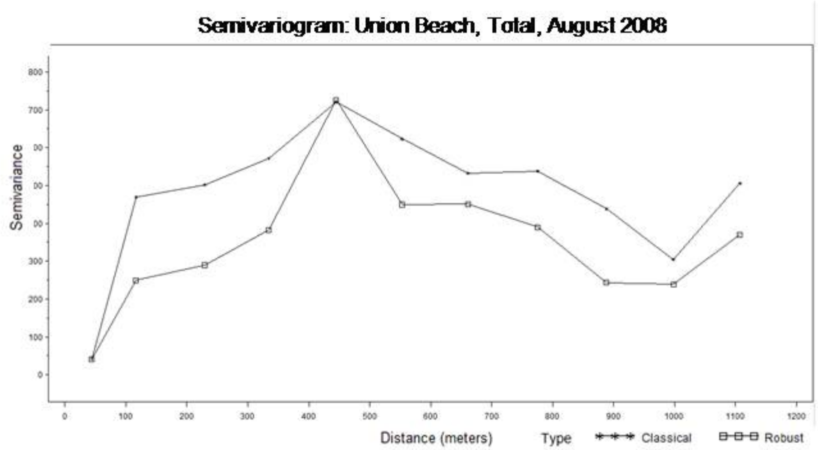

Semivariogram: Cliffwood Beach, Total, July 2008
Semivariogram: Keansburg, Total, July 2008
Semivariance
Type
Classical
Robust
Semivariogram: Keyport, Total, July 2008
Semivariogram: N. Middletown, Total, July 2008
Semivariance
Distance (meters)
Distance (meters)
Type
Classical
Robust


Semivariogram: Union Beach, Total, July 2008
Semivariance
Distance (meters)
Type
Classical
Robust

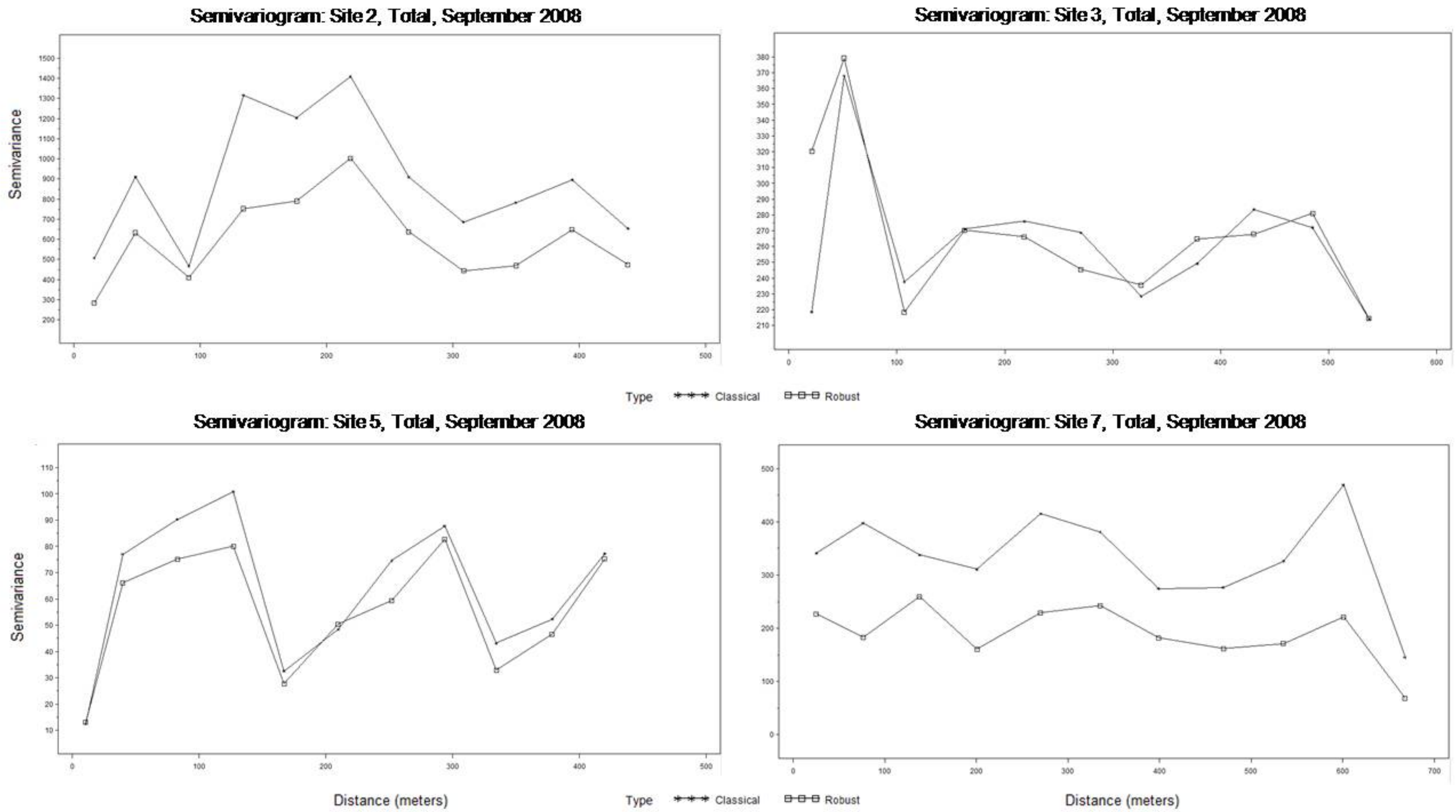
August 2008, Total

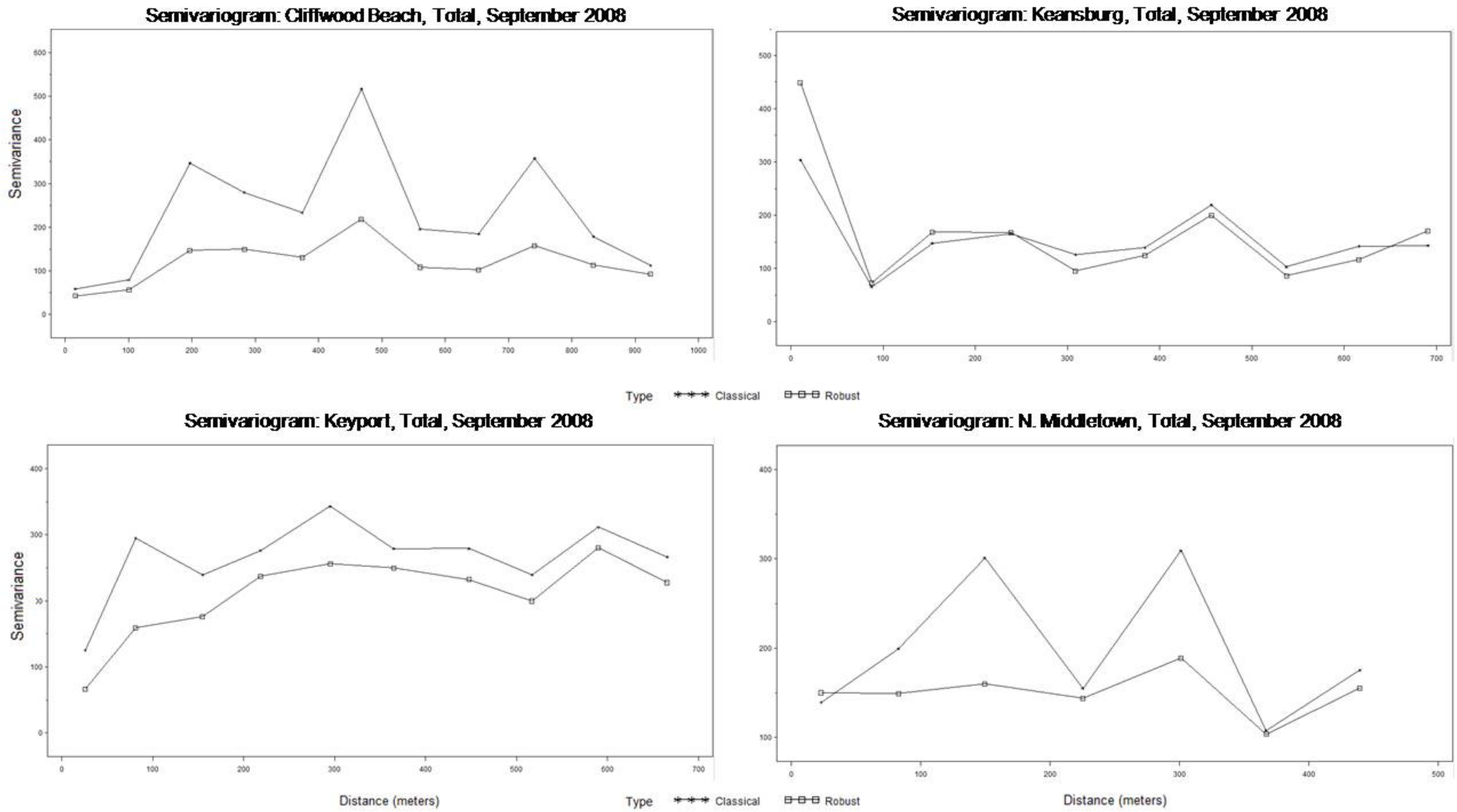


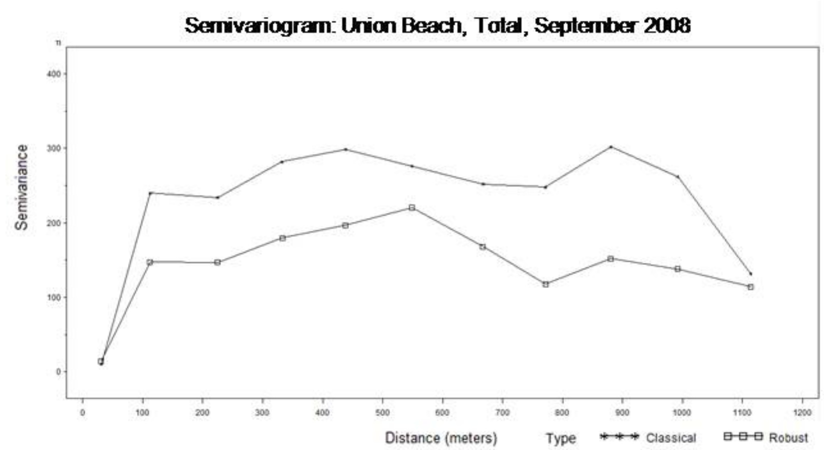

Semivariogram: Cliffwood Beach, Total, August 2008
Semivariogram: Keansburg, Total, August 2008
Semivariance
Type
Classical
Robust
Semivariogram: Keyport, Total, August 2008
Semivariogram: N. Middletown, Total, August 2008
Semivariance
Distance (meters)
Type
Classical
Robust
Distance (meters)


Semivariogram: Union Beach, Total, August 2008
Semivariance
Distance (meters)
Type
Classical
Robust

