PRACTICE OF EPIDEMIOLOGY

Identifying West Nile Virus Risk Areas: The Dynamic Continuous-Area Space-Time System

Constandinos N. Theophilides[1], Sean C. Ahearn[1], Sue Grady[1], and Mario Merlino[2]

[1] Center for Advanced Research of Spatial Information, Hunter College, City University of New York, New York, NY.
[2] Office of Policy and Planning, New York City Department of Health and Mental Hygiene, New York, NY.

The Dynamic Continuous-Area Space-Time (DYCAST) system was developed to identify and prospectively monitor high-risk areas for West Nile virus in New York, New York (New York City). The system is based on a geographic model that uses a localized Knox test to capture the nonrandom space-time interaction of dead birds, as an indicator of an intense West Nile virus amplification cycle, within a 1.5-mile (2.41-km) buffer area and 21-day moving window. The Knox analysis is implemented as an interpolation function to create a surface of probabilities over a grid of 1,400 cells overlaying New York City. The model's parameters were calibrated using year 2000 data and information on the vector-host transmission cycle. The DYCAST system was implemented in a geographic information system and used operationally in year 2001. It successfully identified areas of high risk for human West Nile virus infection in areas where five of seven human cases resided, at least 13 days prior to the onset of illness, and proved that it can be used as an effective tool for targeting remediation and control efforts.

arboviruses; geographic information system; space-time clustering; West Nile virus

Abbreviations: DYCAST, Dynamic Continuous-Area Space-Time; MAUP, modifiable areal unit problem.

In 1999, West Nile virus made its first appearance in the Western Hemisphere and New York, New York (New York City). By 2001, 68 New York City area residents were diagnosed with laboratory-confirmed West Nile virus infection, and there were nine deaths (1, 2). Remediation and control measures were implemented to reduce the numbers of mosquitoes, identified as the vectors of West Nile virus (3).

West Nile virus is a mosquito-borne flavivirus belonging to the Japanese encephalitis virus serocomplex. In New York City, the *Culex pipiens* species is the primary vector of West Nile virus as evidenced by its high competency to transmit West Nile virus, its abundance, and its strong ornithophilic habits (4–6). The primary hosts are native bird species that lack West Nile virus immunity, particularly the American crow and blue jay. Pigeons, although abundant in New York City, have very low infection rates. Humans are dead-end hosts, and their infections are most likely incidental and the

result of a spillover effect. Spillover effects occur when the transmission cycle between mosquitoes and birds intensifies (a process termed the amplification cycle), or when the avian pool is reduced and species of mosquitoes that are opportunistic feeders act as a bridge vector for transmitting West Nile virus to humans (4, 7).

In New York City, during the 1999 season, remediation and control efforts were based on positive West Nile virus human infections, and in the 2000 season, efforts were based on laboratory confirmation of West Nile virus in mosquitoes and birds, which had a delay time of up to 2 weeks. Recently, the New York City Department of Health implemented an in-house laboratory for faster positive mosquito identification. Although these approaches provide a definite confirmation of West Nile virus, they lack timeliness, because positive results in mosquitoes may not appear until West Nile virus activity has substantially intensified. In addition,

Reprint requests to Constandinos Theophilides, Center for Advanced Research of Spatial Information, Hunter College, City University of New York, 695 Park Avenue, New York, NY 10021 (e-mail: ctheo@geo.hunter.cuny.edu).

they identify West Nile virus activity only at discrete point locations, where mosquito traps were positioned and dead birds were found. An important issue for the New York City Department of Health was how to use data points to identify the areal extent of West Nile virus activity in a timely fashion. To address this issue, the New York City Department of Health in cooperation with the Center for Advanced Research of Spatial Information (CARSI) Laboratory of Hunter College, City University of New York, embarked on an effort in January 2001 to develop an area-based system to identify areas of West Nile virus activity that could lead to human infection, for targeting remediation and control efforts. The system needed to be both prospective and dynamic, while providing the New York City Department of Health with the geographic extent of West Nile virus activity.

A widely used methodology for monitoring West Nile virus activity relies on dead crow densities (8, 9) or the number of dead crows per unit area. This approach has a number of limitations: 1) It does not ascribe statistical meaning to the results, leading to the selection of an arbitrary cutoff (critical) density for judging high-risk areas for West Nile virus; 2) density is highly susceptible to reporting bias; 3) density calculations for areas that vary in size, shape, and scale rely on a false assumption of uniformity of crow densities throughout the region, resulting in the aggregation problem of the modifiable areal unit problem (MAUP) (10); 4) densities calculated using kernel functions are subject to edge effects (11), although corrections for edges are available when the process is assumed to be stationary and isotropic (12); and 5) density measures ignore the pathology and ecology of the West Nile virus transmission cycle.

Other methodologies used to model infectious diseases generally factor only the temporal and not the spatial component of epidemics (13). Those methods that do account for space often use single unpartitioned areas or nonoverlapping spatial units specifically designed for unrelated purposes (i.e., census tracts). The latter can lead to an artificial split of clustered data. In addition, these studies are often retrospective and require the use of controls, the selection of which may be biased when knowledge of the ecology of an infectious disease is incomplete (14). There is limited research on prospective methods that account for both space and time localities (15–17).

Knox (18, 19) proposed a method that allows for statistical testing of the interaction of incidents of infectious disease in space and time that does not suffer from the limitation of density measures that use an arbitrary critical value of density for determining risk localities. The Knox statistic is calculated by pairing all possible data points (e.g., location in space and time of the death of birds) within a clearly defined geographic area and temporal window and testing them against assigned values of what is "close" in space and time. The number of close space-time data pairs is compared with what would be expected if there were no interaction of space and time, and a probability of nonrandom space-time interaction is determined. When the probability is less than 0.05, the likelihood of space-time interaction is significant.

The Knox test is widely used in epidemiologic studies (20–24). However, many of these studies are retrospective with the intent of evaluating infectious etiology from disease incidence in single unpartitioned areas. Only recently has the Knox test been implemented prospectively for local regions (15–17). Rogerson (17) used the Knox test on data updated prospectively with a cumulative sum method; however, this method did not account for dynamic spatial phenomena that exhibit spatial movement in time (12). In the case of West Nile virus in New York City, prospective monitoring is necessary for targeting remediation and control efforts, and dynamic monitoring is essential for tracking the changing spatiality of viral activity. Dynamic monitoring reduces false positive-risk areas where viral activity has subsided, it keeps remediation and control efforts focused on the current viral "hot spots," and it can be used to monitor the efficacy of remediation and control initiatives.

