# Landscape-Level Spatial Patterns of West Nile Virus Risk in the Northern Great Plains

Ting-Wu Chuang,* Christine W. Hockett, Lon Kightlinger, and Michael C. Wimberly

*Geographic Information Science Center of Excellence, South Dakota State University, Brookings, South Dakota;
South Dakota Department of Health, Pierre, South Dakota*

*Abstract.* Understanding the landscape-level determinants of West Nile virus (WNV) can aid in mapping high-risk areas and enhance disease control and prevention efforts. This study analyzed the spatial patterns of human WNV cases in three areas in South Dakota during 2003–2007 and investigated the influences of land cover, hydrology, soils, irrigation, and elevation by using case–control models. Land cover, hydrology, soils, and elevation all influenced WNV risk, although the main drivers were different in each study area. Risk for WNV was generally higher in areas with rural land cover than in developed areas, and higher close to wetlands or soils with a high ponding frequency. In western South Dakota, WNV risk also decreased with increasing elevation and was higher in forested areas. Our results showed that the spatial patterns of human WNV risk were associated with landscape-level features that likely reflect variability in mosquito ecology, avian host communities, and human activity.

## INTRODUCTION

West Nile virus (WNV) invaded the northern Great Plains in 2002, and since that time this region has become one of the most significant WNV hotspots in the United States.[1] South Dakota has been an area of particularly high risk, accounting for 3.5% (278 of 7,978) of the total WNV neuroinvasive human cases in the United States during 2003–2007 although the state has less than 0.3% of the national population. This virus is transmitted by mosquitoes via complex enzootic cycles involving avian populations and opportunistic epidemic cycles in humans and horses.[2,3] *Culex tarsalis* Coquillett is the major vector across the northern Great Plains region and in South Dakota.[4,5] The influences of the spatial patterns of environmental determinants associated with vector and host habitats are widely recognized, and there has been considerable research on the spatial epidemiology of WNV over the past decade. However, relatively few studies have focused on the sparsely-populated rural areas of the northern Great Plains, which have unique climate, physiography, land cover, and human geography compared with other parts of the United States. Thus, WNV research in the northern Great Plains has the potential to increase our understanding of how this disease persists across a diverse range of environments and to contribute to disease control and prevention in a region where WNV remains a significant public health threat.

Multiple environmental drivers can affect WNV amplification and transmission to humans. For instance, the effects of climatic variability on inter-annual fluctuations of mosquito abundance and disease transmission have been described in many studies.[6–8] Weather-driven hydrologic models have been applied to analyze and predict WNV human infections in Florida and Colorado.[9,10] In addition to temporal variability in weather, spatial patterns of land cover, hydrological factors, soils, and socioeconomic status also affect WNV risk because of their influences on mosquito habitat suitability, avian host communities, and human exposure to infected mosquitoes. Eisen and others found that human WNV incidence was associated with irrigated agriculture in north central Colorado,[11] and strong correlations between the seasonal distributions of *Cx. tarsalis* and elevation gradients were found in the same area.[12] Studies in California found space-time clusters of WNV positive mosquito pools and WNV positive dead birds, highlighting the influences of environmental heterogeneity on WNV transmission.[13,14] The increased number of home foreclosures and abandoned swimming pools affected urban mosquito populations and WNV risk in Kern County, California.[6,15] Ruiz and others found that the risk factors for WNV in the Chicago region included vegetation, age of house, income, distance to WNV-positive dead birds, race, mosquito abatement, and geologic factors.[16] An association of lower income areas with higher WNV incidence has been described in Orange County, California.[17]

Most of these studies were conducted in urban, suburban, or mountainous areas. In contrast, South Dakota is located in the Great Plains and has distinctive landscapes and environmental characteristics. The cities in the eastern part of South Dakota sit on the Coteau des Prairies and in the James River lowlands, and are surrounded by mosaics of cropland and grassland. The hydrology is characterized by a few large river basins, extensive shallow aquifers, and numerous small wetlands. In contrast, the landscapes of western South Dakota vary along a gradient from arid, low-elevation rangelands to mesic forests at higher elevations in the Black Hills. A regional analysis of the spatial patterns of WNV incidence found that risk was associated with specific climatic conditions and was higher in counties with large rural populations and high percentages of irrigated agriculture.[18] A national-level study demonstrated a correlation between WNV human incidence and grassland in the Upper Plains region.[19] However, more focused landscape-level studies are needed to more clearly elucidate the effects of land cover on WNV risk.

During 2003–2007, 553 cases were reported in and around the cities of Sioux Falls, Rapid City, and Aberdeen, South Dakota. The availability of geocoded case data from the South Dakota Department of Health provided a unique opportunity to conduct a relatively fine scale study of the environmental determinants of WNV, and to contrast the results over three diverse landscapes. Therefore, we conducted a case control study to analyze the effects of land cover types, irrigation activity, elevation, wetlands, and soil ponding frequency on WNV human incidence in these three areas during 2003–2007.

* Address correspondence to Ting-Wu Chuang, Geographic Information Science Center of Excellence, South Dakota State University, 1021 Medary Avenue, Wecota Hall, Brookings, SD 57007. E-mail: ting-wu.chuang@sdstate.edu

Our overarching hypothesis was that human risk was correlated with these landscape-level variables because of their influences on vector populations, avian host communities, and human activities influencing exposure to WNV. The two main objectives of this study were to 1) assess the relative importance of landscape-level environmental determinants of WNV within each study area, and 2) use the best models to map the spatial pattern of WNV risk within each study area.