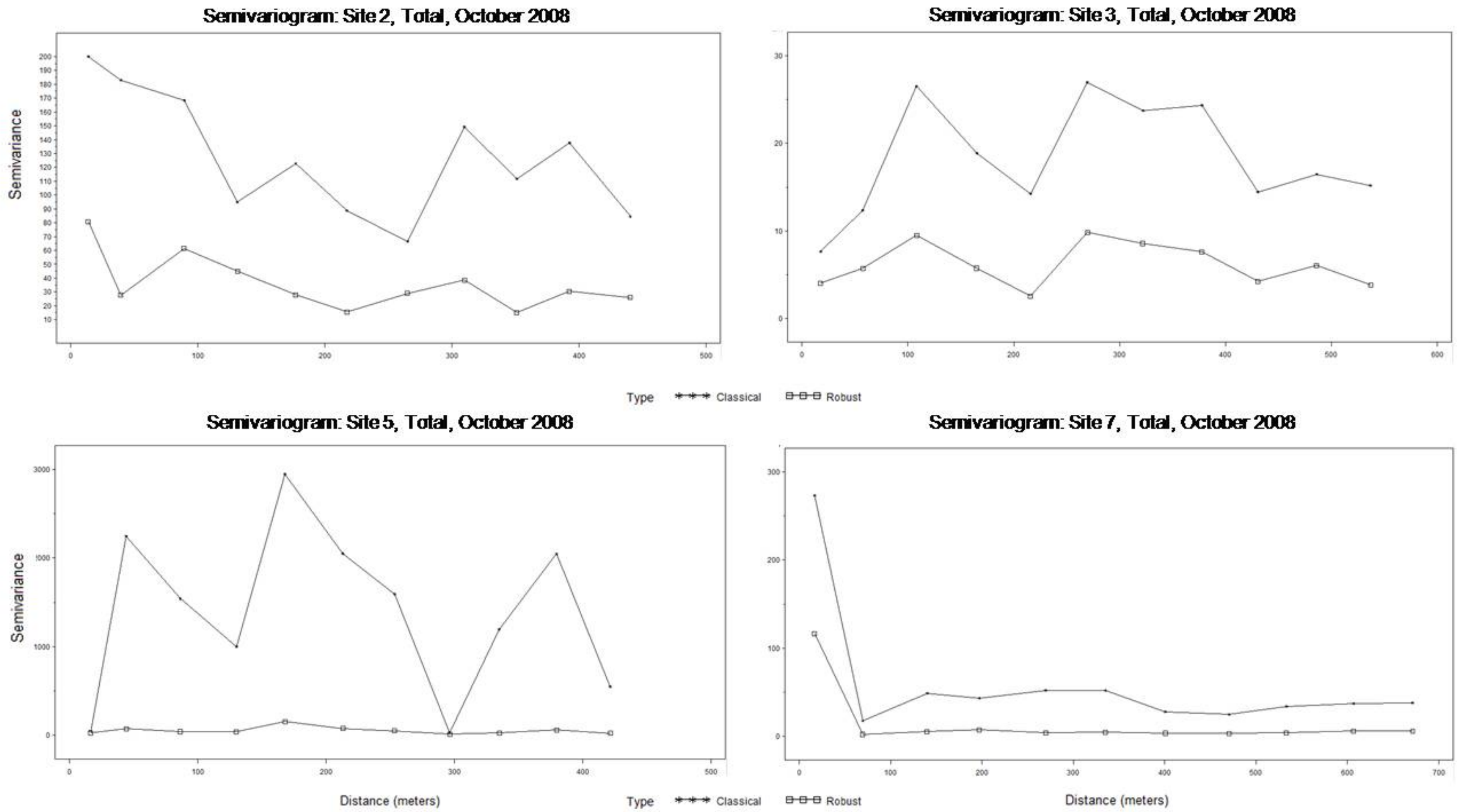
September 2008, Total

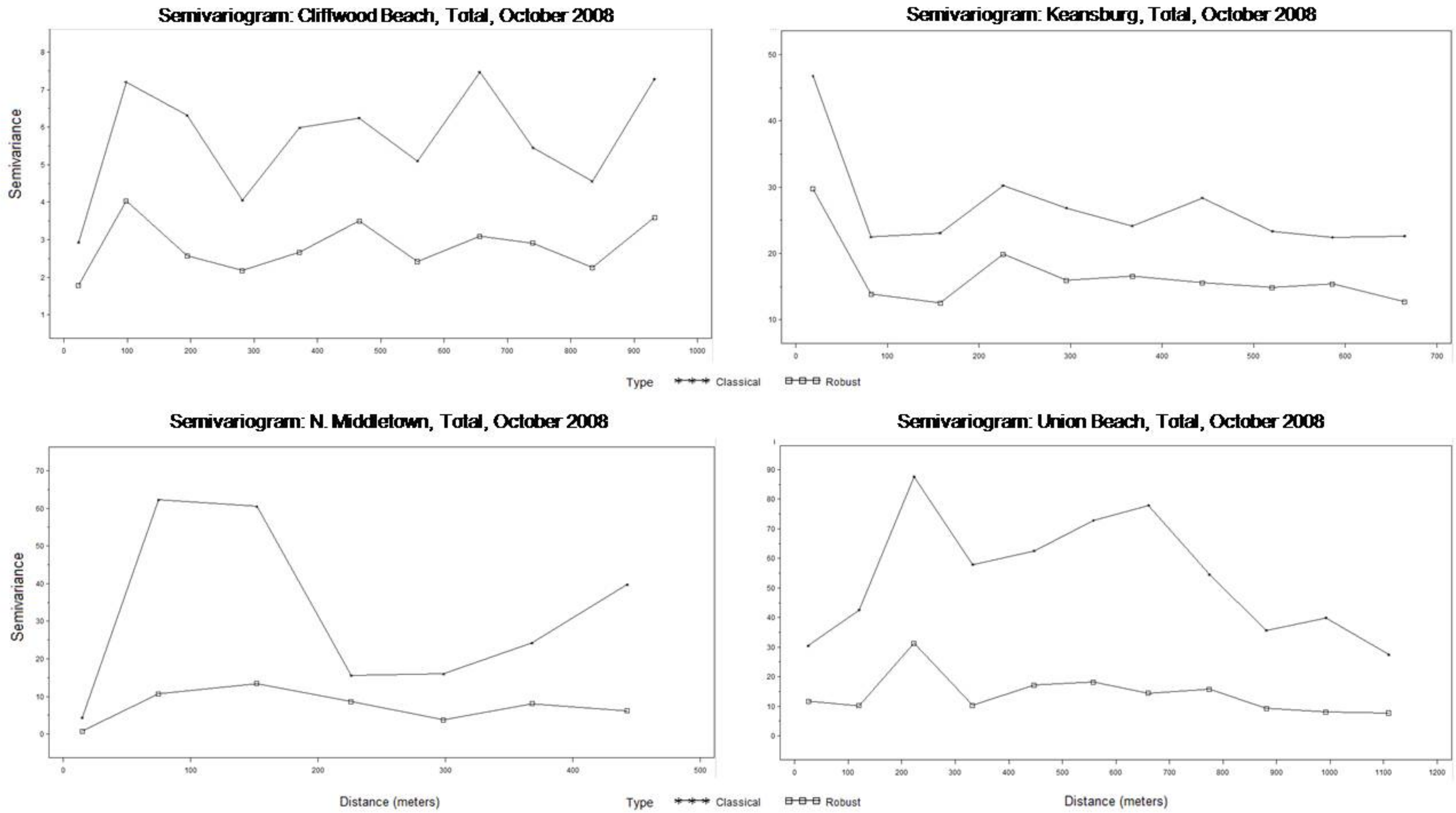



Semivariogram: Cliffwood Beach, Total, September 2008
Semivariogram: Keansburg, Total, September 2008
Semivariance
Type Classical Robust
Semivariogram: Keyport, Total, September 2008
Semivariogram: N. Middletown, Total, September 2008
Semivariance
Distance (meters)
Type Classical Robust
Distance (meters)


Semivariogram: Union Beach, Total, September 2008
Semivariance
Distance (meters)
Type
Classical
Robust