The Dynamic Continuous-Area Space-Time (DYCAST) system was developed to identify and prospectively monitor high-risk areas for West Nile virus, and it was used to assist in guiding the remediation and control efforts of the New York City Department of Health. The DYCAST model is prospective and dynamic, and it relies on the Knox test to statistically assess the significance of space-time interaction and, hence, risk. Reporting bias is decreased using this approach, because the decision to identify a high-risk area is not dependent on a cumulative density measure but on a statistical assessment of the significance of space-time interaction. In fact, significant space-time interaction can occur in low-density situations, and nonsignificant interaction can occur in high-density situations. The Knox analysis is implemented as an interpolation function to create a surface of probabilities over a grid of 1,400 cells overlaying New York City. Each grid cell is assigned a probability based on a Knox analysis of dead birds within a 1.5-mile (2.41-km) radius of its centroid (spatial domain) and a 21-day moving window (temporal domain) preceding each day's daily run. The DYCAST model was calibrated using year 2000 data on dead bird and human West Nile virus incidence reports and was implemented and tested operationally with year 2001 surveillance data.

MATERIALS AND METHODS

Data

Data on human cases and dead bird reports were provided by the New York City Department of Health for the years 2000 and 2001. Dead bird reports were collected by passive surveillance, relying on public reporting through telephone and Internet entry systems. All dead bird reports were geocoded (assigned a coordinate to the street address of where the dead bird was found) and entered into the geographic information system, GE Smallworld technology (25). Attribute information, such as species and found date, was attached to each record. The complete DYCAST system, including the Knox methodology for space-time interaction of dead bird reports, was modeled and written in Magik programming language (25).

SPACE

FIGURE 1.   Construction of the matrix as suggested by Knox (18).

Methodology

The major assumptions of the model used in this study are as follows: 1) West Nile virus is a continuous phenomenon across space; 2) humans are infected at their place of residence; 3) nonrandom space-time interaction of bird deaths is attributed to West Nile virus infection; and 4) each dead bird has an equal opportunity of being reported.

Knox test.   The Knox method is used to test for no interaction of incidents in space and time within a clearly defined spatial and temporal domain. Closeness in space and time is based on a set of criteria (e.g., ecology), and pairs of data points are tested as to which of four categories they fall into: close in space only; close in time only; close in space and time; or close in neither space nor time. Knox (19) suggests the construction of a 2×2 contingency table as shown in figure 1. $T(o_{11})$ is the test statistic or the actual number of pairs found close in space and time, and it is calculated as

$$T(o_{11}) = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} s_{ij} t_{ij},$$

where s_{ij}, t_{ij} is 1 when the ith-jth pair is close in space and/or time, respectively, and 0 otherwise.

The Knox statistic compares the observed number of pairs close in space and time with the expected number of pairs close in space and time under a random process, given that $s = o_{11} + o_{21}$ pairs were found close in space and that $t = o_{11} + o_{12}$ pairs were found close in time (19). The expected number of pairs is calculated as

$$E(T) = \left(\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} t_{ij} \sum_{i-1}^{n-1} \sum_{j=i+1}^{n} s_{ij} \right) / N.$$

The variance of the Knox statistic was developed by Barton and David (26).

Model calibration.   Residential location and the presumed date of West Nile virus infection in humans were used as the basis for model calibration, because they were considered to be the most reliable indicators of amplified West Nile virus activity. A Knox test was performed on all dead birds, except pigeons and unknown species, found within a 1.5-mile (2.41 km) spatial domain of the residential location of each human case and within a 21-day temporal domain prior to the case's presumed date of West Nile virus infection. The presumed

date of human infection was estimated as 7 days prior to the reported date of onset of symptoms, which is above the mean incubation period of hospitalized patients, 5.3 days (27), and within the range of human infection, 3–15 days (28). The 1.5-mile (2.41 km) buffer represented local areas of relatively high risk based on twice the feeding distance of *C. pipiens*, 0.68 miles (1.09 km) (29). The 21-day window accounted for two infectious cycles in birds (e.g., infected birds die within 7 days) and the possibility of a spillover effect. Statistical significance was evaluated at different combinations of critical space-time parameters within the spatial and temporal domains.

Selection of critical parameters.   A challenging task in the Knox methodology is the selection of critical parameters. Because of the uncertainty in their statistical significance, studies will usually set a "range" and systematically perform the Knox test over the span of the range (23, 24). However, the inference of space-time interaction, based on these ranges over the same data set, results in multiple testing (14, 24). Additionally, a purely statistical decision for space-time parameter selection disregards factors inherent to the nature of the phenomena that are being studied.

In the DYCAST model, the selection of values for the critical parameters happens in the calibration phase and is based in part on ecologic considerations. Critical distance (or measure of space) does not exceed 0.75 miles (1.2 km), reflecting the limited mobility of ill birds and avoiding a distance close to the spatial domain that would reduce the test to one of temporal clustering. Critical time lies between 2 and 7 days, reflecting the period within which infected birds experience limited mobility and die. The Knox test was therefore run for the distances of 0.25 miles (0.4 km), 0.5 miles (0.8 km), and 0.75 miles (1.2 km) and the times of 3 and 6 days, producing six combinations of critical parameters. The actual number of close pairs of dead birds was counted, and the expected number of close pairs and variance (30) were calculated. The probability of significance was assessed using 1) Poisson (if one or more of the Knox contingency table cells had less than five pairs) or chi-square (if all the Knox contingency table cells had five or more pairs) distribution (19, 20, 30), 2) a normal approximation (30), and 3) 1,000 Monte Carlo random permutations as adapted by Mantel (31) at the $p = 0.05$ level. The critical parameter combination that resulted in significance of space-time interaction of bird deaths within the spatial and temporal domains (1.5-mile buffer, 21 days prior to presumed infection date) for the greatest number of the year 2000 human cases was chosen as the optimal critical parameter combination and set for all future city-wide analyses.

Spatial design.   The city-wide spatial design of the model was used for both the retrospective calibration phase in the year 2000 and for prospective implementation in the year 2001. The spatial design consists of laying a 0.5-mile grid across New York City, running the Knox test on each cell centroid ($n = 1,400$) as if it were the center of a high-risk area, and assigning to the cell the resulting probability. The localized Knox test uses a 1.5-mile radius as its spatial domain and a 21-day window as its temporal domain. The rationale for using a 1.5-mile-radius domain to evaluate the probability for a 0.5-mile grid cell is that West Nile virus activity is a contin-

TABLE 1. Significance testing results of dead bird interaction by critical parameter set for each year 2000 human case in New York City