## MATERIALS AND METHODS

**Study areas.** We examined three study areas in South Dakota, each consisting of a central city and its surrounding rural landscapes. These areas are referenced throughout the paper by the name of the major city. The Sioux Falls study area included Lincoln and Minnehaha Counties in eastern South Dakota. Sioux Falls is the largest city in South Dakota, with a population of 153,888 (2010 Census). It is located at the southern tip of the Couteau des Prairies in the valley of the Big Sioux River and is surrounded by an agricultural landscape of cropland, pasture, and hayfields. The Aberdeen study area encompassed all of Brown County, which is located in northeastern South Dakota, and the city of Aberdeen has a population of 26,091 (2010 Census). The surrounding landscapes encompass drift plains and glacial lake basins of the James River valley and are dominated by cropland, pasture, and hayfields. The Rapid City study area included portions of Lawrence, Meade, and Pennington Counties in western South Dakota. It is the second largest city in South Dakota, with a population of 67,956 (2010 Census). Elevations range from 662 to 2,179 meters, with higher elevations characterized by more precipitation and cooler summer temperature than lower elevations. Landscapes are mosaics of grassland and cropland at lower elevations and are dominated by forests at higher elevations in the Black Hills. Irrigated agriculture is relatively rare in all these areas, comprising only 0.7% of the land area in Sioux Falls, 0.8% of the land area in Aberdeen, and 2.9% of the land area in Rapid City.

**Human cases of WNV and control points generation.** To assess the spatial associations of WNV risk with hypothesized landscape-level drivers, we developed a Geographic Information System (GIS) database containing WNV cases, control points, land cover, irrigation, hydrology, soils, and elevation. All GIS datasets were converted into a custom Albers Equal Area projection for South Dakota. Unless otherwise noted, all geoprocessing steps were carried out using ArcGIS 9.3 software (ESRI, Redlands, CA).

The geocoded locations of the residences of de-identified reported human WNV cases during 2003–2007 were provided by the South Dakota Department of Health. The definition of WNV human cases followed the guidelines of the Centers for Disease Control and Prevention (Atlanta, GA) and was confirmed by serologic testing (IgM enzyme-linked immunosorbent assay) at the South Dakota Department of Health. Both WNV fever and neuroinvasive cases were reported in South Dakota and included in the analysis. Research protocols were approved by the South Dakota State University Institutional Review Board (IRB-1003005-EXM). The household address of WNV cases were used for geocoding. The geocoding success rate (80.8%) across the entire state was verified and geocoding procedures were described in a pre-

vious study.[20] Most geocoding failures were concentrated on the Indian reservations in the central part of South Dakota, and the success rates for our study areas were considerably higher: 98.3% in Sioux Falls, 91% in Aberdeen, and 93.1% in Rapid City, respectively. There were a total of 172 cases in Aberdeen, 114 cases in Sioux Falls, and 195 cases in Rapid City.

For a spatial case–control study, control points should be selected from the same geographic domain as the cases with probabilities proportional to the densities of households. Therefore, control points were generated as an inhomogeneous Poisson process accounting for census block-level population density from the 2000 census. The numbers of control points were 370, 347, and 543 in Aberdeen, Sioux Falls, and Rapid City, respectively. Similar methods have been applied by other studies.[21] The control points should not be located in areas without households, such as pastures or agricultural fields, because we used the residential addresses of cases for geocoding. For the Sioux Falls and Aberdeen study areas, we further applied a land cover mask created from the 2001 National Land Cover Database (NLCD) so that control point selection was limited to developed space where households are most likely to occur. We did not apply this masking technique in the Rapid City study area because it encompassed large expanses of forested exurban areas, which contain residences but were not identified as developed area by the NLCD.[22] The Sampling Tools function in the Hawth's Analysis Tools was applied in the ArcGIS 9.3 environment to generate the control points of the three cities.

**Land cover and irrigation data.** Land cover can serve as an indicator of mosquito habitats, avian habitats, or human activities that affect exposure to mosquitoes. We retrieved land cover information from the 2001 NLCD as a 30 meter × 30 meter raster GIS dataset. A 200-meter buffer zone was created for each case and control point and was applied to calculate the percentage of specific land cover types by using the Zonal Statistics tool. Five land cover types, including urban, open developed space, cropland, grass/hay, and forest, were summarized from the 16 classes in the original database. Urban was composed of developed areas at different levels of intensity, including residential areas, commercial/industrial regions, and mixtures of constructed materials and vegetation. Open developed space was separated from other developed space because it is composed of parks and golf-courses with large proportions of vegetation. The cultivated crops type included annual crops, primarily corn and soybeans. The grass/hay type included pastures and hayfields dominated by graminoid or herbaceous vegetation. The forest type consisted of deciduous, evergreen, and mixed forests combined. A database of irrigation permits and the geographic coordinates of their associated water draw points were obtained from the South Dakota Department of Environment and Natural Resources. The distances from every case and control point to the nearest draw point were calculated using the Euclidian Distance tool.

**Hydrologic data.** Wetlands are a preferred habitat for many migratory birds because of the availability of water, food resources, and nesting.[23] The aggregation of avian populations may increase the rate of WNV amplification by increasing the interaction between the host-seeking mosquito and resting birds.[24,25] The presence of wetlands may also be indicative of areas that pond frequently and contain smaller breeding

sites after periods of heavy rain. We used wetland GIS data from National Wetland Inventory dataset produced by the U.S. Fish and Wildlife Service. We focused our analysis on emergent wetlands, which are the most abundant type of wetland in South Dakota and are characterized by the presence of herbivorous vegetation and include a variety of herbaceous marshes, fens, swales and, wet meadows.[26] The distances from every case and control point to the nearest emergent wetland were calculated by using the Proximity Toolset.