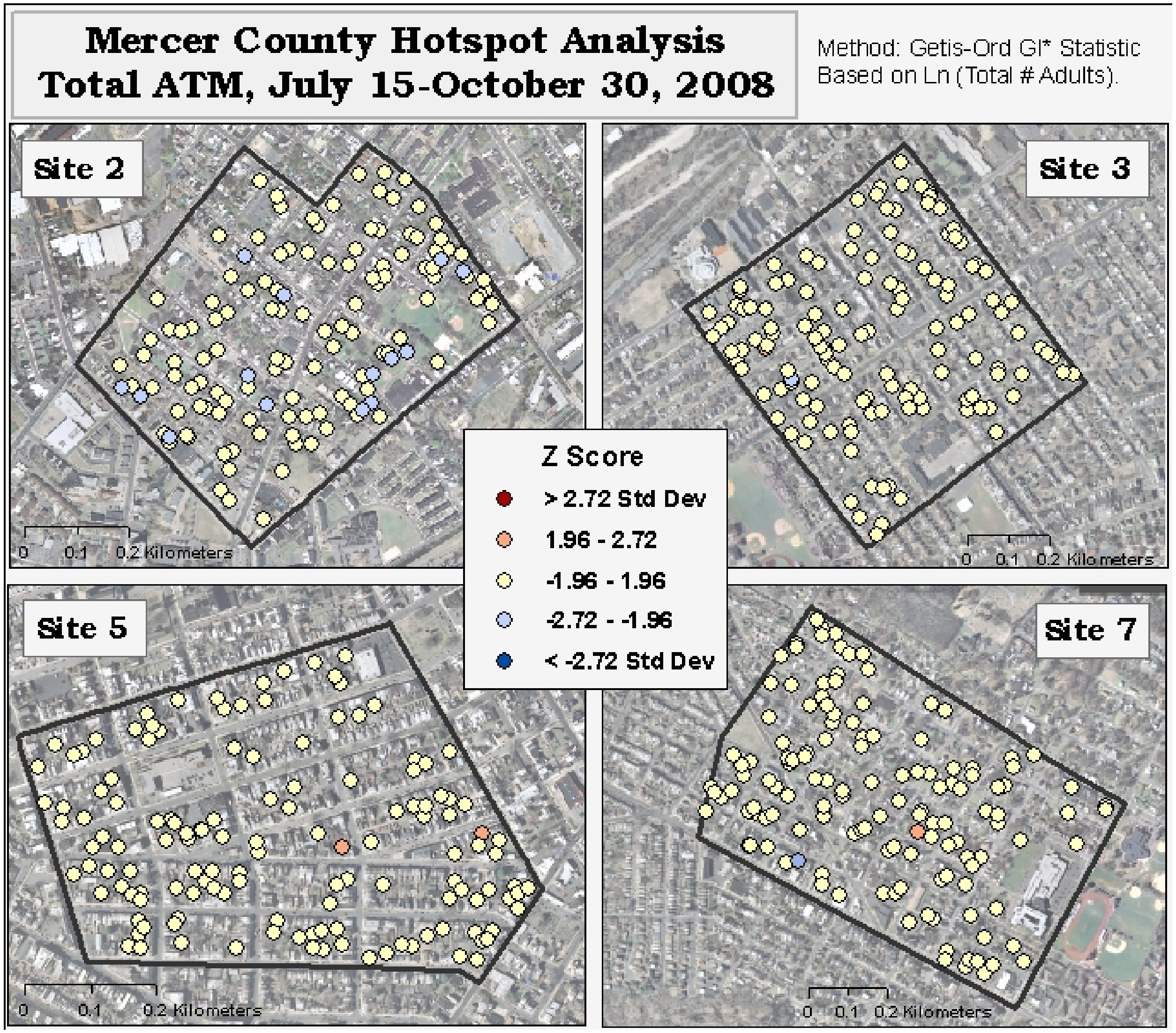
October 2008, Total

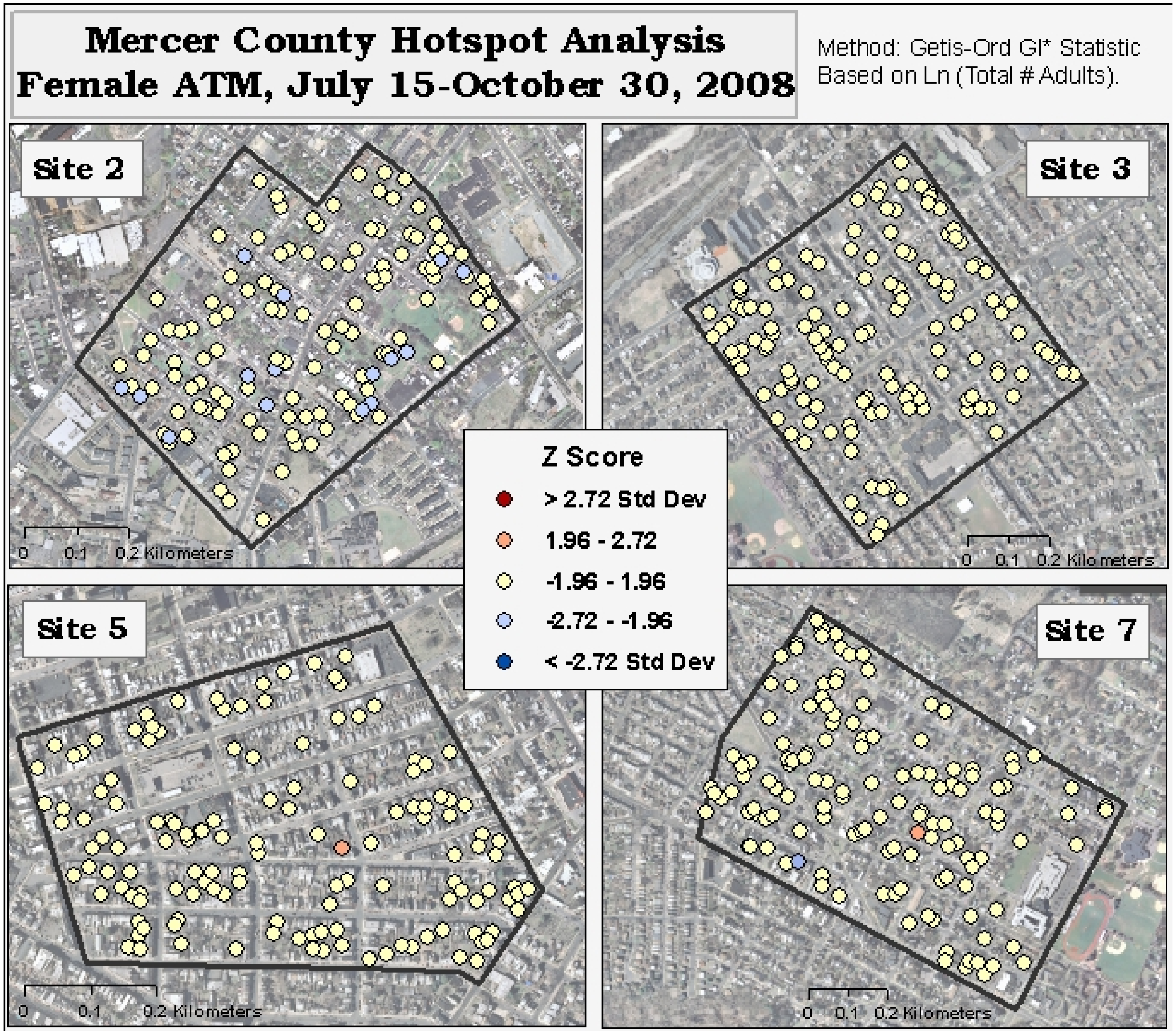


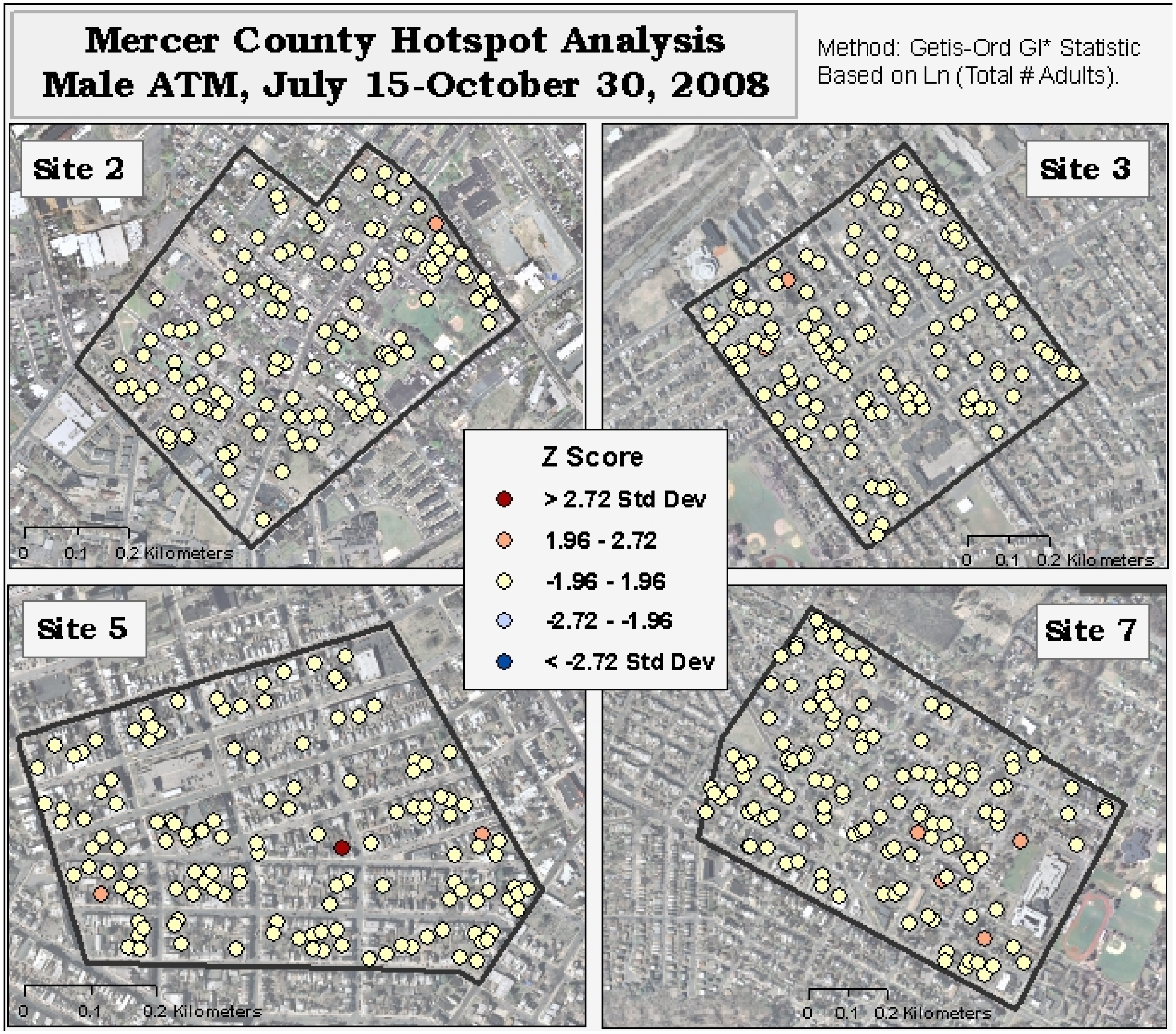


Semivariogram: Cliffwood Beach, Total, October 2008
Semivariogram: Keansburg, Total, October 2008
Semivariance
Type
Classical
Robust
Semivariogram: N. Middletown, Total, October 2008
Semivariogram: Union Beach, Total, October 2008
Semivariance
Distance (meters)
Distance (meters)
Type
Classical
Robust