Case identification no. and significance test†	Onset date	No. of birds	Distance and time					
			0.4 km		0.8 km		1.2 km	
			3 days	6 days	3 days	6 days	3 days	6 days
Case 484	July 20, 2000	35						
Poisson or chi-square distribution			0.183	0.299	0.074*	0.990	0.589	0.047*
Normal distribution			0.304	0.405	0.341	0.479	0.466	0.556
Monte Carlo trial			0.026*	0.001*	0.001*	0.001*	0.001*	0.003*
Case 624	July 28, 2000	45						
Poisson or chi-square distribution			0.129	0.381	0.110	0.113	0.007*	0.165
Normal distribution			0.718	0.604	0.636	0.598	0.658	0.544
Monte Carlo trial			0.371	0.072*	0.112	0.070*	0.056*	0.003*
Case 625	August 4, 2000	171						
Poisson or chi-square distribution			0.001*	0.001*	0.146	0.627	0.141	0.987
Normal distribution			0.140	0.314	0.438	0.485	0.461	0.506
Monte Carlo trial			0.001*	0.001*	0.001*	0.001*	0.001*	0.001*
Case 759	August 8, 2000	103						
Poisson or chi-square distribution			0.086	1.000	0.468	0.109	1.000	0.018*
Normal distribution			0.327	0.459	0.455	0.445	0.505	0.440
Monte Carlo trial			0.001*	0.001*	0.001*	0.001*	0.001*	0.001*
Case 1,284	August 10, 2000	79						
Poisson or chi-square distribution			0.004*	1.000	0.011*	0.111	0.001*	0.001*
Normal distribution			0.194	0.495	0.333	0.443	0.218	0.382
Monte Carlo trial			0.001*	0.001*	0.001*	0.001*	0.001*	0.001*
Case 879	August 15, 2000	54						
Poisson or chi-square distribution			0.064	0.002*	0.471	0.115	1.000	0.078
Normal distribution			0.160	0.152	0.402	0.378	0.490	0.403
Monte Carlo trial			0.001*	0.001*	0.031*	0.027*	0.012*	0.001*
Case 912	August 15, 2000	215						
Poisson or chi-square distribution			0.001*	0.001*	0.001*	0.060*	0.001*	0.038*
Normal distribution			0.023*	0.211	0.229	0.417	0.289	0.439
Monte Carlo trial			0.001*	0.001*	0.001*	0.001*	0.001*	0.001*

Table continues

uous phenomenon and, therefore, should be modeled as a continuous surface rather than as a collection of discrete adjoining regions. The creation of surfaces has traditionally been accomplished using a function that interpolates the value of a cell based on its neighborhood (32). The grid's centroid spacing of 0.5 miles (0.8 km) (0.707 miles diagonally) was selected to be less than half the feature of interest (e.g., the 1.5-mile range of *C. pipiens*) to avoid information loss during the interpolation process (33). Similar grid size selection procedures are widely used in remote sensing (34). Treating West Nile virus activity as a surface also avoids the MAUP associated with the arbitrary partitioning of space.

RESULTS

Model calibration

Critical parameter calibration. The critical parameters of 0.25 miles (0.4 km) and 3 days were selected as the best

space-time combination for use in this model's Knox test. These parameters demonstrated agreement in significance of the Poisson or chi-square assessment with the Monte Carlo trial in nine (64 percent) of the 14 human cases (although this comparison is not adjusted for chance agreement). Furthermore, although the normal approximation showed only four significant cases of space-time interaction, they were all found at the combination of 0.25 miles (0.4 km) and 3 days (table 1). When the critical time was extended to 6 days, agreement between the Poisson or chi-square test and the Monte Carlo trial was observed in only seven (50 percent) human cases. As the critical distance was extended to 0.5 miles (0.8 km) and the critical time was kept at 3 days, agreement in significance was observed in only four (29 percent) human cases. At 0.5 miles (0.8 km) and 6 days, there was agreement in only one human case. When critical distance was increased to 0.75 miles (1.2 km) and critical time was kept at 3 days, there was agreement in four (29 percent)

TABLE 1.   Continued

| Case identification no. and significance test	Onset date	No. of birds	Distance and time					
			0.4 km		0.8 km		1.2 km	
			3 days	6 days	3 days	6 days	3 days	6 days
Case 1,165	August 15, 2000	159						
Poisson or chi-square distribution			0.001*	0.002*	0.003*	0.239	1.000	0.123
Normal distribution			0.022*	0.250	0.301	0.448	0.505	0.435
Monte Carlo trial			0.001*	0.001*	0.001*	0.001*	0.001*	0.001*
Case 876	August 18, 2000	71						
Poisson or chi-square distribution			0.003*	0.007*	0.715	0.566	0.602	0.238
Normal distribution			0.083	0.195	0.450	0.542	0.535	0.543
Monte Carlo trial			0.001*	0.001*	0.001*	0.011*	0.001*	0.001*
Case 981	August 23, 2000	90						
Poisson or chi-square distribution			1.000	0.079	0.392	0.097	0.190	0.119
Normal distribution			0.552	0.659	0.433	0.419	0.436	0.451
Monte Carlo trial			0.088	0.285	0.001*	0.001*	0.001*	0.001*
Case 1,031	August 27, 2000	39						
Poisson or chi-square distribution			0.001*	0.001*	0.003*	0.016*	0.013*	0.044*
Normal distribution			0.023	0.107	0.163	0.327	0.285	0.395
Monte Carlo trial			0.010*	0.001*	0.001*	0.001*	0.001*	0.001*
Case 1,002	August 31, 2000	22						
Poisson or chi-square distribution			0.420	0.586	0.140	0.128	0.159	0.053*
Normal distribution			0.327	0.391	0.304	0.347	0.364	0.357
Monte Carlo trial			0.112	0.264	0.113	0.050*	0.149	0.179
Case 1,023	September 2, 2000	181						
Poisson or chi-square distribution			0.002*	1.000	0.988	0.124	1.000	0.001*
Normal distribution			0.159	0.484	0.473	0.563	0.501	0.605
Monte Carlo trial			0.001*	0.001*	0.001*	0.001*	0.001*	0.001*
Case 1,157	September 13, 2000	16						
Poisson or chi-square distribution			0.023*	0.048*	1.000	0.439	0.706	0.209
Normal distribution			0.032*	0.076	0.507	0.615	0.565	0.618
Monte Carlo trial			0.001*	0.001*	0.052*	0.316	0.001*	0.221

* $p < 0.05$.
† Spatial and temporal domains: buffer, 1.5 miles (2.41 km); window, 21 days.

human cases, and at 0.75 miles (1.2 km) and 6 days, agreement increased to six (43 percent) human cases. This decrease and then increase in the numbers of significant space-time interaction areas suggest that, as the critical distance approaches the 1.5-mile buffer size, most pairs are close in space, and the Knox test is reduced to a test of pairs close in time. Hence, the probability results refer to a related but different test quantity. In summary, the critical parameter combination of 0.25 miles (0.4 km) and 3 days proved to be the most appropriate for this study because of its strong likelihood of avoiding false negatives.