**Soil conditions.** Mosquito breeding requires the availability of suitable water bodies. In addition to precipitation, soil characteristics are also an important determinant of the spatial patterns of breeding sites and mosquito abundance.[27,28] We used GIS soil polygon data from the Soil Survey Geographic Database to obtain information on ponding frequency, which indicates the potential for ponding in different areas. The distances from every case and control point to the nearest soil map unit with a ponding frequency > 75% were calculated by using the Proximity Toolset.

**Elevation data.** Elevation was used in the analysis of Rapid City as an indicator of the major climatic and vegetation gradients the in the study area. A 30 meter × 30 meter digital evaluation model raster file for Rapid City was generated from the National Elevation Dataset. Average elevation was calculated within each case and control buffer zone using the Zonal Summary tool. Elevation was not used in Sioux Falls or Aberdeen because the topographic relief was extremely low in these areas. The unit of elevation was converted to 100-meter intervals (divided by 100) to facilitate the interpretation of the odds ratios in the case–control models.

**Statistical models.** Logistic regression models were used to analyze the binary dependent variable (cases = 1, controls = 0). Odds ratios were calculated from the estimated coefficients to indicate the influences of the environmental variables on WNV infection risk. All the distance variables were classified as binary variables where distances falling within the lowest quartile were coded as one and all other distances were coded a zero. We used multi-model inference to compare a set of models containing different combinations of environmental variables.[29] In the Aberdeen and Sioux Falls model, four types of variables were considered for inclusion in these models. The first variable was land cover percentages, including urban, open developed space, grass/hay, and cropland. The second variable was irrigation, measured as distance to draw point. The third variable was hydrology, measured by distance to emergent wetland. The fourth variable was soil conditions, measured by distance to soils with ponding frequency > 75%.

In the Rapid City study area, forest was added to the list of potential land cover types and elevation was added as a potential variable. Because urban land cover was correlated with the other land cover variables (variance inflation factor > 2.5), the models contained either urban land cover alone or a combination of the other land cover variables to avoid multicollinearity. We examined all combinations of these variables, resulting in a total of 38 candidate models for the Sioux Falls and Aberdeen study areas and 142 models for the Rapid City study area. The models were evaluated using the corrected Akaike's Information Criterion (AICc), which provided a metric of fit that was penalized for the number of parameters in the model relative to the sample size.[29] Lower AICc indicated better model fit. Statistical modeling was carried out by using SAS 9.2 (SAS Institute, Inc., Cary, NC).

To generate maps of the relative risk of WNV, we selected the best model for each study area and used it to predict WNV risk using the equation

$$p_i = \frac{1}{1 + \exp\left(-\beta_0 - \sum_{j=1}^{v} \beta_j X_{ij}\right)}$$

where $p_i$ is the predicted probability at point $i$, $\beta_0$ is the intercept, $\beta_j$ are the coefficients for ndependent variable $j$, $v$ is the number of independent variables, and $X_{ij}$ are the values of independent variable $j$ at point $i$. In a case–control study, the values of $p_i$ cannot be interpreted as true probabilities. However, mapping $p_i$ still provides a valid representation of the relative WNV risk in different geographic areas. For this reason, we labeled the maps as gradients from low to high risk and did not report the actual values of $p_i$.

Risk maps were generated by using the Raster Calculator by applying these equations to 30-meter raster datasets that contained the same variables used to fit the models. The percentage cover of each land cover type within a 200-meter radius of each raster cell was computed by using the Focal Statistics tool. We also computed the distance from each raster cell to the nearest irrigation draw point, emergent wetland, and soil map unit with > 75% ponding frequency. These distances were converted to binary variables.

## RESULTS

The five-year cumulative WNV infection rates were 698 per 100,000 in Aberdeen, 327 per 100,000 in Rapid City, and 92 per 100,000 in Sioux Falls. Overall, the Aberdeen area had the highest infection rate during the study period and the lowest population among the three cities. The distribution of land cover types, elevation, and distances to irrigation draw points, emergent wetlands, and soils with high ponding frequency are shown in Table 1. The percentage of land cover types in the control group demonstrated similar patterns in the Aberdeen and Sioux Falls areas, whereas higher percentages of forest and urban were shown in the Rapid City area. The distances from the control points to the different hydrologic features, soil conditions, and irrigation in Rapid City were larger than the other two areas. These results illustrated the strong differences in land cover and physiography between the eastern and western sides of South Dakota.

The five models with the best fit (lowest AICc values) for each of the three study areas are shown in Table 2. In Aberdeen, the top five models were all close competitors as indicated by the difference in the Akaike's Information Criterion corrected values and the Akaike weights. The best model included open development, grass/hay, and emergent wetlands. However, a more parsimonious model with just grass/hay and emergent wetlands had only a slightly higher AICc and lower Akaike weight. All five of the top models for Aberdeen contained grass/hay and emergent wetlands. Risk for WNV increased with the percentage of grass/hay and proximity to emergent wetlands and decreased with the percentage of open development, but the 95% confidence interval for the open development odds ratio overlapped one (Table 3). Overall, there was strong support for the effects of grass/hay land cover and proximity to wetlands in Aberdeen, but relative weak support for an effect of open development.