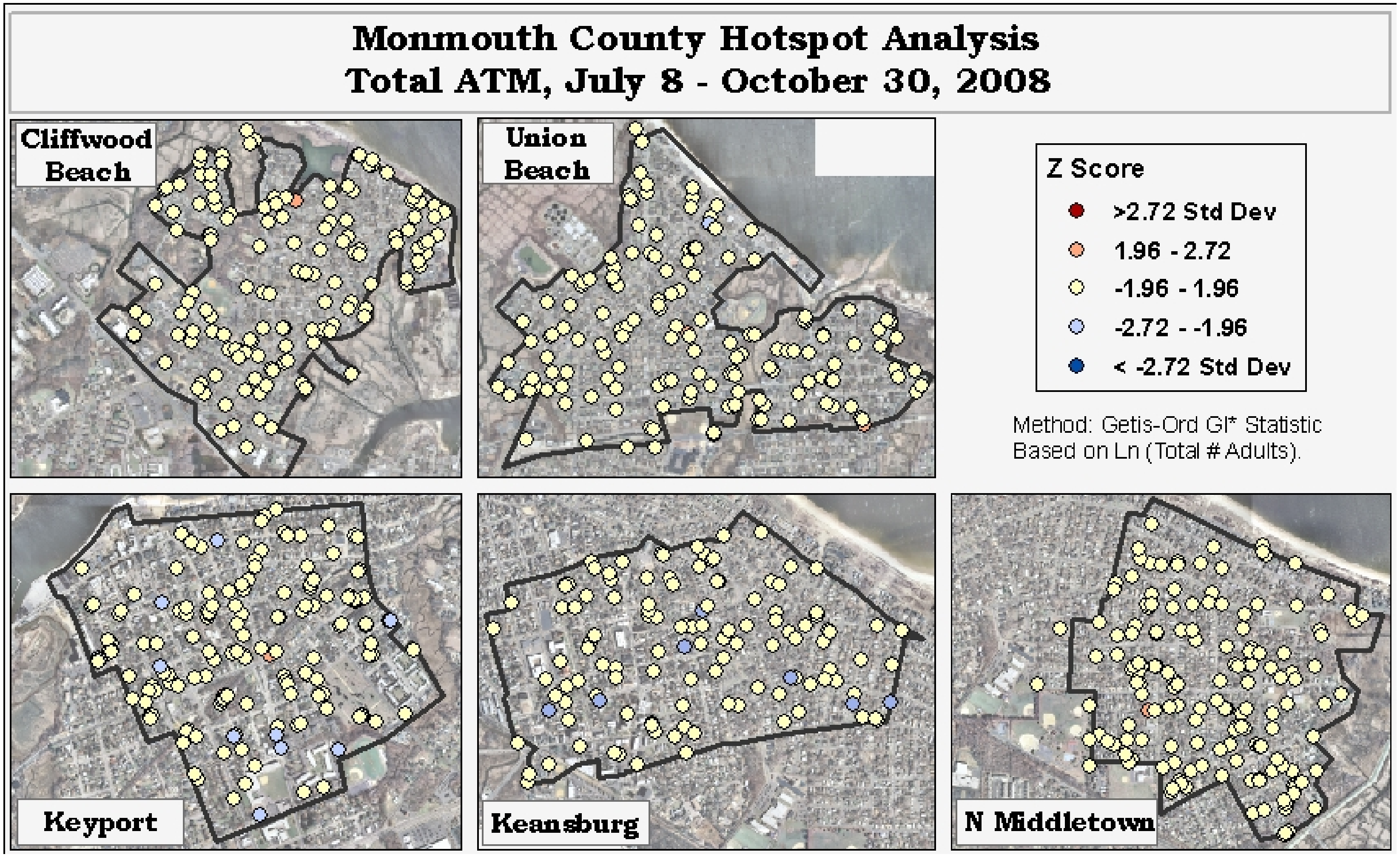
APPENDIX 2: HOT SPOT MAPS

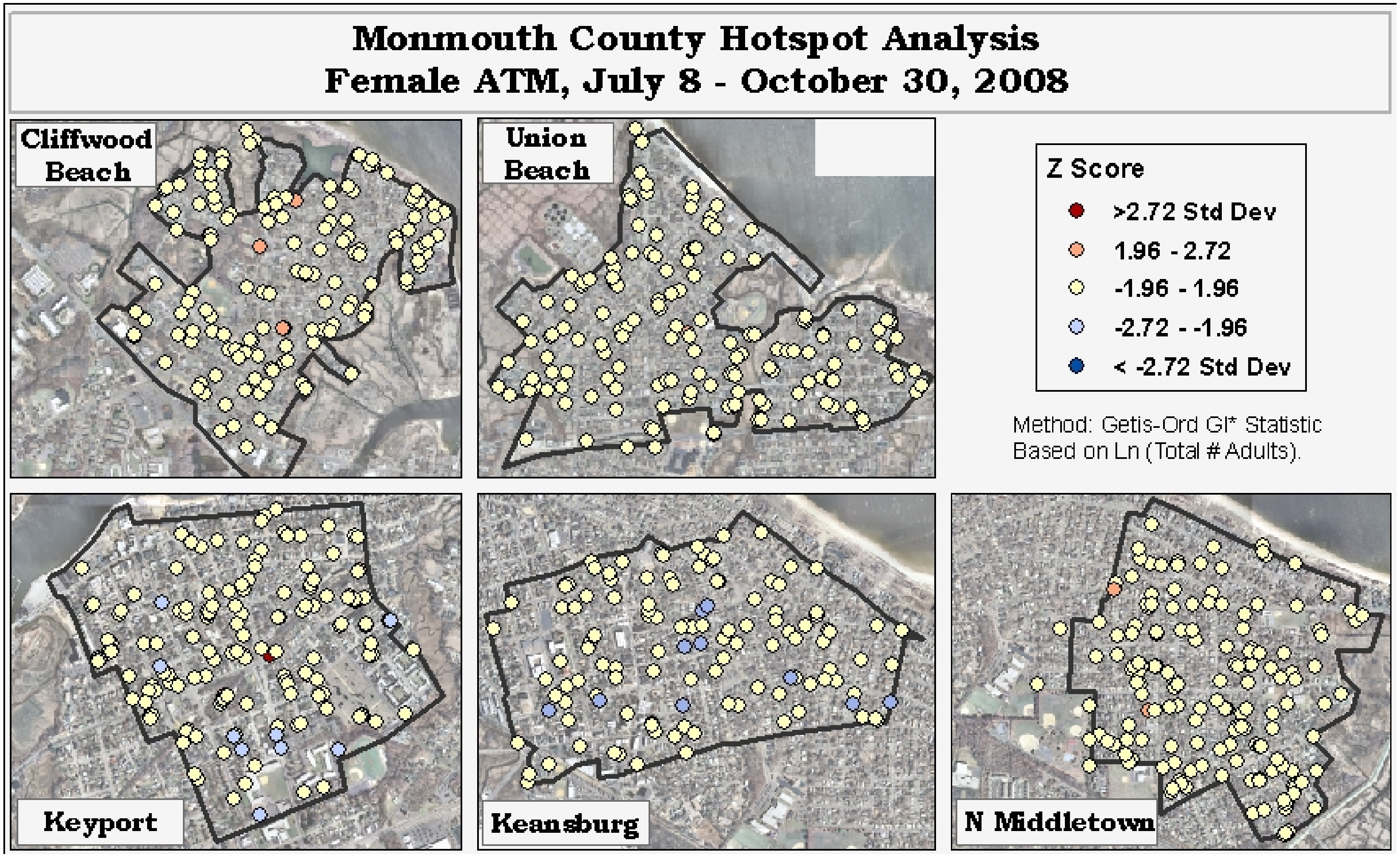


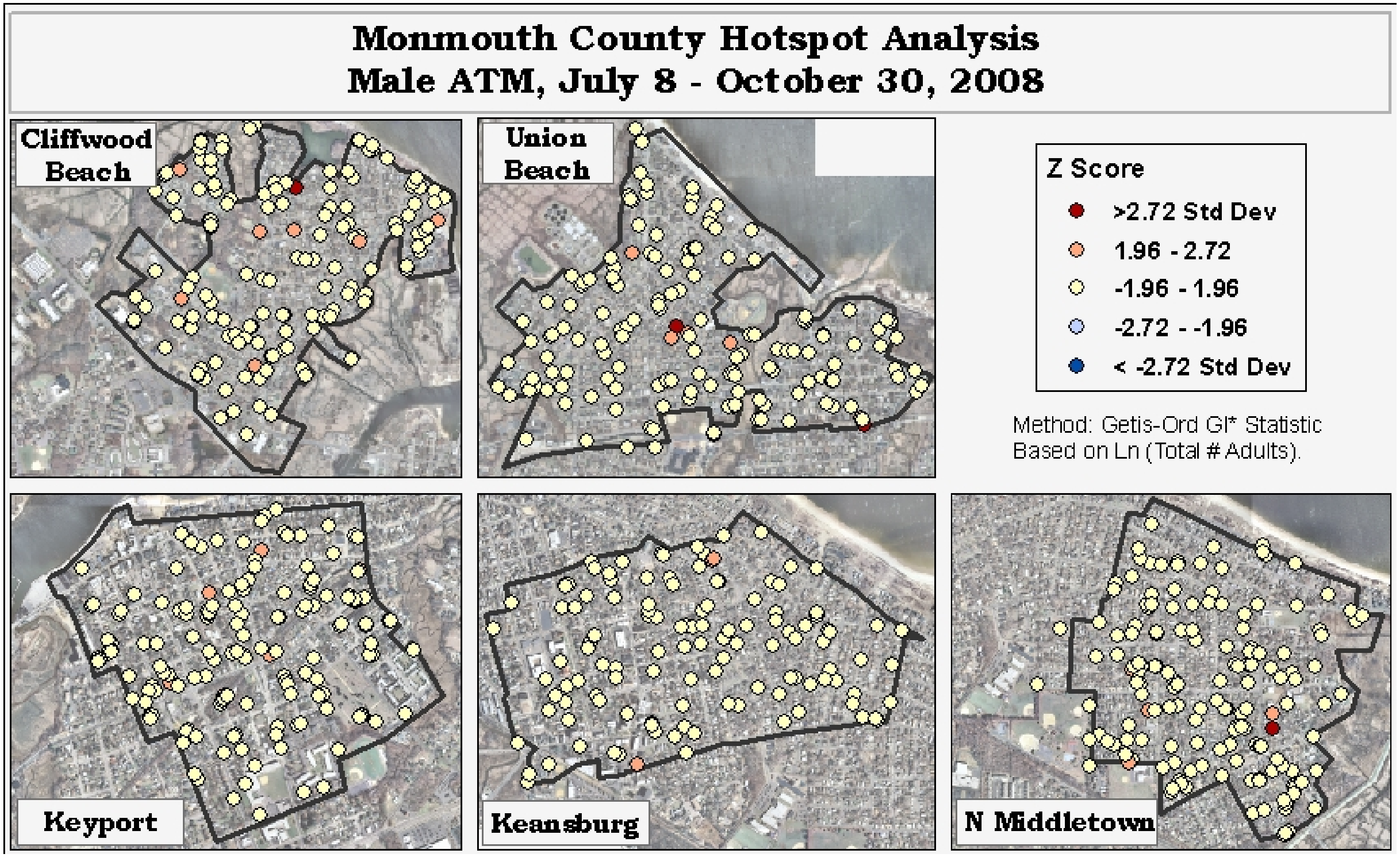

Mercer County Hotspot Analysis
Female ATM, July 15-October 30, 2008
Method: Getis-Ord Gi* Statistic
Based on Ln (Total # Adults).
Site 2
Site 3
Site 5
Site 7
Z Score
> 2.72 Std Dev
1.96 - 2.72
-1.96 - 1.96
-2.72 - -1.96
< -2.72 Std Dev
0 0.1 0.2 Kilometers

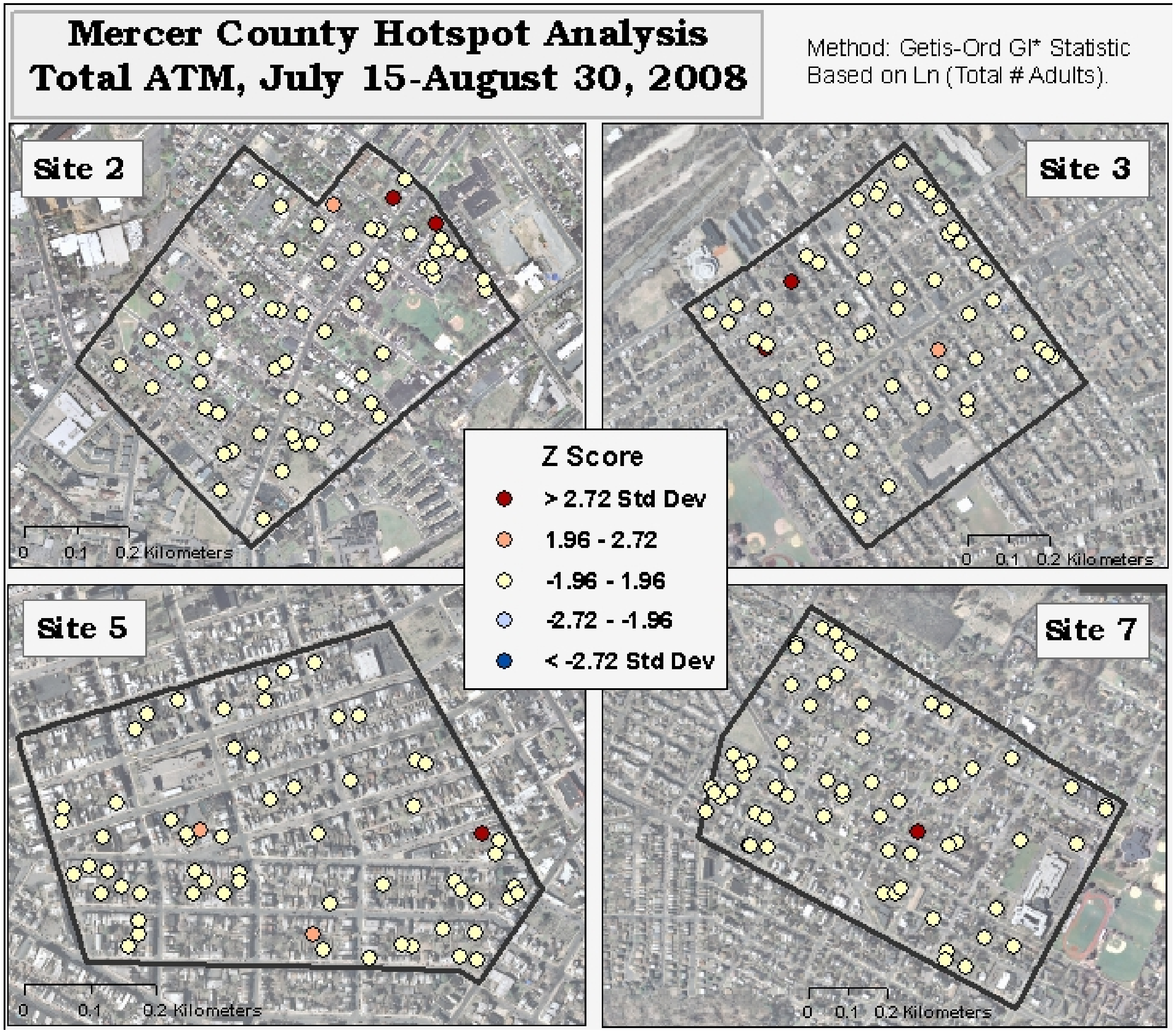


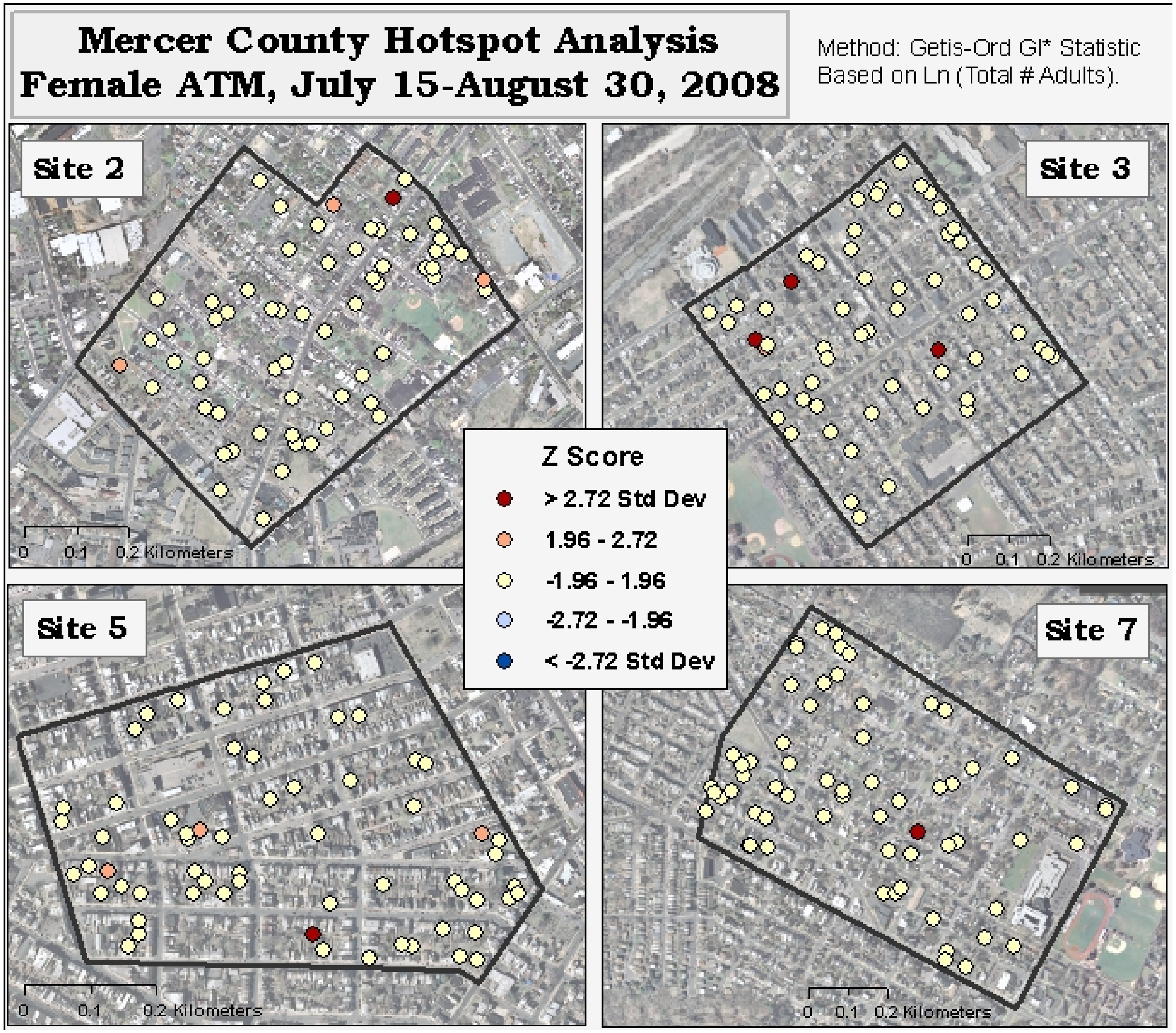
Mercer County Hotspot Analysis
Male ATM, July 15-October 30, 2008
Method: Getis-Ord GI* Statistic
Based on Ln (Total # Adults).
Site 2
Site 3
Site 5
Site 7
Z Score
> 2.72 Std Dev
1.96 - 2.72
-1.96 - 1.96
-2.72 - -1.96
< -2.72 Std Dev
0 0.1 0.2 Kilometers