Significance testing.   The Poisson and chi-square tests for significance assume that the data used for testing are independent of each other. Data pairs that share points in the Knox procedure violate this assumption of independence. In this study, the violation proved to be particularly acute when the number of data points used (birds) was small. This problem may explain the disagreement shown between the Poisson or chi-square test and the Monte Carlo trial for small numbers of birds. For example, with case 1,002, 22 birds were found within the 1.5-mile buffer and 21-day window. Significance was shown with the Monte Carlo trial at 0.5 miles (0.8 km) and 6 days ($p = 0.05$), but the Poisson or chi-square test showed no significance ($p = 0.128$). An examination of the seven close pairs formed from these 22 birds showed that six pairs shared at least one data point, indicating statistical instability due to a violation in the pair independence assumption. The results of the retrospective city-wide analysis on year 2000 data demonstrated additional problems associated with the diverging results of the Poisson or chi-square test and the Monte Carlo trial and the dependencies at smaller bird numbers. Therefore, a threshold was
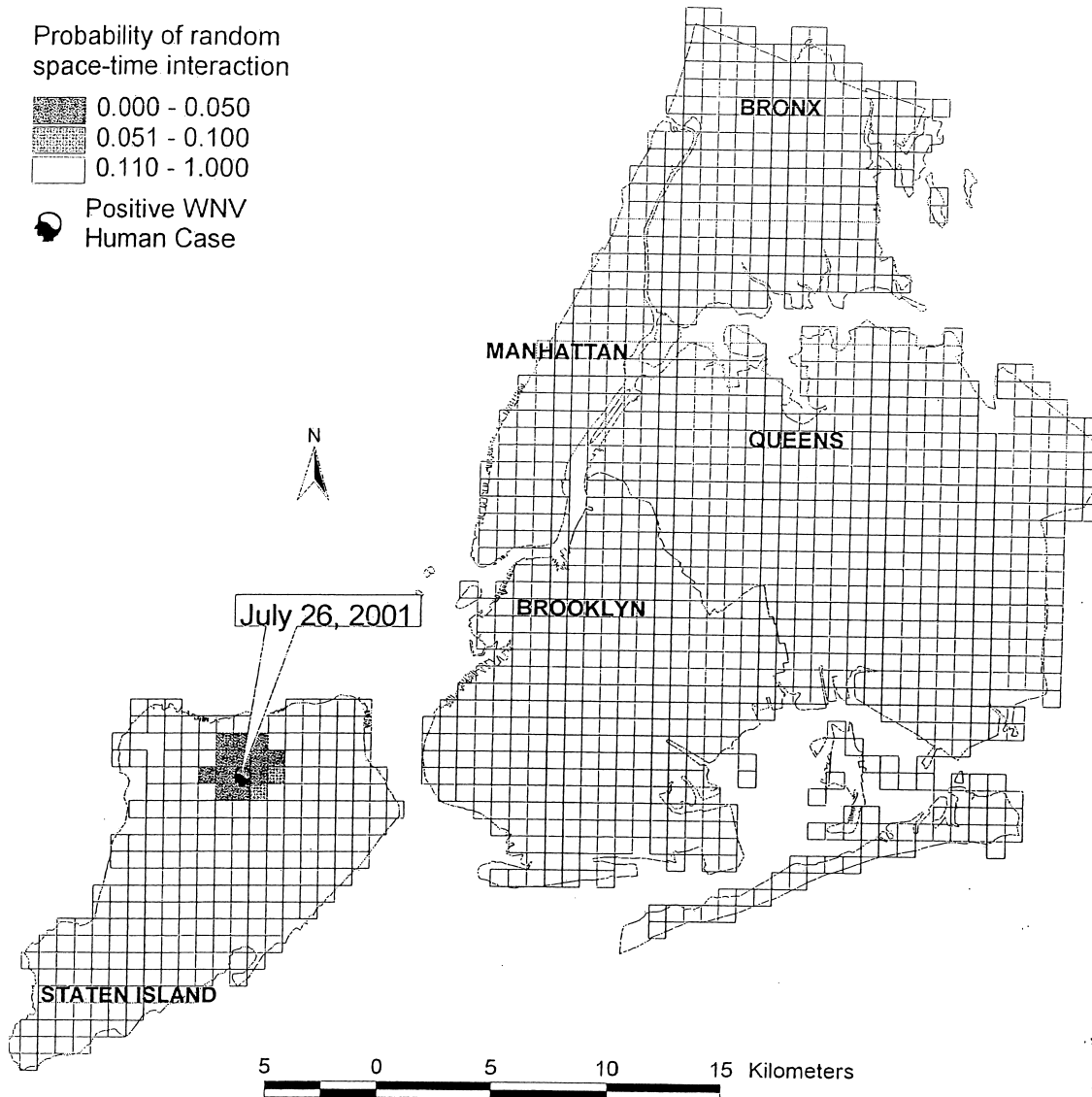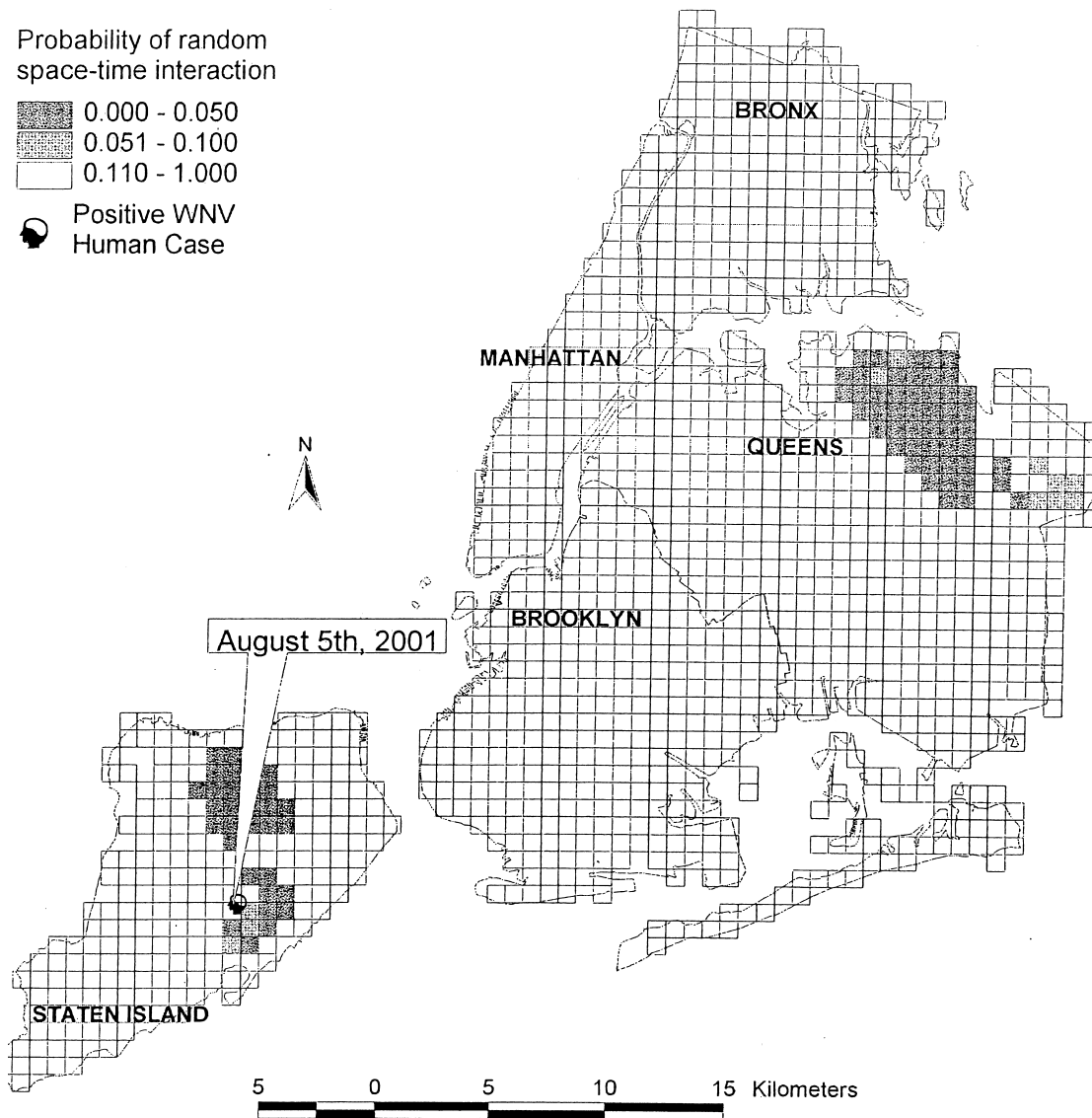
FIGURE 2.   July 2, 2001, New York, New York. Localized Knox test analysis with the following parameters: buffer, 1.5 miles (2.41 km); temporal window, 21 days; close-in space, 0.4 km; close-in time, 3 days; and bird threshold, 25 birds. Data source: New York City Department of Health and New York City map. Note the onset date of the human case that is July 26, 2001, in the identified high-risk area. WNV, West Nile virus.