TABLE 1
Descriptive statistics of environmental variables in the three study areas*

| | Aberdeen | | Sioux Falls | | Rapid City | |
|---|---|---|---|---|---|---|
| | Case (n = 172) | Control (n = 370) | Case (n = 114) | Control (n = 347) | Case (n = 195) | Control (n = 543) |
| Variables | Mean (SD) | Mean (SD) | Mean (SD) | Mean (SD) | Mean (SD) | Mean (SD) |
| Land cover (%) | | | | | | |
| Urban | 46.7 (33.6) | 55.0 (29.9) | 43.9 (30.8) | 58.4 (28.2) | 60.2 (34.5) | 61.5 (34.3) |
| Open developed space | 23.4 (16.5) | 27.2 (17.7) | 29.6 (18.0) | 28.7 (17.3) | 15.3 (17.0) | 18.0 (17.8) |
| Grass/hay | 18.4 (29.2) | 8.8 (20.7) | 7.4 (16.5) | 5.2 (15.1) | 10.5 (18.4) | 8.8 (17.8) |
| Cropland | 8.5 (20.4) | 6.6 (18.7) | 16.7 (28.2) | 6.9 (18.7) | 4.0 (14.3) | 1.4 (7.5) |
| Forest | 1.0 (2.8) | 0.7 (2.3) | 1.4 (4.5) | 0.5 (2.3) | 5.0 (14.7) | 6.5 (17.4) |
| Hydrology (meter) | | | | | | |
| Emergent wetland | 376.8 (308.9) | 447.4 (322.5) | 476.7 (450.5) | 711.1 (578.4) | 1,039.8 (643.7) | 1,350.6 (995.2) |
| Forest and shrub wetland | 1,376.6 (779.4) | 1,447.2 (868.4) | 1,516.0 (981.2) | 1,695.5 (978.5) | 3,014.1 (2,763.0) | 2,992.0 (2,417.7) |
| Pond | 822.4 (470.2) | 851.5 (414.8) | 795.4 (513.4) | 968.6 (582.0) | 731.1 (431.0) | 865.8 (441.9) |
| Lake and river | 1,577.8 (1,328.9) | 1,695.8 (1,704.1) | 2,589.1 (3,299.7) | 1,932.0 (2,764.8) | 1,770.3 (1,011.3) | 2,006.0 (1,085.1) |
| Soil conditions (meter)† | | | | | | |
| Ponding frequency > 15% | 437.7 (363.4) | 509.1 (346.1) | 1,091.6 (1,794.2) | 1,235.9 (1,224.1) | – | – |
| Ponding frequency > 50% | 570.3 (380.8) | 587.4 (334.8) | – | – | – | – |
| Ponding frequency > 75% | 573.4 (381.3) | 594.3 (337.5) | 1,495.4 (2,082.5) | 1,578.5 (1,575.6) | 8,130.3 (4,037.2) | 12,517.4 (8,660.3) |
| Irrigation (meter) | | | | | | |
| Water draw point | 2,902.8 (2,906.7) | 2,969.6 (4,171.5) | 2,800.9 (3,425.4) | 2,248.7 (2,699.5) | 1,971.1 (1,261.1) | 5,934.7 (8,558.9) |
| DEM (meter) | – | – | – | – | | |
| Elevation | – | – | – | – | 1,017.6 (78.2) | 1,077.0 (160.6) |

*DEM = digital elevation model.
†Sioux Falls data does not include ponding frequency between 50% and 75%. Rapid City data does not include ponding frequency between 15% and 75%.

In Sioux Falls, the best model contained percent urban area and ponding frequency. Other models were relatively weak competitors, with the next best model having a difference in the Akaike's Information Criterion corrected value of 1.97. Percent urban development was included in the two best models, and ponding frequency was included in all five of the best models. Risk for WNV decreased with percentage of urban area and increased with proximity to areas with a high ponding frequency (Table 3). Overall, there was strong support for effects of urban area and proximity to soils with high ponding frequency in Sioux Falls.

In Rapid City, the best model contained elevation, percent cropland, percent forest, and ponding frequency. However, a more parsimonious model with just elevation, percent forest, and ponding frequency had only a slightly higher AICc and lower Akaike weight. All five of the top models for Rapid City contained elevation, forest, and ponding frequency. Risk for WNV increased with percent cropland, increased with percentage forest, decreased with elevation, and increased with ponding frequency. However, the 95% confidence interval for the cropland odds ratio overlapped one (Table 3). Overall, there was strong evidence for effects of forest, elevation, and ponding frequency in Rapid City and weaker evidence for an effect of cropland.

The predicted values of risk probability were calculated from the best models, and the WNV risk prediction maps are shown in Figure 1. In Sioux Falls and Aberdeen, low-risk areas typically fell inside the city limits, whereas high-risk areas were located in more rural areas. In Sioux Falls, high-risk areas were concentrated in the Prairie Coteau and the James River Lowland in the western portion of the study areas. In contrast, WNV risk was relatively low in the better-drained

TABLE 2
Model selection results based on corrected Akaike's Information Criterion*

| Study areas | Models | Parameters | Environmental variables | AICc | DAIC | Weight |
|---|---|---|---|---|---|---|
| Aberdeen | | | | | | |
| | 1 | 3 | ODE, grass/hay, emergent wetland | 660.286 | 0 | 0.22 |
| | 2 | 2 | Grass/hay, emergent wetland | 660.599 | 0.313 | 0.18 |
| | 3 | 4 | ODE, emergent wetland, ponding frequency > 75% | 662.284 | 1.998 | 0.08 |
| | 4 | 4 | ODE, grass/hay, cropland, emergent wetland | 662.284 | 1.998 | 0.08 |
| | 5 | 3 | Grass/hay, cropland, emergent wetland | 662.425 | 2.139 | 0.07 |
| Sioux Falls | | | | | | |
| | 1 | 2 | Urban, ponding frequency > 75% | 494.288 | 0 | 0.48 |
| | 2 | 3 | Urban, emergent wetland, ponding frequency > 75% | 496.261 | 1.973 | 0.18 |
| | 3 | 3 | ODE, cropland, ponding frequency > 75% | 498.389 | 4.101 | 0.06 |
| | 4 | 2 | Cropland, ponding frequency > 75% | 498.637 | 4.349 | 0.05 |
| | 5 | 4 | ODE, grass/hay cropland, ponding frequency > 75% | 499.321 | 5.033 | 0.04 |
| Rapid City | | | | | | |
| | 1 | 4 | Elevation, cropland, forest, ponding frequency > 75% | 813.441 | 0 | 0.12 |
| | 2 | 3 | Elevation, forest, ponding frequency > 75% | 814.217 | 0.776 | 0.08 |
| | 3 | 5 | Elevation, ODE, cropland, forest, ponding frequency > 75% | 814.518 | 1.077 | 0.07 |
| | 4 | 4 | Elevation, ODE, forest, ponding frequency > 75% | 814.911 | 1.470 | 0.06 |
| | 5 | 5 | Elevation, grass/hay, cropland, forest, ponding frequency > 75% | 815.095 | 1.654 | 0.05 |

*AICc = corrected Akaike's Information Criterion; DAIC = difference in Akaike's Information Criterion; ODE = open developed space.