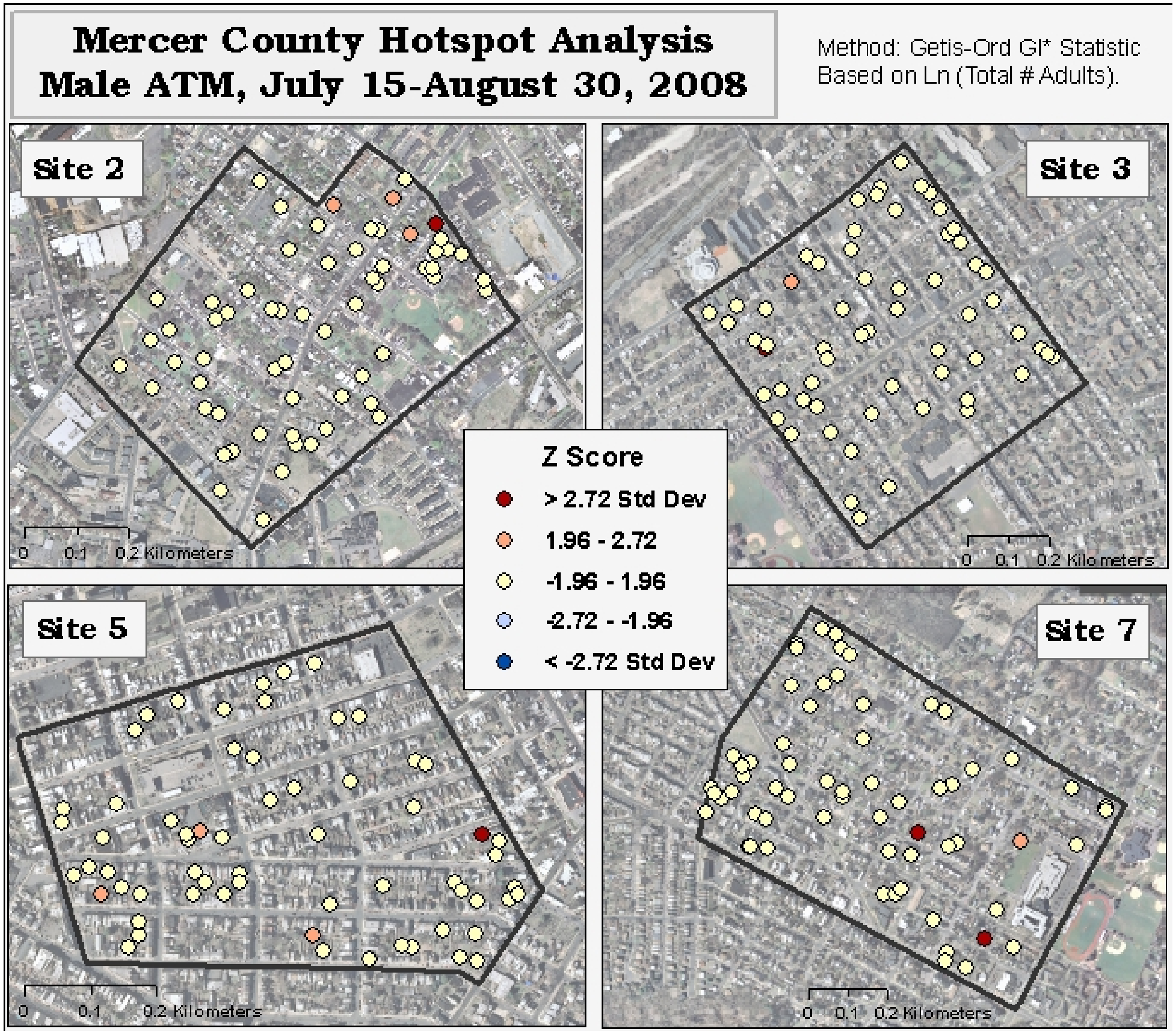

Monmouth County Hotspot Analysis
Total ATM, July 8 - October 30, 2008
Cliffwood Beach
Union Beach
Z Score
>2.72 Std Dev
1.96 - 2.72
-1.96 - 1.96
-2.72 - -1.96
< -2.72 Std Dev
Method: Getis-Ord GI* Statistic
Based on Ln (Total # Adults).
Keyport
Keansburg
N Middletown

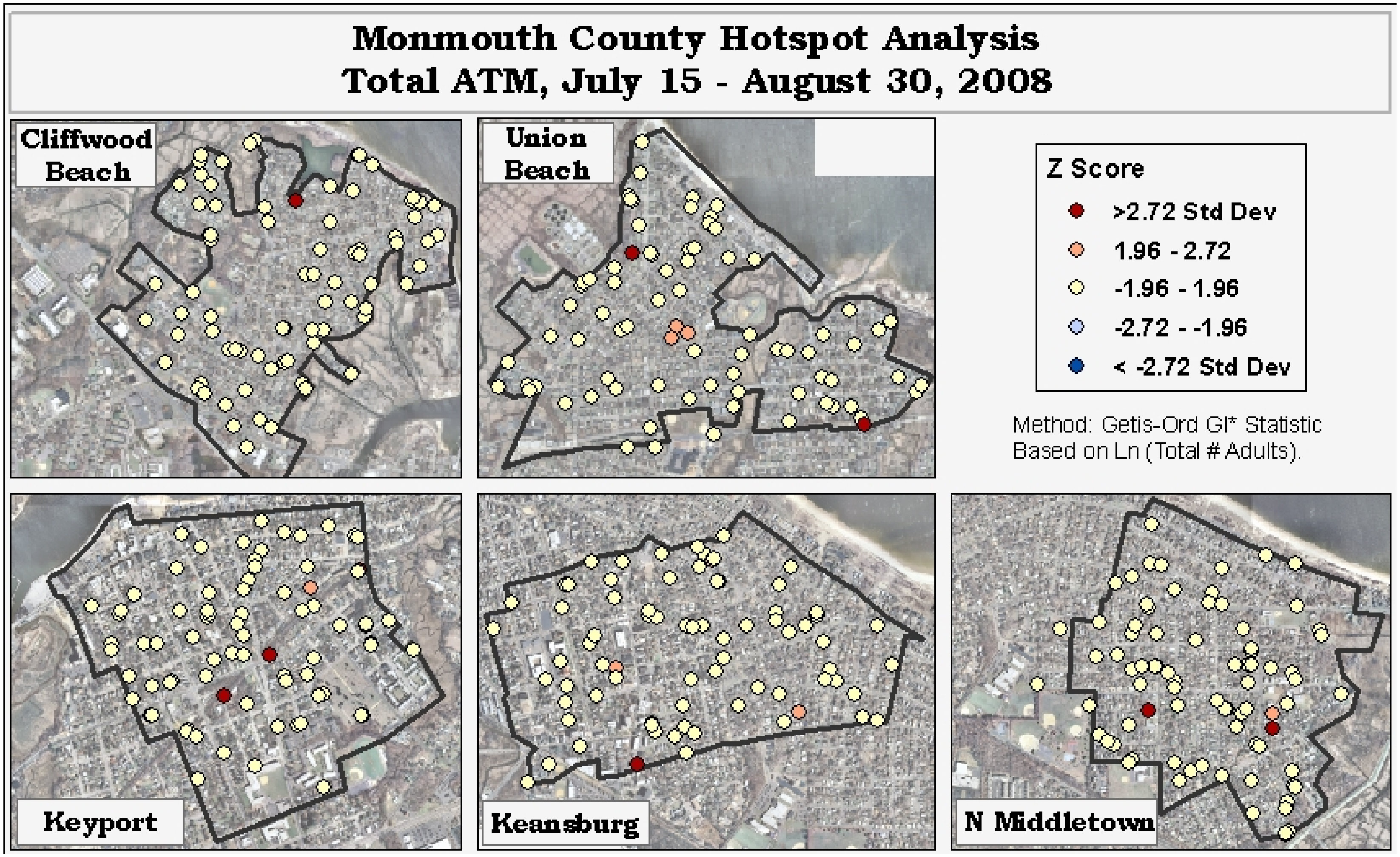

Monmouth County Hotspot Analysis
Female ATM, July 8 - October 30, 2008
Cliffwood Beach
Union Beach
Z Score
>2.72 Std Dev
1.96 - 2.72
-1.96 - 1.96
-2.72 - -1.96
< -2.72 Std Dev
Method: Getis-Ord Gi* Statistic
Based on Ln (Total # Adults).
Keyport
Keansburg
N Middletown

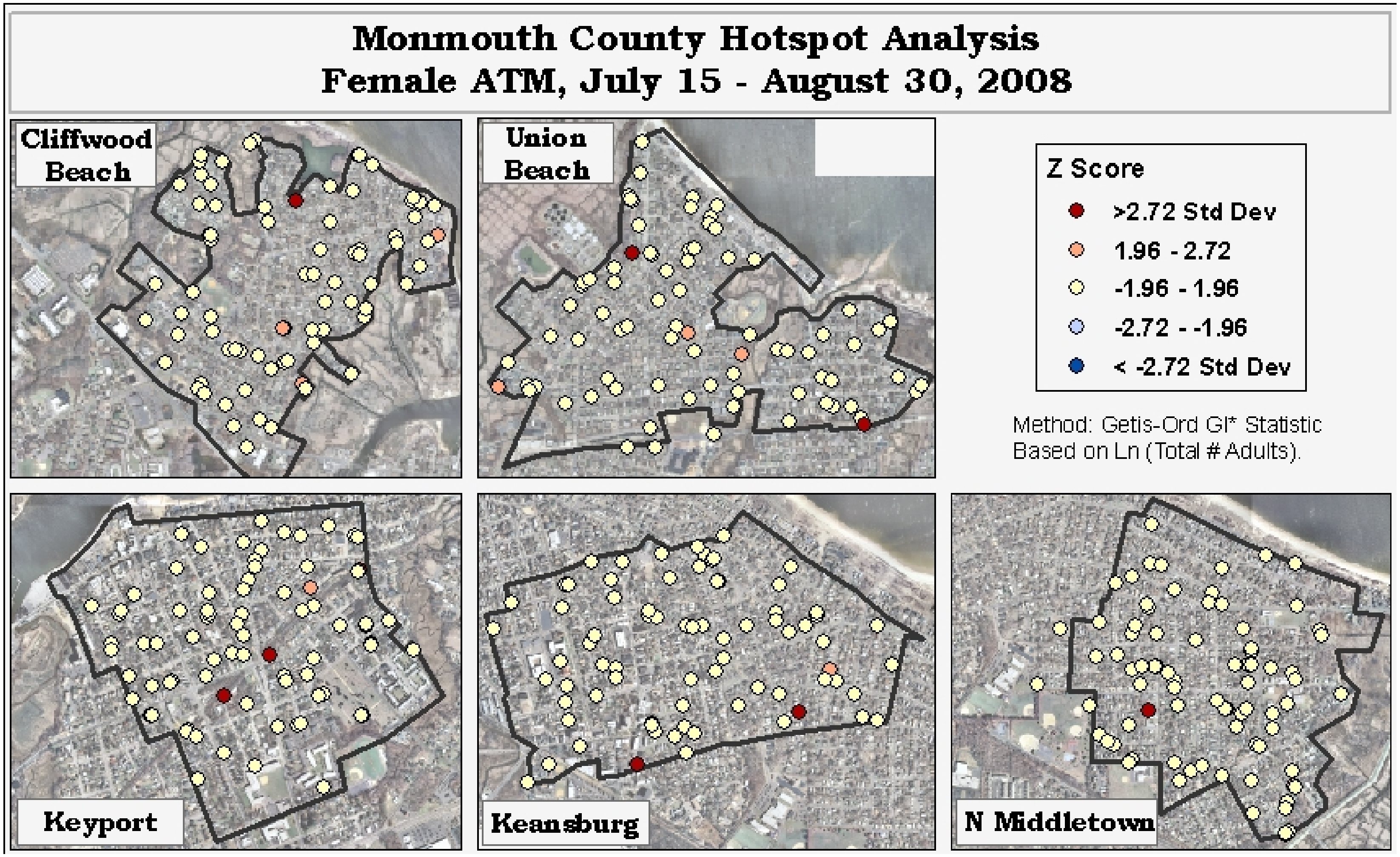

Monmouth County Hotspot Analysis
Male ATM, July 8 - October 30, 2008
Cliffwood Beach
Union Beach
Z Score
>2.72 Std Dev
1.96 - 2.72
-1.96 - 1.96
-2.72 - -1.96
< -2.72 Std Dev
Method: Getis-Ord Gi* Statistic
Based on Ln (Total # Adults).
Keyport
Keansburg
N Middletown