set at 25 birds and was based on 1) the convergence between the Poisson or chi-square and the Monte Carlo probabilities as the numbers of birds increased, 2) the attempt to reduce pair dependencies that could bias the results in both the Poisson or chi-square and the Monte Carlo methods, and 3) the attempt to minimize false positives at lower bird numbers. For example, a statistically significant result of one close pair with only two data points would not raise the risk (and should not raise the risk alarm) in that area. Because the results of the Poisson or chi-square test with the Monte Carlo trial were similar after the threshold was implemented, all year 2001 analyses were run using only the Poisson or chi-square significance test.

Model implementation

The results of this study demonstrate that the DYCAST model was successful in identifying areas of high risk for West Nile virus at least 13 days prior to the onset of illness in five of the seven human cases in the year 2001. One of the missed cases appeared in an area that was indicated by the model 3 days after the person's onset of illness; however, this case occurred after dead bird surveillance was virtually halted on September 11, 2001, because of the terrorist attack on the World Trade Center buildings in New York City.

A high-risk area for West Nile virus was identified on July 2nd in northern Staten Island, 25 days prior to the first human case of West Nile virus with onset of illness on July
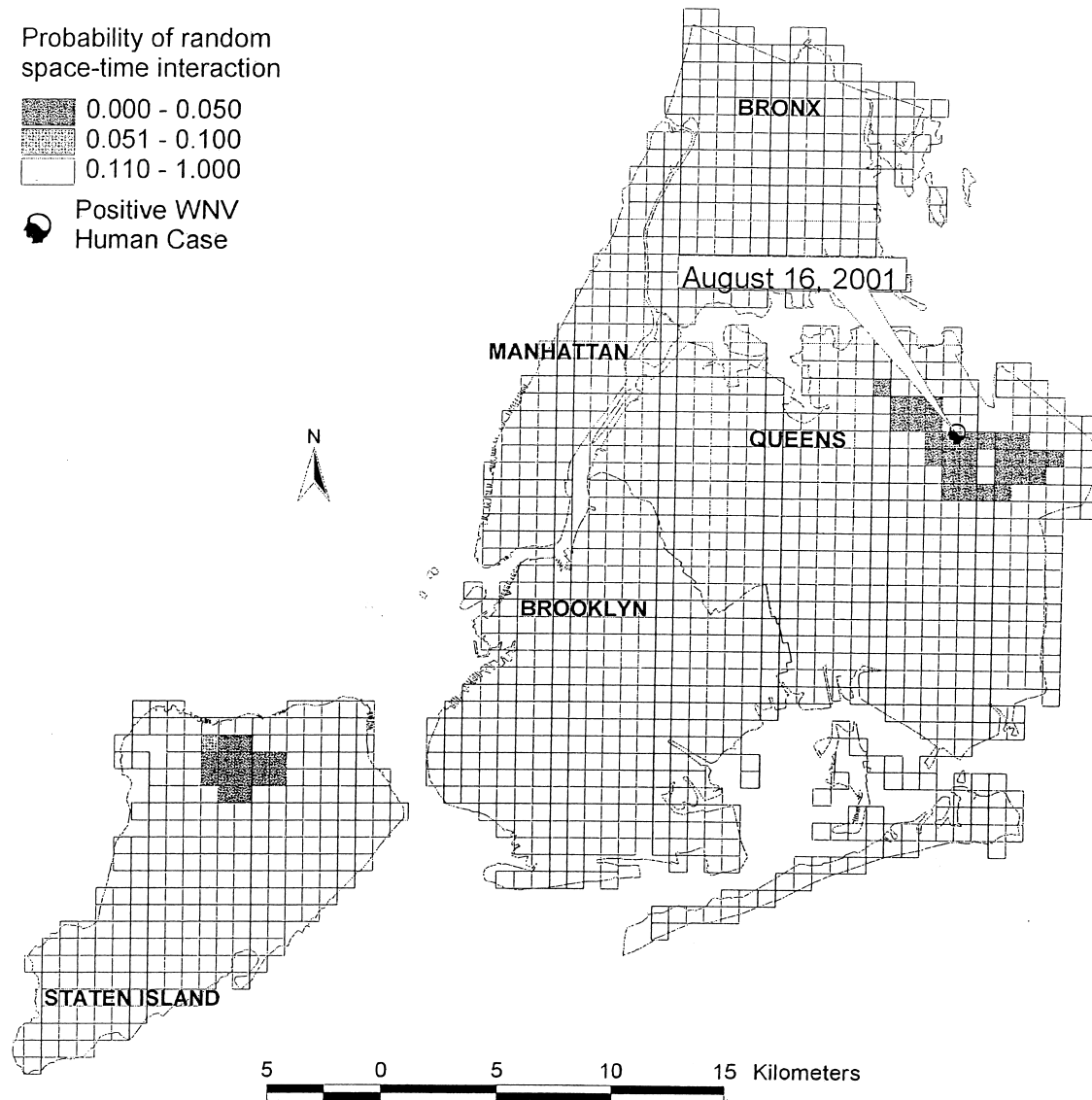
FIGURE 3. July 19, 2001, New York, New York. Localized Knox test analysis with the following parameters: buffer, 1.5 miles (2.41 km); temporal window, 21 days; close-in space, 0.4 km; close-in time, 3 days; and bird threshold, 25 birds. Data source: New York City Department of Health and New York City map. Note the onset date of the human case that is August 5, 2001, in the identified high-risk area. WNV, West Nile virus.