TABLE 3
Odds ratios for environmental variables from best-fitting logistic regression models in the three study areas*

| Study area | Parameters | OR | Lower 95% CI† | Upper 95% CI |
|---|---|---|---|---|
| Aberdeen | ODE | 0.41 | 0.13 | 1.31 |
| | Grass/hay | 2.81 | 1.24 | 6.36 |
| | Emergent wetland | 1.74 | 1.12 | 2.71 |
| Sioux Falls | Urban | 0.24 | 0.12 | 0.49 |
| | Ponding frequency > 75% | 1.94 | 1.21 | 3.13 |
| Rapid City | Cropland | 3.86 | 0.79 | 18.88 |
| | Forest | 5.12 | 1.40 | 18.71 |
| | DEM (100 meters) | 0.58 | 0.45 | 0.76 |
| | Ponding frequency > 75% | 1.50 | 1.01 | 2.25 |

*OR = odds ratio; CI = confidence interval; ODE = open developed space; DEM = digital elevation model.
†Statistical significance was achieved if the 95% CI did not overlap the value 1.

soils of the Loess Prairies in the eastern portion of the Sioux Falls study area. In Aberdeen, WNV risk was highest in the grass/hay dominated Drift Plains ecoregion in the western portion of the study area and lower in the row-crop dominated Glacial Lakes Basins to the east. In Rapid City, WNV was highest in the low-elevation plains and lowest at higher elevations in the Black Hills. The area under the receiver operating characteristics curve statistic was 0.61 for Aberdeen, 0.66 for Sioux Falls, and 0.66 for Rapid City, for maps generated using the best models, demonstrating moderate predictive power.

## DISCUSSION

This study found relationships between human WNV infection and landscape-scale environmental variables in the three most populated areas in South Dakota. We focused on relatively static land cover variables, hydrological features, and soil conditions that do not change over time. Interannual variability in weather also drives temporal variability in WNV risk and can lead to hot spots of WNV risk occurring in different locations in different years. However, our analysis enabled us to highlight specific locations and environmental characteristics that have consistently higher WNV risk when examined over multiple years. These environmental variables serve as indicators of potential vector or host habitats and human activity. Irrigated agriculture has been highlighted as an important risk factor for WNV in other regions,[30] but irrigation did not emerge as an important predictor in any of our study areas. This finding likely reflects the fact that irrigated agriculture is relatively uncommon in our study areas, and corroborates our previous finding that mosquito abundance was not associated with irrigation in Sioux Falls.[31]

In Aberdeen, the associations of WNV cases with grass/hay land cover and emergent wetlands indicated the human-vector interactions tended to occur close to these landscape features. Our previous entomologic study supports the finding that the host-seeking behavior of *Cx. tarsalis* is associated with grass/hay land cover in eastern South Dakota.[31] The associations we found with emergent wetlands contradict the general assumption that *Cx. tarsalis* typically breeds in smaller pools and is not associated with larger wetlands.[32,33] However, proximity to wetlands might serve as a general indicator of landscapes with poorly drained soils that are likely to also contain smaller temporary pools. Wetlands also attract a variety of bird species and may serve

as amplification foci that increase the infection rate of mosquito populations.

Sioux Falls has a larger, more urbanized area than Aberdeen and is surrounded by a landscape that has a larger amount of cropland and a more interspersed pattern of cropland and grass/hay than the Aberdeen study area. In this landscape, we found that WNV risk was generally lower in urban areas and higher in rural areas regardless of whether they were dominated by grass/hay or croplands. The more dispersed pattern of grasslands and greater human exposure to mosquitoes in croplands likely reduced our ability to detect an association with grass/hay in this study area, despite the association of host-seeking mosquitoes with grass/hay land cover.[31] In the best model, proximity to soils with a high ponding frequency was a significant risk factor for predicting human risk (Table 2). This variable was also present in all five of the top models for Rapid City and in one of the five top models for Aberdeen, suggesting that ponding frequency from Soil Survey Geographic Database soils data may have value as a general indicator of locations that have a high suitability for *Cx. tarsalis* breeding and therefore high risk for WNV amplification and transmission. However, this hypothesis would need to be tested with additional field data on mosquito breeding sites and larval abundance.

Rapid City had more complicated relationships between WNV risk and environmental variable because of its diverse landscapes. Elevation served as an indirect gradient that captured variability in temperature (lower temperatures at higher elevations) and land cover (higher conifer forest cover at higher elevations) and was the most important predictor in this study area. Similar to Sioux Falls, cropland and soil conditions were important factor in western South Dakota. However, we did not see a major reduction of WNV risk in urban areas, possibly because Rapid City is located at the intersection of the Black Hills and Great Plains areas and landscape diversity is much higher than in eastern South Dakota. Interestingly, a positive association of WNV risk with forest was identified in Rapid City after adjusting for elevation. In the drier low-elevation landscapes within this study area, tree cover may be an indicator of more mesic habitat that provide breeding habitat for mosquitoes and also support communities of avian hosts.