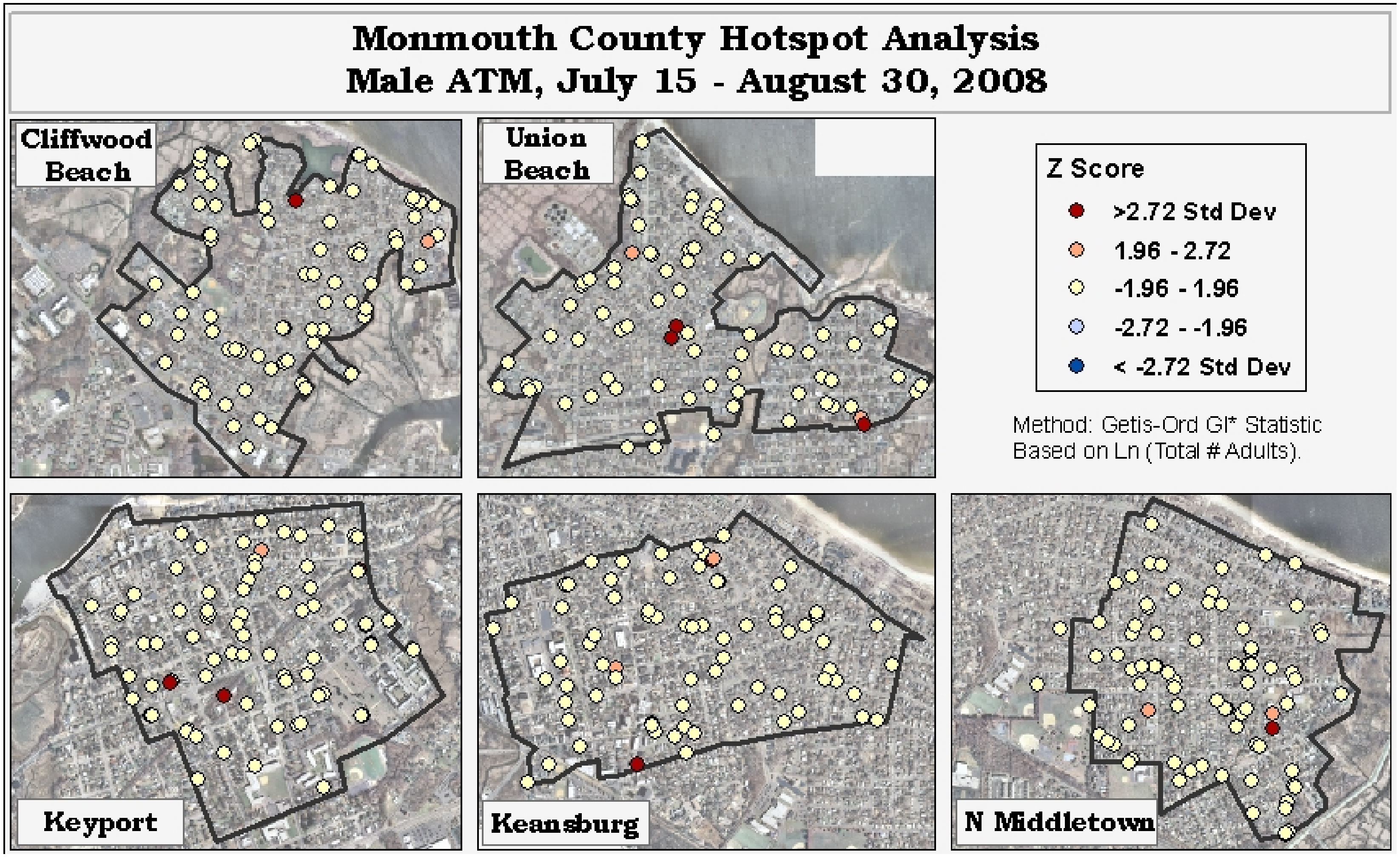

Mercer County Hotspot Analysis
Total ATM, July 15-August 30, 2008
Method: Getis-Ord GI* Statistic
Based on Ln (Total # Adults).
Site 2
Site 3
Site 5
Site 7
Z Score
> 2.72 Std Dev
1.96 - 2.72
-1.96 - 1.96
-2.72 - -1.96
< -2.72 Std Dev
0 0.1 0.2 Kilometers

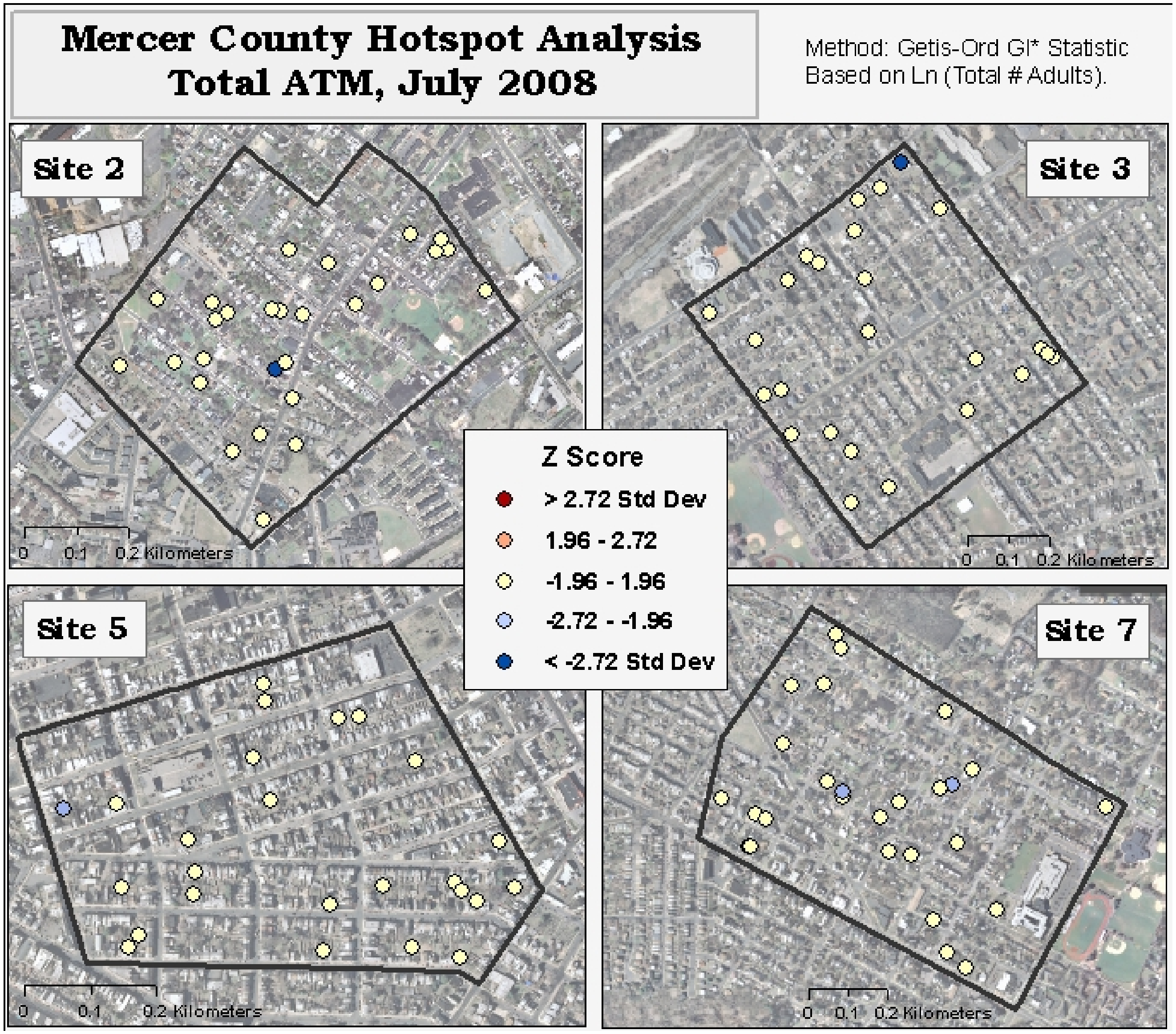

Mercer County Hotspot Analysis
Female ATM, July 15-August 30, 2008
Method: Getis-Ord Gi* Statistic
Based on Ln (Total # Adults).
Site 2
Site 3
Site 5
Site 7
Z Score
> 2.72 Std Dev
1.96 - 2.72
-1.96 - 1.96
-2.72 - -1.96
< -2.72 Std Dev
0 0.1 0.2 Kilometers

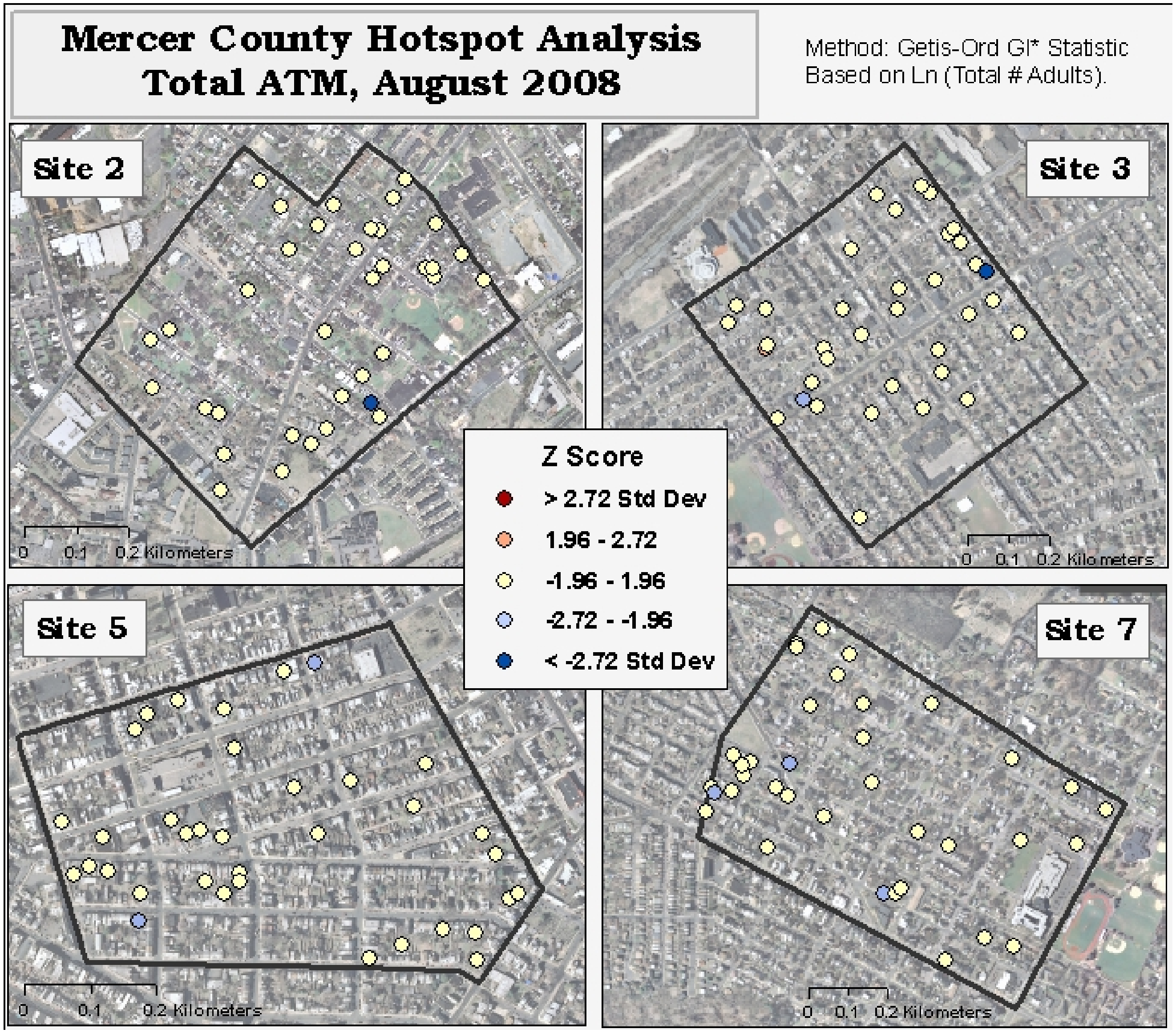

Mercer County Hotspot Analysis
Male ATM, July 15-August 30, 2008
Method: Getis-Ord Gi* Statistic
Based on Ln (Total # Adults).
Site 2
Site 3
Site 5
Site 7
Z Score
> 2.72 Std Dev
1.96 - 2.72
-1.96 - 1.96
-2.72 - -1.96
< -2.72 Std Dev
0 0.1 0.2 Kilometers