26th (figure 2). South of this location, another area of significant space-time interaction of dead birds appeared 17 days prior to the second human case with onset of illness on August 5th (figure 3). In northeastern Queens, an area of high risk for West Nile virus appeared on July 9th, 38 days prior to the third human case with onset of illness on August 16th (figure 4). Subsequently, significant space-time interaction of dead birds was observed in southwestern Brooklyn on August 20th, 13 days prior to the fourth human case with onset of illness on September 2nd (figure 5). This area of high risk for West Nile virus in Brooklyn was an expansion of an area first observed on August 4th, which persisted and expanded until September 20th. In northern Queens, signifi-

cant space-time interaction of dead birds was seen from August 3rd through August 28th, 35 days prior to the fifth human case with onset of illness September 7th.

Two days later, on September 9th, a sixth human case of West Nile virus was diagnosed in central Brooklyn, where the DYCAST system failed to capture significant space-time interaction at any time. In the aftermath of the terrorist attacks on the World Trade Center buildings in New York City, the model failed to capture significant interaction, reflecting decreased surveillance of dead birds. During this time, all areas showing significant city-wide interaction gradually receded and disappeared completely 9 days later on September 22nd. These findings demonstrate both the

FIGURE 4. July 19, 2001, New York, New York. Localized Knox test analysis with the following parameters: buffer, 1.5 miles (2.41 km); temporal window, 21 days; close-in space, 0.4 km; close-in time, 3 days; and bird threshold, 25 birds. Data source: New York City Department of Health and New York City map. Note the onset date of the human case that is August 16, 2001, in the identified high-risk area. WNV, West Nile virus.

temporal sensitivity and the dynamics of the model. The 9-day interval was not necessarily a disadvantage of the DYCAST system, as there was interest in both current and prior areas of significant interaction for targeting remediation and control efforts. Furthermore, a decrease in birds during this time could potentially have caused mosquitoes to switch feeding habits to mammals, increasing the risk for humans.

When bird surveillance did resume in early October, the model demonstrated significant space-time interaction on October 8th in lower Manhattan, just 2 days after the seventh and last human case of West Nile virus with onset of illness on October 6th (figure 6). This same area of high risk

expanded and remained significant until October 31st, the last date for which we conducted the analysis.

A review of these analyses also demonstrates significant space-time interaction of dead birds at the edges of all the boroughs, although a potential edge effect could have existed when the edge caused the reporting numbers in a buffer to drop below the threshold ($n = 25$). However, this was not a problem in this study. Edge effect concerns were also viewed in the perspective of the amplification cycle, which suggests that a reduced number of birds at the edges (i.e., birds will not exist as mosquito targets in the water) will depress the amplification cycle in the buffer.
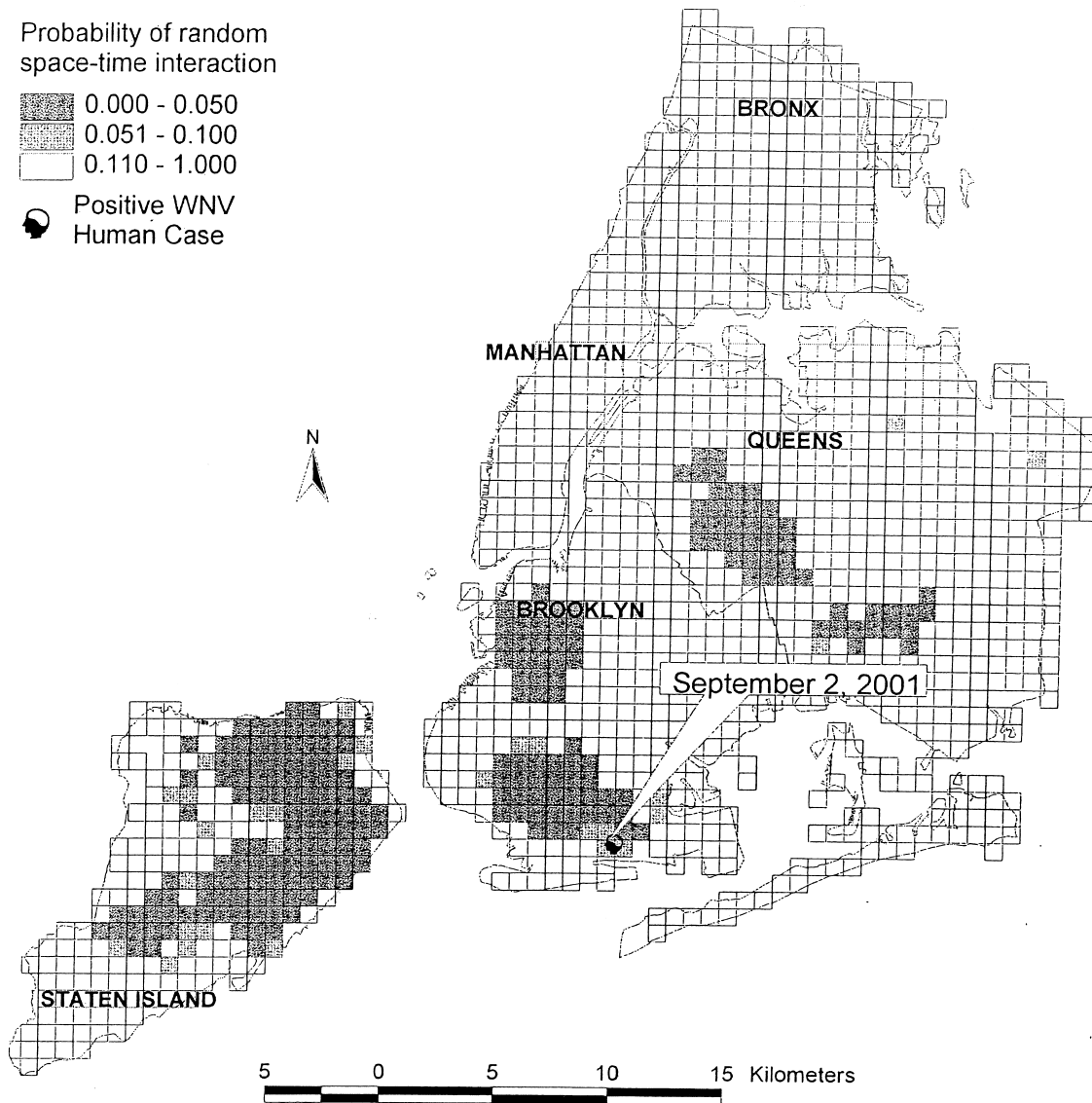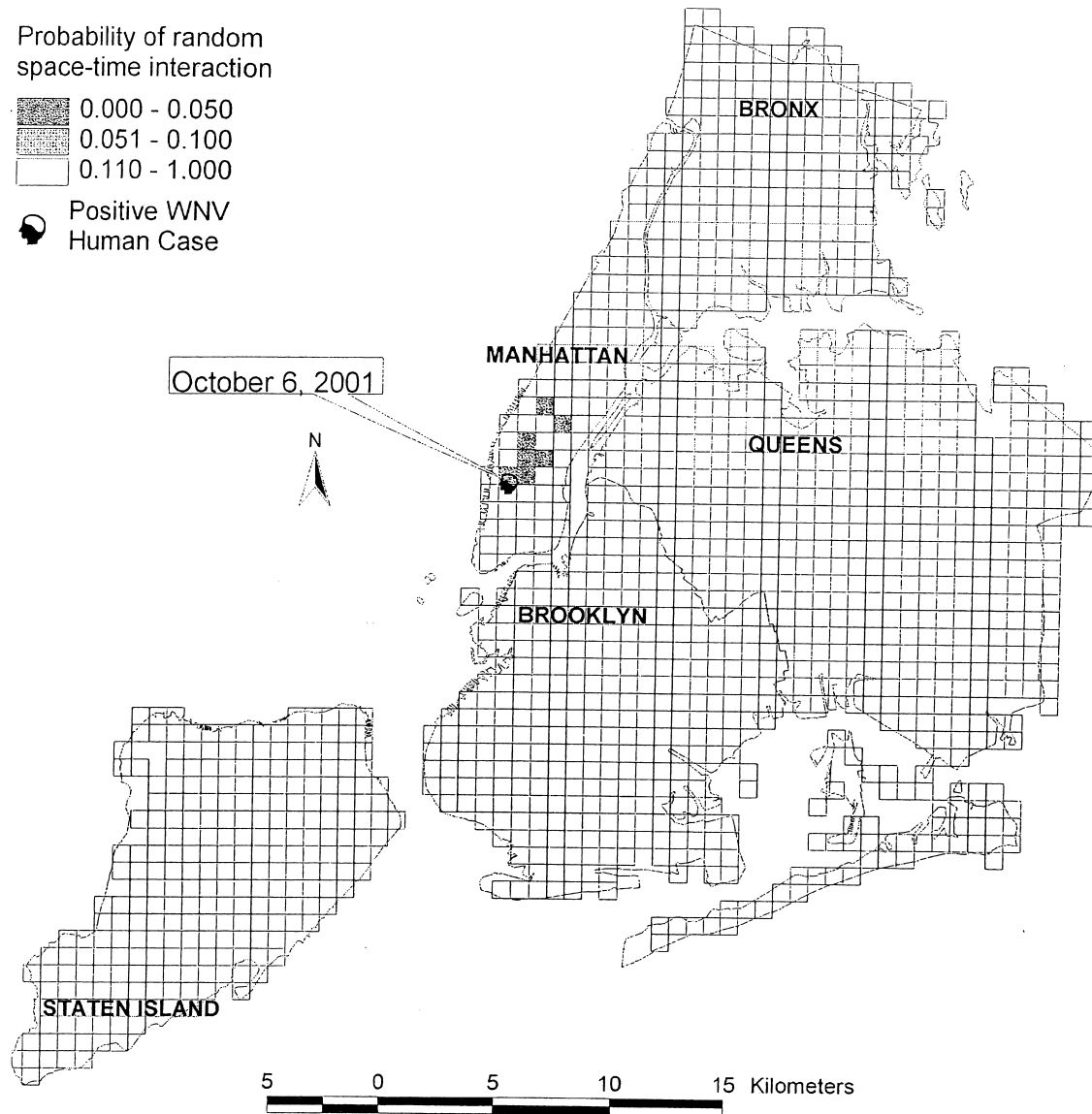
FIGURE 5.   August 23, 2001, New York, New York. Localized Knox test analysis with the following parameters: buffer, 1.5 miles (2.41 km); temporal window, 21 days; close-in space, 0.4 km; close-in time, 3 days; and bird threshold, 25 birds. Data source: New York City Department of Health and New York City map. Note the onset date of the human case that is September 2, 2001, in the identified high-risk area. WNV, West Nile virus.