After the United States was invaded by WNV in 1999, there have been many attempts to predict the spatial and temporal patterns of WNV risk. One common approach has been to use dead bird reports as indicators of human risk in the surrounding areas. For instance, previous studies have proposed that dead crow sightings and density are useful as the early warning indicators of WNV infection in New York.[34–36] Alternately, spatial aggregation among WNV human cases, WNV-positive vectors, and WNV-positive birds has been reported in Davis, California,[37] highlighting the potential or monitoring WNV risk through surveillance of mosquito populations and dead bird reports. However, because dead bird reporting systems are sensitive to human population density and public awareness of WNV, they are difficult to establish and maintain in sparsely populated areas such as the Northern Great Plains. Similarly, mosquito surveillance is expensive and infeasible across large rural areas. In areas like the northern Great Plains, understanding the landscape-level correlates of WNV offers an alternative method for identifying high-risk areas.
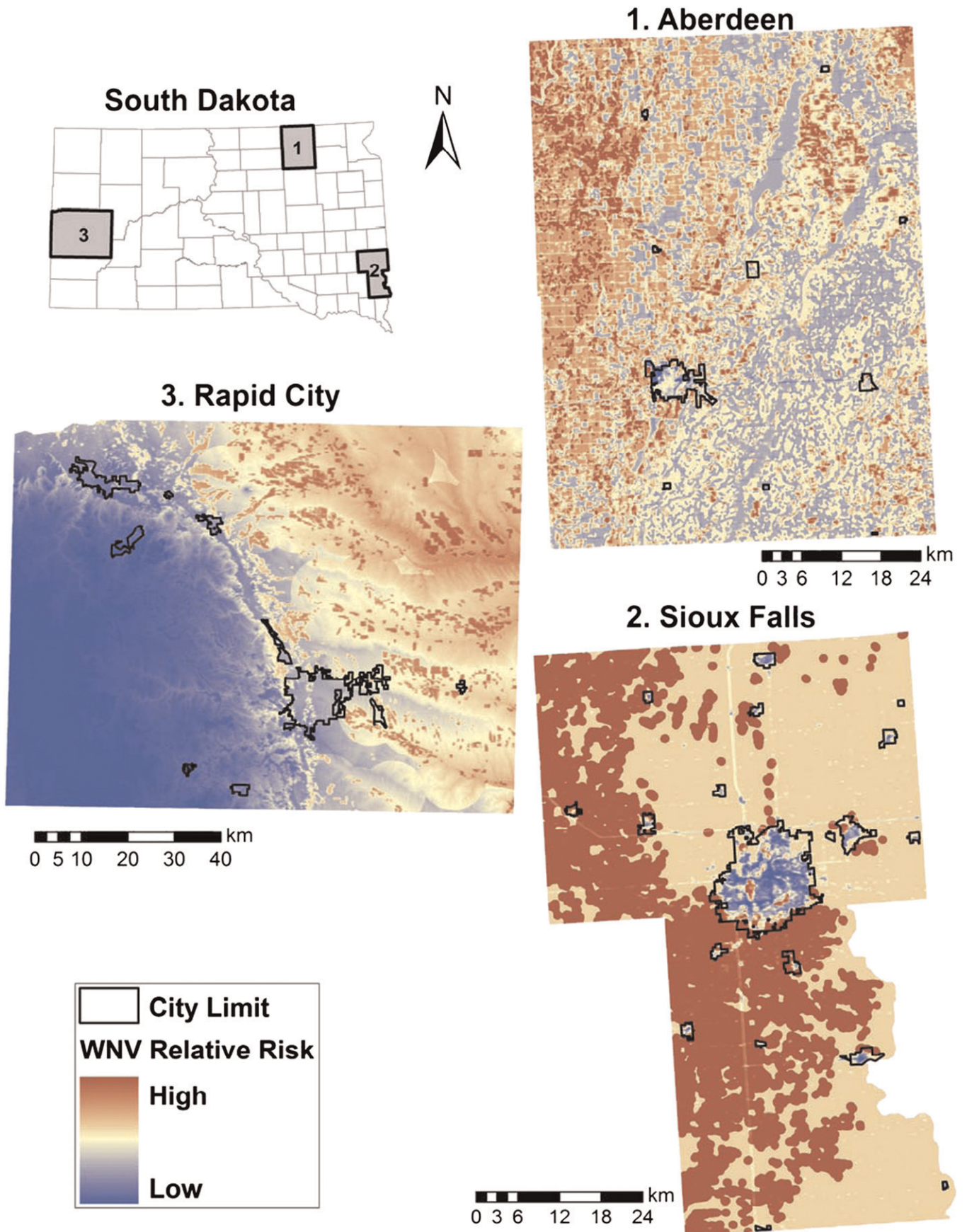
FIGURE 1. Predicted West Nile virus risk maps for the three areas in South Dakota generated from logistic regression models based on landscape-level environmental variables.

Previous studies have also demonstrated relationships between landscape-level environmental factors and WNV risk. Ruiz and others showed that the highest WNV incidence rates in Chicago and Detroit occurred in inner suburbs with intermediate vegetation and population density.[38] Infection rates for WNV in wild birds were highest in urban/suburban areas in a study Georgia.[39] Brown and others investigated WNV disease incidence in the northeastern United States and found higher risk in the urban areas with intermediate forest cover.[40] All of these studies were conducted in the eastern or midwestern United States. In contrast, our study focused on rural portions of the northern Great Plains that have distinctive ecology and encompass an important national-level hot spot for WNV.[41] Our study found associations with land cover and with soils or hydrology at all three study sites. Although we consistently found positive association with rural land cover types and indicators of poorly-drained soils (wetlands or high ponding frequency), there was no general model that was applicable across the entire region. The heterogeneity of topographic features, geographic variability in host communities, and the potential for local adaption of vectors and pathogens limit our abilities to make global inferences across broader regions. However, our findings of WNV associations with rural habitats and poorly drained soils at multiple sites suggest more general relationships that could be tested at additional locations in the northern Great Plains.