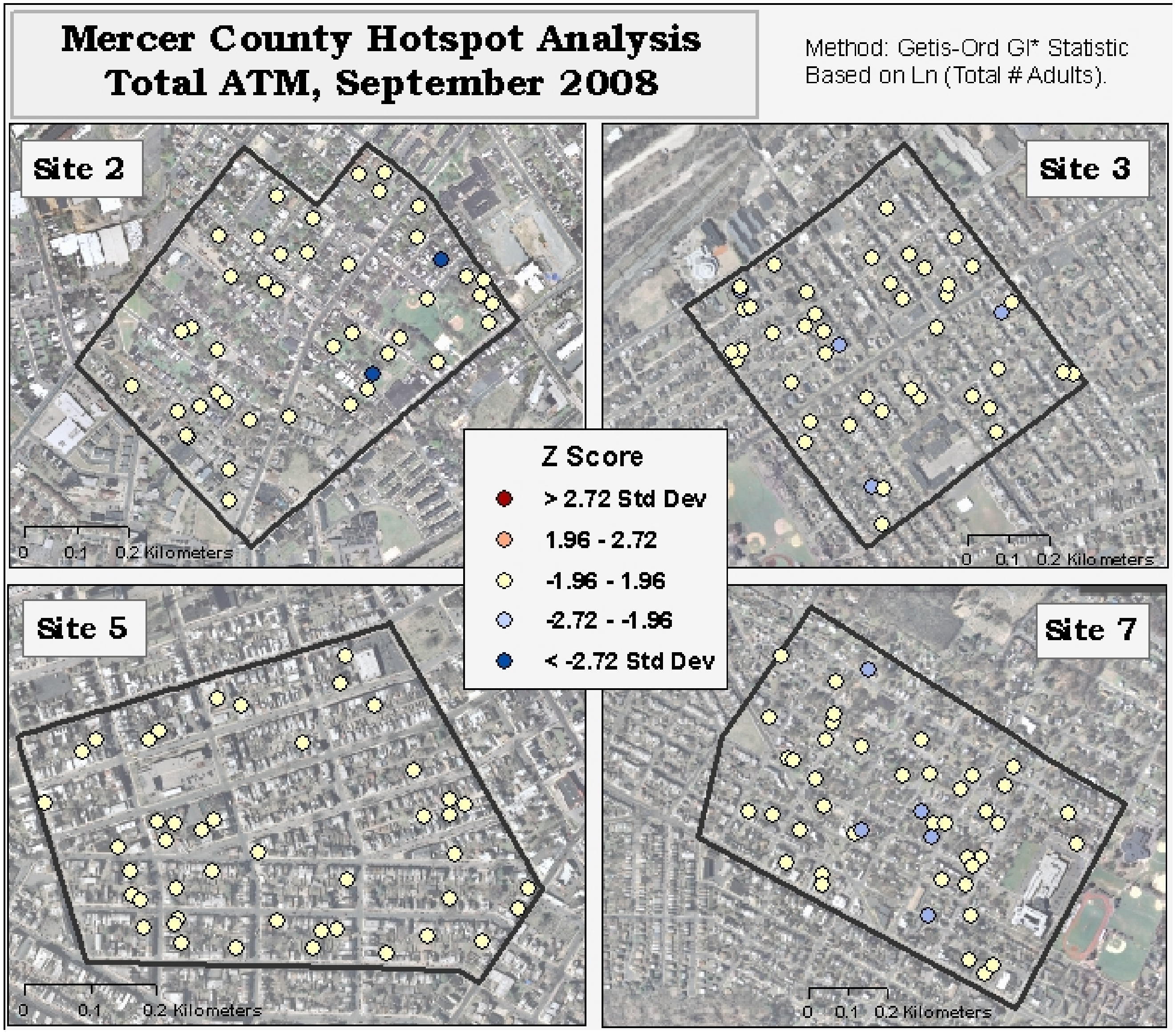

Monmouth County Hotspot Analysis
Total ATM, July 15 - August 30, 2008
Cliffwood Beach
Union Beach
Z Score
>2.72 Std Dev
1.96 - 2.72
-1.96 - 1.96
-2.72 - -1.96
< -2.72 Std Dev
Method: Getis-Ord Gi* Statistic
Based on Ln (Total # Adults).
Keyport
Keansburg
N Middletown


Monmouth County Hotspot Analysis
Female ATM, July 15 - August 30, 2008
Cliffwood Beach
Union Beach
Z Score
>2.72 Std Dev
1.96 - 2.72
-1.96 - 1.96
-2.72 - -1.96
< -2.72 Std Dev
Method: Getis-Ord Gi* Statistic
Based on Ln (Total # Adults).
Keyport
Keansburg
N Middletown

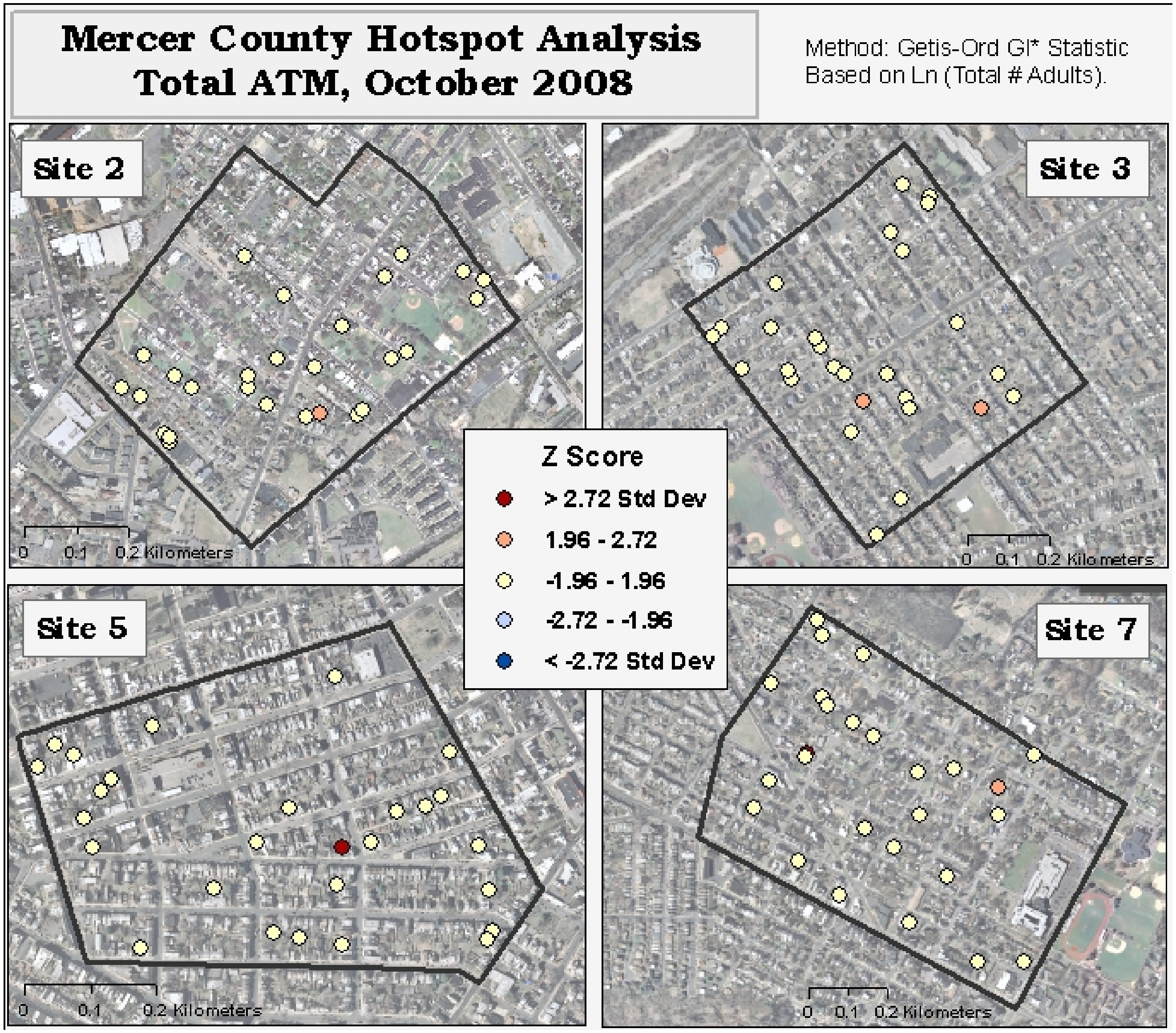

Monmouth County Hotspot Analysis
Male ATM, July 15 - August 30, 2008
Cliffwood Beach
Union Beach
Z Score
>2.72 Std Dev
1.96 - 2.72
-1.96 - 1.96
-2.72 - -1.96
< -2.72 Std Dev
Method: Getis-Ord Gi* Statistic
Based on Ln (Total # Adults).
Keyport
Keansburg
N Middletown


Mercer County Hotspot Analysis
Total ATM, July 2008
Method: Getis-Ord Gi* Statistic
Based on Ln (Total # Adults).
Site 2
Site 3
Site 5
Site 7
Z Score
> 2.72 Std Dev
1.96 - 2.72
-1.96 - 1.96
-2.72 - -1.96
< -2.72 Std Dev
0 0.1 0.2 Kilometers

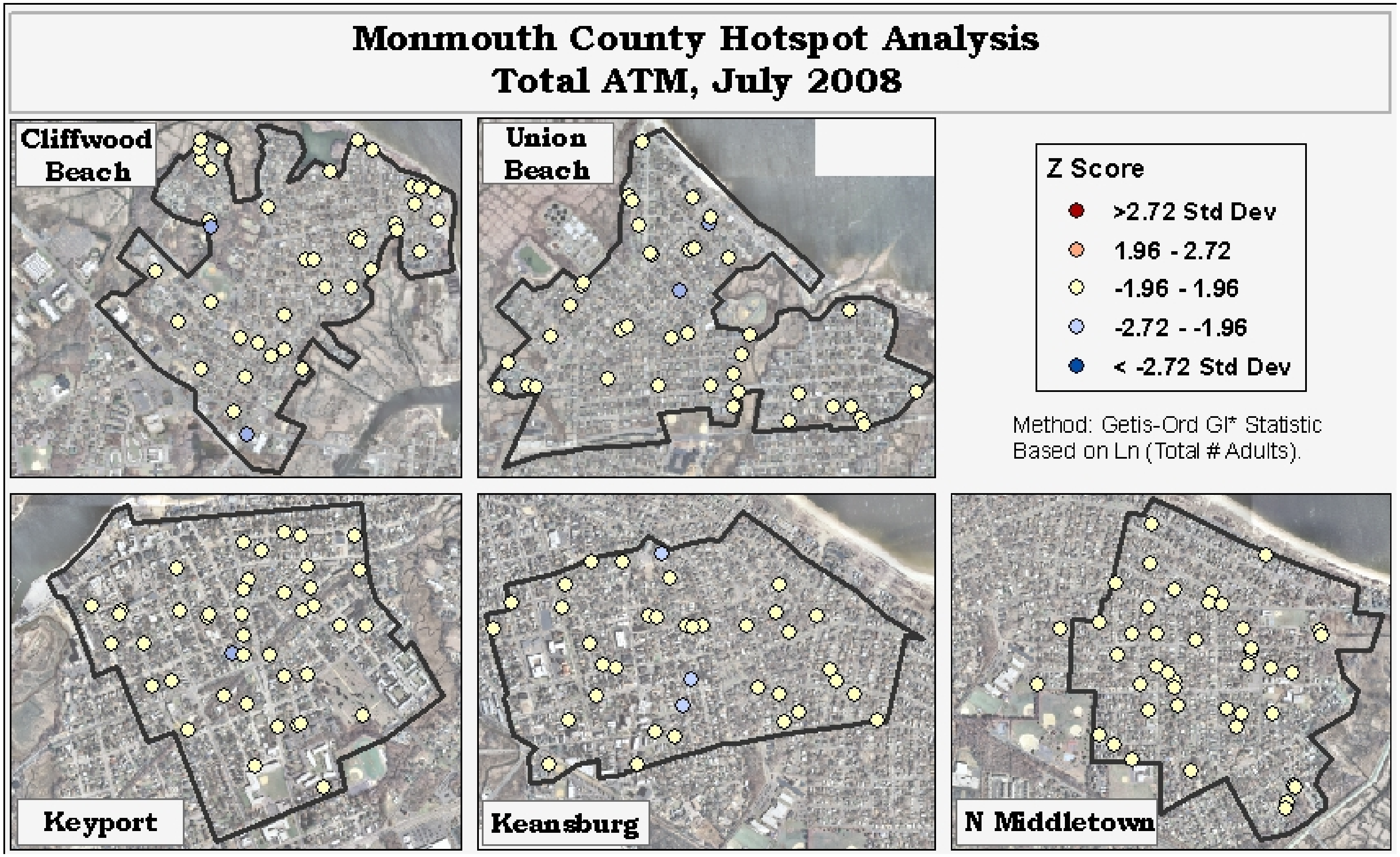

Mercer County Hotspot Analysis
Total ATM, August 2008
Method: Getis-Ord Gi* Statistic
Based on Ln (Total # Adults).
Site 2
Site 3
Site 5
Site 7
Z Score
> 2.72 Std Dev
1.96 - 2.72
-1.96 - 1.96
-2.72 - -1.96
< -2.72 Std Dev
0 0.1 0.2 Kilometers