Sensitivity analysis.   A sensitivity analysis was performed on data for September 1, 2001, because this date reflected the peak of the season's West Nile virus activity. The results of the analysis using the critical parameter combinations of 0.25 miles (0.4 km), 0.5 miles (0.8 km), and 0.75 miles (1.2 km) with 3 and 6 days are summarized in table 2. This table shows the number of cells that were identified as being high-risk areas on the basis of the dead bird interaction effects with the respective parameters, at the $p = 0.05$ and the $p = 0.1$ significance levels.

The more constrained critical parameters (0.25 miles (0.4 km) and 3 days) produced the maximum number of cells in the

city-wide grid that were identified as high risk. Furthermore, the temporal critical parameter appears to be more important than the spatial critical parameter, because the combination of 0.25 miles (0.4 km) and 6 days produced fewer high-risk cells than did the combination of 0.5 miles (0.8 km) and 3 days. There is also a decrease in the number of high-risk cells with increasing space distance. Critical parameters with the least number of significant buffers appear at the 0.75-mile (1.2-km) and 6-day critical parameter combination. This sensitivity analysis validates the use of the most constrained critical parameters to minimize false negatives. Despite differences in the number of high-risk cells, all critical combinations resulted

FIGURE 6. October 9, 2001, New York, New York. Localized Knox test analysis with the following parameters: buffer, 1.5 miles (2.41 km); temporal window, 21 days; close-in space, 0.4 km; close-in time, 3 days; and bird threshold, 25 birds. Data source: New York City Department of Health and New York City map. Note the onset date of the human case that is October 6, 2001, in the identified high-risk area. WNV, West Nile virus. (Retrospectively because of September 11, 2001).

in high-risk cells in the same approximate areas with minor differences in extent. This finding demonstrates that the model is highly robust with respect to its critical parameters.

In summary, the DYCAST system was stable over the range of ecologically constrained parameters, and the most constrained parameters (0.25 miles (0.4 km) and 3 days) were the best for initiating remediation and control activities, because they minimized false negatives.

DISCUSSION

In conclusion, in the year 2001, the DYCAST system successfully identified areas of high risk for West Nile virus

in humans at least 13 days prior to their onset of illness in five of the seven human cases. The reason for this success is the unique combination of geographic principles, ecologic knowledge of the West Nile virus vector-host transmission cycle, and statistical analyses to overcome the limitations observed in previous studies to model West Nile virus.