This study has several limitations that must be considered when interpreting the results. Because our spatial models demonstrated only moderate discriminatory power, they are more useful for highlighting broad trends in WNV risk than pinpointing specific locations where cases will occur. Case locations are based on geocoded residence addresses because it is impossible to identify the exact locations where persons acquired infection. Therefore, uncertainty about the location of exposure adds to the error in our models, but should not bias our inferences about environmental risk factors. Furthermore, our efforts to identify environmental correlates of WNV risk are limited by the spatial resolution, attribute resolution, and accuracy of the underlying GIS datasets. Most of our variables were derived from national-level datasets, which provide broad spatial coverage but are typically not optimal for specific, localized applications. Higher-resolution and more accurate regional datasets could enable us to develop more consistent models of WNV risk, and the present study can serve as a starting point for identifying key environmental variables and improving their spatial representation. Finally, our models did not include climate or weather variables because our aim was to evaluate relatively static environmental influences on disease transmission over multiple years. However, future modeling efforts can build on this study to integrate landscape and weather variables and develop spatial-temporal forecasts of human WNV risk to enhance disease prevention efforts and improve mosquito control programs.

Disclaimer: The content of this article are solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Allergy and Infectious Diseases or the National Institutes of Health.

Authors' addresses: Ting-Wu Chuang and Michael C. Wimberly, Geographic Information Science Center of Excellence, South Dakota State University, Brookings, SD, E-mails: chtingwu@gmail.com and michael.wimberly@sdstate.edu. Christine W. Hockett and Lon Kightlinger, South Dakota Department of Health, Pierre, SD, E-mails: Christine.Hockett@state.sd.us and Lon.Kightlinger@state.sd.us.

## REFERENCES

1. Lindsey NP, Staples JE, Lehman JA, Fischer M, 2010. Surveillance for human West Nile virus disease—United States, 1999–2008. *MMWR Surveill Summ 59:* 1–17.
2. Roehrig JT, Layton M, Smith P, Campbell GL, Nasci R, Lanciotti RS, 2002. The emergence of West Nile virus in North America: ecology, epidemiology, and surveillance. *Curr Top Microbiol Immunol 267:* 223–240.
3. Campbell GL, Marfin AA, Lanciotti RS, Gubler DJ, 2002. West Nile virus. *Lancet Infect Dis 2:* 519–529.
4. Turell MJ, Dohm DJ, Sardelis MR, Oguinn ML, Andreadis TG, Blow JA, 2005. An update on the potential of North American mosquitoes (Diptera: Culicidae) to transmit West Nile Virus. *J Med Entomol 42:* 57–62.
5. Bell JA, Mickelson NJ, Vaughan JA, 2005. West Nile virus in host-seeking mosquitoes within a residential neighborhood in Grand Forks, North Dakota. *Vector Borne Zoonotic Dis 5:* 373–382.
6. Reisen WK, Carroll BD, Takahashi R, Fang Y, Garcia S, Martinez VM, Quiring R, 2009. Repeated West Nile virus epidemic transmission in Kern County, California, 2004–2007. *J Med Entomol 46:* 139–157.
7. Reisen WK, Cayan D, Tyree M, Barker CM, Eldridge B, Dettinger M, 2008. Impact of climate variation on mosquito abundance in California. *J Vector Ecol 33:* 89–98.
8. Winters AM, Eisen RJ, Lozano-Fuentes S, Moore CG, Pape WJ, Eisen L, 2008. Predictive spatial models for risk of West Nile virus exposure in eastern and western Colorado. *Am J Trop Med Hyg 79:* 581–590.
9. Shaman J, Day JF, Komar N, 2010. Hydrologic conditions describe West Nile virus risk in Colorado. *Int J Environ Res Public Health 7:* 494–508.
10. Day JF, Shaman J, 2008. Using hydrologic conditions to forecast the risk of focal and epidemic arboviral transmission in peninsular Florida. *J Med Entomol 45:* 458–465.
11. Eisen L, Barker CM, Moore CG, Pape WJ, Winters AM, Cheronis N, 2010. Irrigated agriculture is an important risk factor for West Nile virus disease in the hyperendemic Larimer-Boulder-Weld area of north central Colorado. *J Med Entomol 47:* 939–951.
12. Bolling BG, Barker CM, Moore CG, Pape WJ, Eisen L, 2009. Seasonal patterns for entomological measures of risk for exposure to *Culex* vectors and West Nile virus in relation to human disease cases in northeastern Colorado. *J Med Entomol 46:* 1519–1531.
13. Kwan JL, Kluh S, Madon MB, Reisen WK, 2010. West Nile virus emergence and persistence in Los Angeles, California, 2003–2008. *Am J Trop Med Hyg 83:* 400–412.
14. Nielsen CF, Armijos MV, Wheeler S, Carpenter TE, Boyce WM, Kelley K, Brown D, Scott TW, Reisen WK, 2008. Risk factors associated with human infection during the 2006 West Nile virus outbreak in Davis, a residential community in northern California. *Am J Trop Med Hyg 78:* 53–62.
15. Reisen WK, Takahashi RM, Carroll BD, Quiring R, 2008. Delinquent mortgages, neglected swimming pools, and West Nile virus, California. *Emerg Infect Dis 14:* 1747–1749.
16. Ruiz MO, Tedesco C, McTighe TJ, Austin C, Kitron U, 2004. Environmental and social determinants of human risk during a West Nile virus outbreak in the greater Chicago area, 2002. *Int J Health Geogr 3:* 8.
17. Harrigan RJ, Thomassen HA, Buermann W, Cummings RF, Kahn ME, Smith TB, 2010. Economic conditions predict prevalence of West Nile virus. *PLoS ONE 5:* e15437.