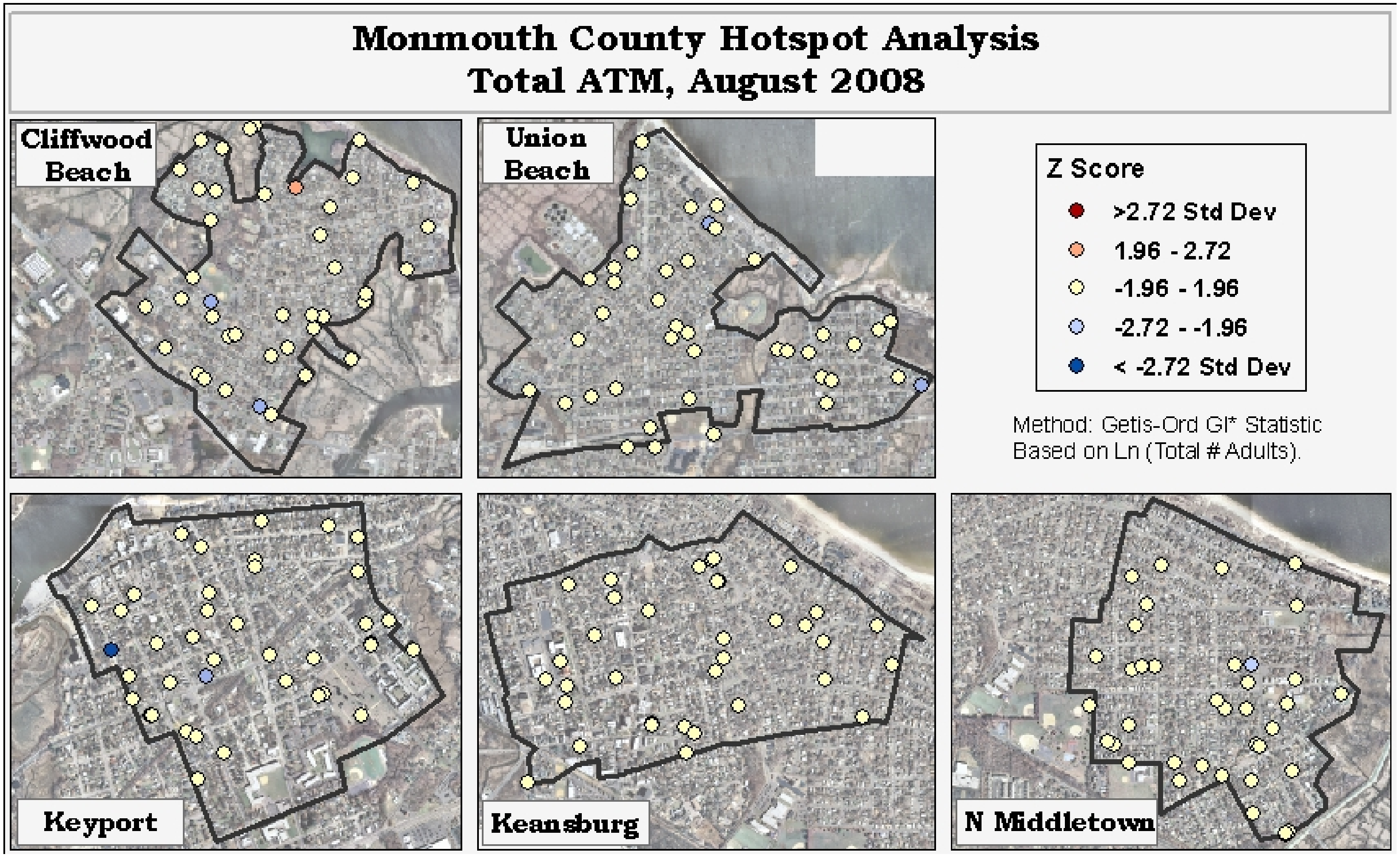
Mercer County Hotspot Analysis
Total ATM, September 2008
Method: Getis-Ord GI* Statistic
Based on Ln (Total # Adults).
Site 2
Site 3
Site 5
Site 7
Z Score
> 2.72 Std Dev
1.96 - 2.72
-1.96 - 1.96
-2.72 - -1.96
< -2.72 Std Dev
0 0.1 0.2 Kilometers

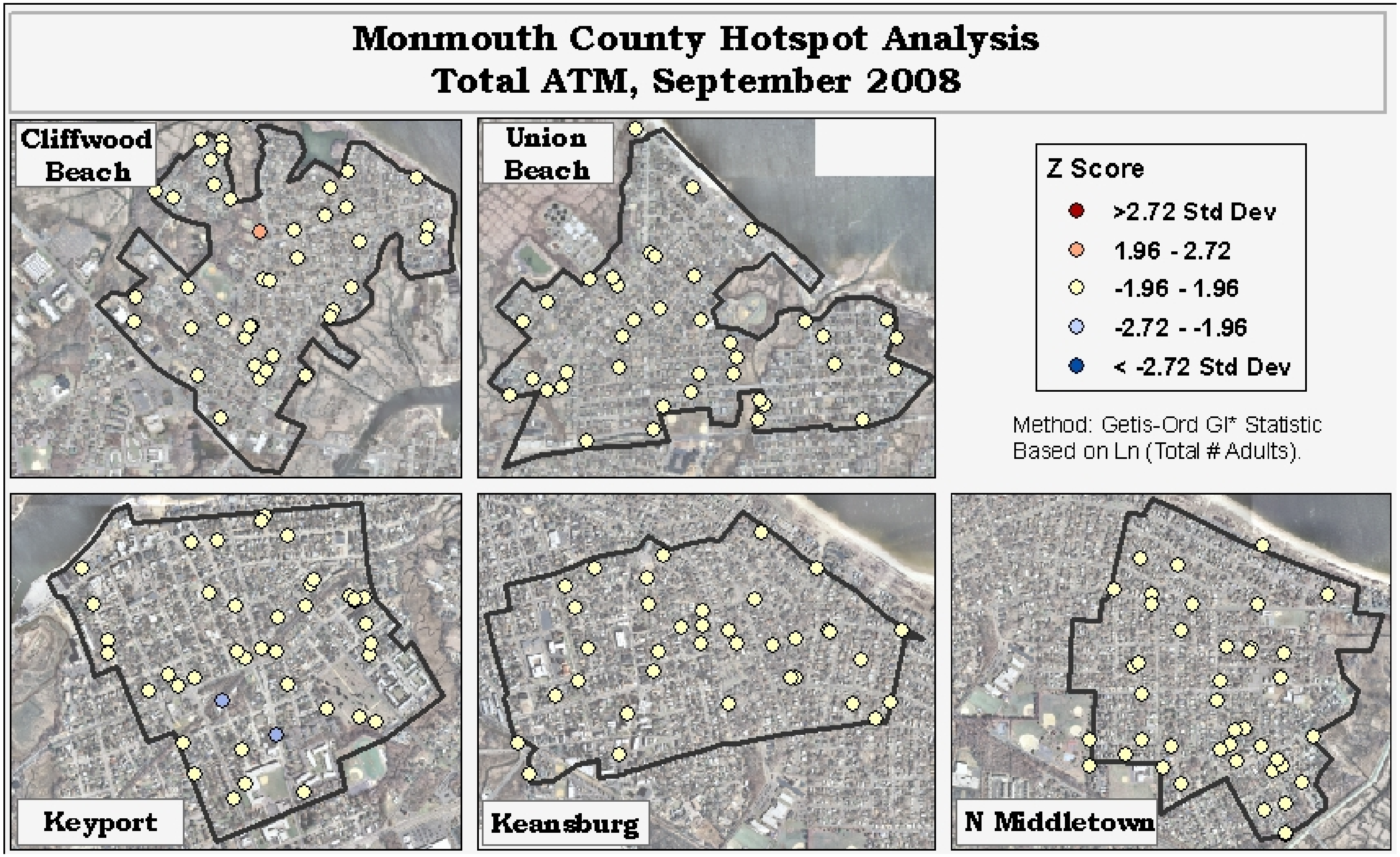


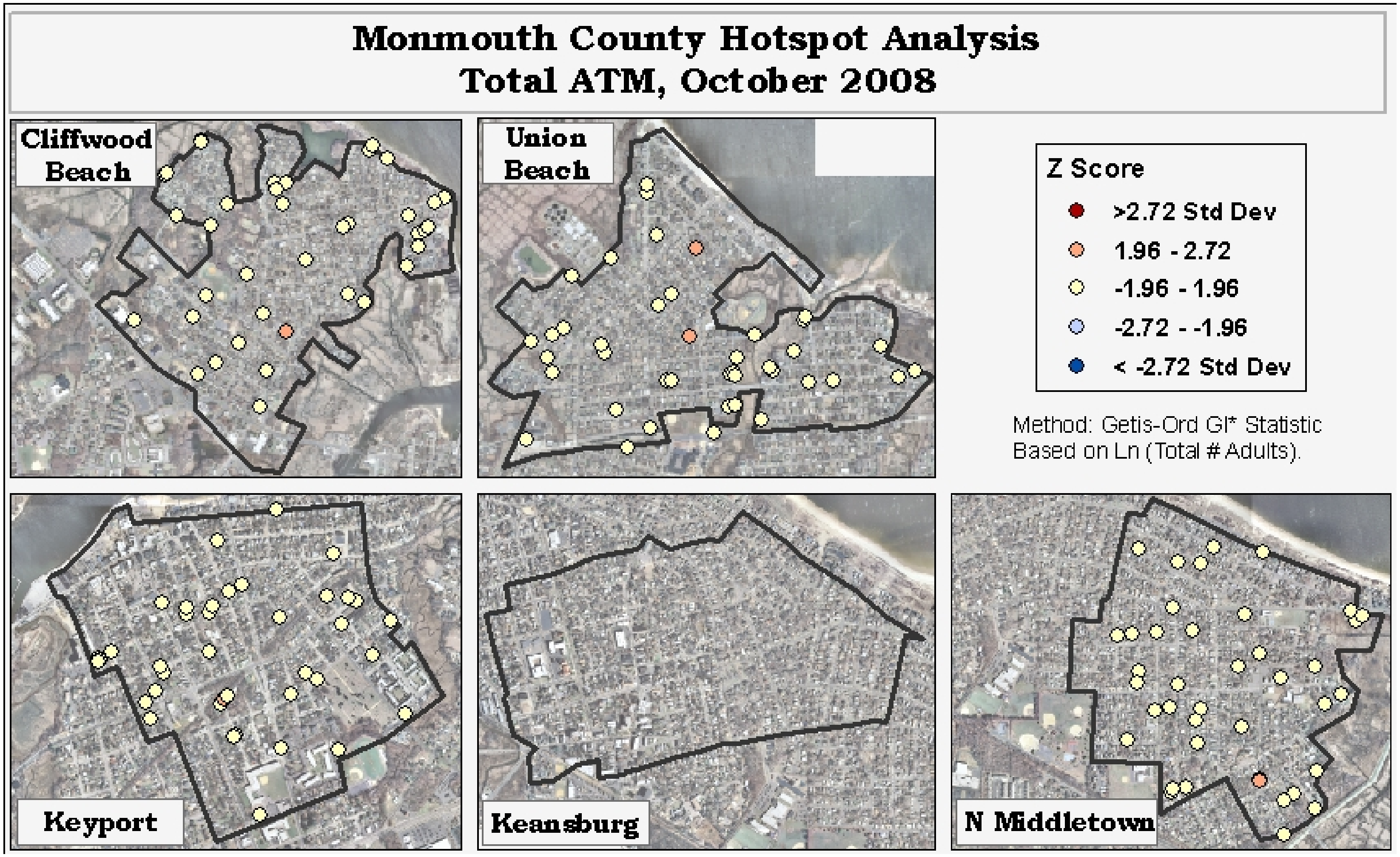
Mercer County Hotspot Analysis
Total ATM, October 2008
Method: Getis-Ord Gi* Statistic
Based on Ln (Total # Adults).
Site 2
Site 3
Site 5
Site 7
Z Score
> 2.72 Std Dev
1.96 - 2.72
-1.96 - 1.96
-2.72 - -1.96
< -2.72 Std Dev
0 0.1 0.2 Kilometers



Monmouth County Hotspot Analysis
Total ATM, July 2008
Cliffwood Beach
Union Beach
Keyport
Keansburg
N Middletown
Z Score
>2.72 Std Dev
1.96 - 2.72
-1.96 - 1.96
-2.72 - -1.96
< -2.72 Std Dev
Method: Getis-Ord Gi* Statistic
Based on Ln (Total # Adults).


Monmouth County Hotspot Analysis
Total ATM, August 2008
Cliffwood Beach
Union Beach
Z Score
>2.72 Std Dev
1.96 - 2.72
-1.96 - 1.96
-2.72 - -1.96
< -2.72 Std Dev
Method: Getis-Ord Gi* Statistic
Based on Ln (Total # Adults).
Keyport
Keansburg
N Middletown


Monmouth County Hotspot Analysis
Total ATM, September 2008
Cliffwood Beach
Union Beach
Z Score
>2.72 Std Dev
1.96 - 2.72
-1.96 - 1.96
-2.72 - -1.96
< -2.72 Std Dev
Method: Getis-Ord Gi* Statistic
Based on Ln (Total # Adults).
Keyport
Keansburg
N Middletown



Monmouth County Hotspot Analysis
Total ATM, October 2008
Cliffwood Beach
Union Beach
Z Score
>2.72 Std Dev
1.96 - 2.72
-1.96 - 1.96
-2.72 - -1.96
< -2.72 Std Dev
Method: Getis-Ord Gi* Statistic
Based on Ln (Total # Adults).
Keyport
Keansburg
N Middletown