The innovations of the DYCAST system include the use of a statistical measure to assess areas of high risk for West Nile virus rather than the selection of an arbitrary "critical" density and the use of data regarding the ecology of the host and vectors in the parameter calibration phase that attunes the model to the real world phenomenon. In addition, the DYCAST system did not suffer from MAUP inconsisten-

TABLE 2.   Sensitivity analysis for critical parameters, September 1, 2001, New York, New York

Distance and p values	Time			
	3 days		6 days	
	No.	%	No.	%
0.025 mile (0.4 km)				
p value < 0.05	254	18.8	156	11.5
p value = 0.05–0.1	26	1.9	34	2.5
Nonsignificant	1,072	79.3	1,162	85.9
Total	1,352	100.0	1,352	100.0
0.05 mile (0.8 km)				
p value < 0.05	168	12.4	115	8.5
p value = 0.05–0.1	26	1.9	28	2.0
Nonsignificant	1,158	85.7	1,209	89.3
Total	1,352	100.0	1,352	100.0
0.75 mile (1.2 km)				
p value < 0.05	121	8.9	59	4.4
p value 0.05–0.1	40	3.0	48	3.6
Nonsignificant	1,191	88.2	1,245	92.1
Total	1,352	100.0	1,352	100.0

cies, because it was designed as a continuous area system that used a neighborhood interpolation technique in accordance with centroid spacing selection principles. Moreover, the identification of areas of high risk for West Nile virus in locations of low population density (i.e., Staten Island) demonstrates no visible evidence of reporting bias when a minimum bird threshold was calculated and implemented. Edge effects were reduced to cases where an edge spuriously caused the number of dead bird data points not to reach the appropriate 25-bird threshold. Finally, the dynamic nature of the system provided timely identification and incorporation of temporal effects.

Despite the success of the DYCAST system, certain issues require further research. Two points arising from the calibration results include the issue of thresholding and the validity of significance tests. Another important issue that should be addressed is reporting bias in areas of varied socioeconomic characteristics. Finally, the model could be optimized further by incorporating additional data, such as the location of mosquito breeding grounds, environmental conditions, high-risk populations, and feedback mechanisms for further calibration on positive bird and mosquito results and control activities.

REFERENCES

1. New York City Department of Health and Mental Hygiene. City health information. West Nile virus surveillance and control: an update for healthcare providers in New York City. New York, NY: New York City Department of Health, May 2001:1–8. (http://www.ci.nyc.ny.us/html/doh/pdf/chi/chi20-2.pdf).
2. New York City Department of Health and Mental Hygiene. City health information. Vector-borne and zoonotic diseases in New York City—an update. New York, NY: New York Department of Health, June 2002:1–4. (http://www.ci.nyc.ny.us/html/doh/pdf/chi/chi21-1.pdf).
3. Miller JR. The control of mosquito-borne diseases in New York City. J Urban Health 2001;78:359–66.
4. Despommier DD. West Nile story. New York, NY: Apple Trees Productions, 2001.
5. Turell MJ, O'Guinn M, Oliver J. Potential for New York mosquitoes to transmit West Nile virus. Am J Trop Med Hyg 2000; 62:413–14.
6. Means RG. Host preferences of mosquitoes (Diptera: Culcidae) in Suffolk County, NY. Ann Entomol Soc Am 1968;61:116–20.
7. Turell MJ, Sardelis MR, Dohm DJ, et al. Pontential North American vectors of West Nile virus. Ann N Y Acad Sci 2001; 951:317–24.
8. Eidson M, Kramer L, Stone W, et al. Dead bird surveillance as an early warning system for West Nile virus. Emerg Infect Dis 2001;7:631–5.

9. Eidson M, Miller M, Kramer L, et al. Dead crow densities and human cases of West Nile virus, New York State. Emerg Infect Dis 2001;7:662–4.

10. Fotheringham SA, Wong SWD. The modifiable areal unit problem in multivariate statistical analysis. Environ Plan A 1991; 23:1025–44.

11. Baily CT, Gatrel CA. Interactive spatial data analysis. Essex, England: Logman Scientific & Technical, 1995.

12. Lawson A. Statistical methods in spatial epidemiology. New York, NY: John Wiley & Sons, Inc, 2001.

13. Daley DJ, Gani J. Epidemic modeling. Cambridge, United Kingdom: Cambridge University Press, 1999.

14. Kulldorff M, Hjalmars U. The Knox method and other tests for space-time interaction. Biometrics 1999;55:544–52.

15. Kulldorff M. Prospective time periodic geographical disease surveillance using a scan statistic. J R Stat Soc (A) 2001;164: 61–72.

16. Rogerson PA. Surveillance systems for monitoring the development of spatial patterns. J R Stat Soc Assoc 1997;16:2081–93.

17. Rogerson PA. Monitoring point patterns for the development of space-time clusters. J R Stat Soc (A) 2001;164:87–96.

18. Knox GE. Detection of low intensity epidemicity. Application to cleft lip and palate. Br J Prev Soc Med 1963;17:121–7.

19. Knox GE. The detection of space-time iterations. J R Stat Soc (C) 1964;13:25–9.

20. Knox GE. Epidemiology of childhood leukaemia in Northumberland and Durham. Br J Prev Soc Med 1964;18:17–24.

21. Anderson R, Hugander A, Thulin A, et al. Clusters of acute appendicitis: further evidence for an infectious aetiology. Int J Epidemiol 1995;4:829–33.

22. Samuelson U, Johansson C, Carstensen J, et al. Space-time clustering in insulin-dependent diabetes mellitus, South-East Sweden. Int J Med 1994;23:138–42.

23. Gilman EA, Knox EG. Childhood cancers: space-time distribution in Britain. J Epidemiol Community Health 1995;49:158–63.

24. Williams GW. Time-space clustering in disease. In: Cornell GR, ed. Statistical methods for cancer studies. New York, NY: Marcel Decker, Inc, 1984.

25. GE Smallworld core spatial technology, version 3.2.1. Atlanta, GA: General Electric Network Solutions, 2002.

26. Barton DE, David FN. The random intersection of two graphs. In: David FN, ed. Research papers in statistics. New York, NY: Wiley, 1966.

27. Nash D, Mostashari F, Fine A, et al. The outbreak of West Nile virus infection in the New York City area in 1999. N Engl J Med 2001;344:1807–14.

28. West Nile virus. Atlanta, GA: Centers for Disease Control and Prevention, 2002. (http://www.cdc.gov/ncidod/dvbid/westnile/).

29. Savage H, Miller B. House mosquitoes of the USA, *Culex pipiens* complex. Wing Beats 1995;6:8–9.

30. Barton DE, David NF, Fix E, et al. Tests for space-time interaction and a power function. In: LeCam L, Neyman J, eds. Proceedings of the Fifth Berkeley Symposium on Mathematics, Statistics, and Probability. Vol IV. Berkeley, CA: University of California Press, 1967:217–28.

31. Mantel N. The detection of disease clustering and a generalized regression approach. Cancer Res 1967;27:209–20.

32. Tobler W. Smooth pycnophylactic interpolation for geographical regions (with discussion). J Am Stat Assoc 1979;74:519–36.

33. Shannon CE. Communications in the presence of noise. Proc Inst Radio Engineers 1949;37:10–21.

34. Sonka M, Hlavac V, Boyle R. Image processing, analysis, and machine vision. Pacific Grove, CA: Brooks/Cole Publishing Company, 1998.