18. Wimberly MC, Hildreth MB, Boyte SP, Lindquist E, Kightlinger L, 2008. Ecological niche of the 2003 West Nile virus epidemic in the northern great plains of the United States. *PLoS ONE 3:* e3744.

19. Bowden SE, Magori K, Drake JM, 2011. Regional differences in the association between land cover and West Nile virus disease incidence in humans in the United States. *Am J Trop Med Hyg 84:* 234–238.

20. Wey CL, Griesse J, Kightlinger L, Wimberly MC, 2009. Geographic variability in geocoding success for West Nile virus cases in South Dakota. *Health Place 15:* 1108–1114.

21. Eisen RJ, Enscore RE, Biggerstaff BJ, Reynolds PJ, Ettestad P, Brown T, Pape J, Tanda D, Levy CE, Engelthaler DM, Cheek J, Bueno R Jr, Targhetta J, Montenieri JA, Gage KL, 2007. Human plague in the southwestern United States, 1957–2004: spatial models of elevated risk of human exposure to *Yersinia pestis. J Med Entomol 44:* 530–537.

22. Theobald DM, 2001. Land-use dynamics beyond the American urban fringes. *Geogr Rev 91:* 544–564.

23. Tiner RW, Jr., 1984. *Wetlands of the United States: Current Status and Recent Trends*. Washington, DC: U.S. Fish and Wildlife Service, 14–16.

24. Ezenwa VO, Milheim LE, Coffey MF, Godsey MS, King RJ, Guptill SC, 2007. Land cover variation and West Nile virus prevalence: patterns, processes, and implications for disease control. *Vector Borne Zoonotic Dis 7:* 173–180.

25. Hubalek Z, 2008. Mosquito-borne viruses in Europe. *Parasitol Res 103 (Suppl 1):* S29–S43.

26. U.S. Fish and Wildlife Service, 2010. *Wetlands Mapper Documentation and Instructions Manual*. Madison, WI: Division of Habitat and Resource Conservation, Branch of Resource and Mapping Support.

27. Patz JA, Strzepek K, Lele S, Hedden M, Greene S, Noden B, Hay SI, Kalkstein L, Beier JC, 1998. Predicting key malaria transmission factors, biting and entomological inoculation rates, using modelled soil moisture in Kenya. *Trop Med Int Health 3:* 818–827.

28. Shaman J, Stieglitz M, Stark C, Le Blancq S, Cane M, 2002. Using a dynamic hydrology model to predict mosquito abundances in flood and swamp water. *Emerg Infect Dis 8:* 6–13.

29. Burnham KP, Anderson DR, 2002. *Model Selection and Multimodel Inference: A Practical Information Theoretic Approach*. New York: Springer-Verlag.

30. Gates MC, Boston RC, 2009. Irrigation linked to a greater incidence of human and veterinary West Nile virus cases in the United States from 2004 to 2006. *Prev Vet Med 89:* 134–137.

31. Chuang TW, Hildreth MB, Vanroekel DL, Wimberly MC, 2011. Weather and land cover influences on mosquito populations in Sioux Falls, South Dakota. *J Med Entomol 48:* 669–679.

32. Workman PD, Walton WE, 2000. Emergence patterns of *Culex* mosquitoes at an experimental constructed treatment wetland in southern California. *J Am Mosq Control Assoc 16:* 124–130.

33. Mercer DR, Sheeley SL, Brown EJ, 2005. Mosquito (Diptera: Culicidae) development within microhabitats of an Iowa wetland. *J Med Entomol 42:* 685–693.

34. Johnson GD, Eidson M, Schmit K, Ellis A, Kulldorff M, 2006. Geographic prediction of human onset of West Nile virus using dead crow clusters: an evaluation of year 2002 data in New York State. *Am J Epidemiol 163:* 171–180.

35. Eidson M, Miller J, Kramer L, Cherry B, Hagiwara Y, 2001. Dead crow densities and human cases of West Nile virus, New York State, 2000. *Emerg Infect Dis 7:* 662–664.

36. Eidson M, Kramer L, Stone W, Hagiwara Y, Schmit K, 2001. Dead bird surveillance as an early warning system for West Nile virus. *Emerg Infect Dis 7:* 631–635.

37. Nielsen CF, Reisen WK, 2007. West Nile virus-infected dead corvids increase the risk of infection in *Culex* mosquitoes (Diptera: Culicidae) in domestic landscapes. *J Med Entomol 44:* 1067–1073.

38. Ruiz MO, Walker ED, Foster ES, Haramis LD, Kitron UD, 2007. Association of West Nile virus illness and urban landscapes in Chicago and Detroit. *Int J Health Geogr 6:* 10.

39. Gibbs SE, Wimberly MC, Madden M, Masour J, Yabsley MJ, Stallknecht DE, 2006. Factors affecting the geographic distribution of West Nile virus in Georgia, USA: 2002–2004. *Vector Borne Zoonotic Dis 6:* 73–82.

40. Brown HE, Childs JE, Diuk-Wasser MA, Fish D, 2008. Ecological factors associated with West Nile virus transmission, northeastern United States. *Emerg Infect Dis 14:* 1539–1545.

41. Sugumaran R, Larson SR, Degroote JP, 2009. Spatio-temporal cluster analysis of county-based human West Nile virus incidence in the continental United States. *Int J Health Geogr 8:* 43.