07JEI00083
1726-2135/1684-8799

www.iseis.org/jei

Journal of Environmental Informatics 9(1) 4-17 (2007)

Spatial Patterns of Lightning at Different Spatial Scales in the Western United States during August of 1990 – A Case Study Using Geographic Information Systems Technology

Z. Y. Yin[1*], J. Estberg[2], E. J. Hallisey[3] and D. R. Cayan[4]

[1]Marine Science and Environmental Studies, University of San Diego, San Diego, CA 92110, USA
[2]Department of Physics, University of San Diego, San Diego, CA 92110, USA
[3]Division of Public Health, Georgia Department of Human Resources, Atlanta, GA 30303, USA
[4]Scripps Institution of Oceanography, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA

ABSTRACT. Geographic information systems (GIS) have been widely used to study spatial variability in different atmospheric processes. In this study, we used a GIS approach to explore the potential to examine variation patterns of lightning strikes at different scales so that micro-, synoptic-, and planetary-scale processes can be linked in explaining and modeling the distribution patterns of lightning strikes. The data collected by the ground-based lightning detection system for an entire month in the western United States were used as an example. Lightning strike density surfaces were generated using different kernel bandwidths, or search radii. It has been recognized that density surfaces are useful in visual interpretation of spatial patterns at different scales, but there are insufficient data on how well such surfaces can be used in quantitative analysis of point distribution patterns. In our study, the resulting surfaces were compared quantitatively with gridded lightning strikes using meshes, or fishnets, of different cell sizes. The fishnet cell sizes ranged from 1 km for micro-scale processes to 50 km for synoptic- and planetary-scale processes. The results suggest that there is a threshold in the search radius or kernel bandwidth, above which a significant amount of errors would be introduced in quantitative analysis. It could be argued that it is possible to achieve a balance between the need for visual interpretation of distribution patterns and the need for quantitative analysis at different scales. We used the lightning data of August 1990 and digital elevation models of 1 km resolution to perform a case study on the relationship between lightning occurrence and topography. Our results indicate that at different spatial scales, the relationships between lightning density and topography may reflect different processes that influenced the spatial distribution pattern of the lightning occurrence.

Keywords: Lightning density, kernel bandwidth, geographic information systems (GIS)

1. Introduction

This study examines the potential of studying spatial patterns of cloud-to-ground lightning events using density surfaces generated by geographic information systems (GIS). Cloud-to-ground lightning is a frequent natural phenomenon often associated with severe weather conditions, which can cause loss of property and human life. Therefore, it is considered a form of natural hazards requiring risk analysis from both spatial and temporal perspectives (Elsom, 1996, 2001; Coates et al., 1993; Blong, 1997; Berz et al., 2001). From a regional point of view, it is a concern for the utility industry because severe thunderstorms can cause equipment damage and power shortages for large areas. Spatial analysis of lightning occurrence has been useful to assess the safety and reliability of power utility systems (López et al., 1997). It can also pose a severe risk for major public events. For example, Livingston et al. (1996) and Watson and Holle (1996) examined the climatology of lightning for the 1996 Summer Olympic Games in Atlanta, Georgia, a region with frequent thunderstorms during the summer months. Lightning is associated with the occurrence of wildfires as a major ignition mechanism, which again has significant impact on human society and ecosystems (Stocks et al., 2003; Gedalof et al., 2005). The spatial and temporal patterns of wildfires can be linked to the occurrence of lightning events (Hardy, 2005). To understand the occurrence of wildfire, lightning and the associated weather conditions must be examined. Synoptic conditions optimal of thunderstorm occurrence can then be used for predictions of fire occurrences in a large area context (Rorig and Ferguson, 1999, 2002).

Because of its associations with many different atmospheric and surface processes, the spatial pattern of lightning occurrence is of special interest in atmospheric science, fire ecology, disaster planning and management, and the power/energy industry. Under a specific synoptic condition, lightning occurrence associated with one or more thunderstorms may display distinct spatial patterns (Figure 1). Besides numerous studies on the physics of the lightning phenomenon, quantitative information on lightning has been used to estimate convective rainfall at different spatial scales, especially during the warm season when convective rainfall makes up a large

* Corresponding author: zyin@sandiego.edu

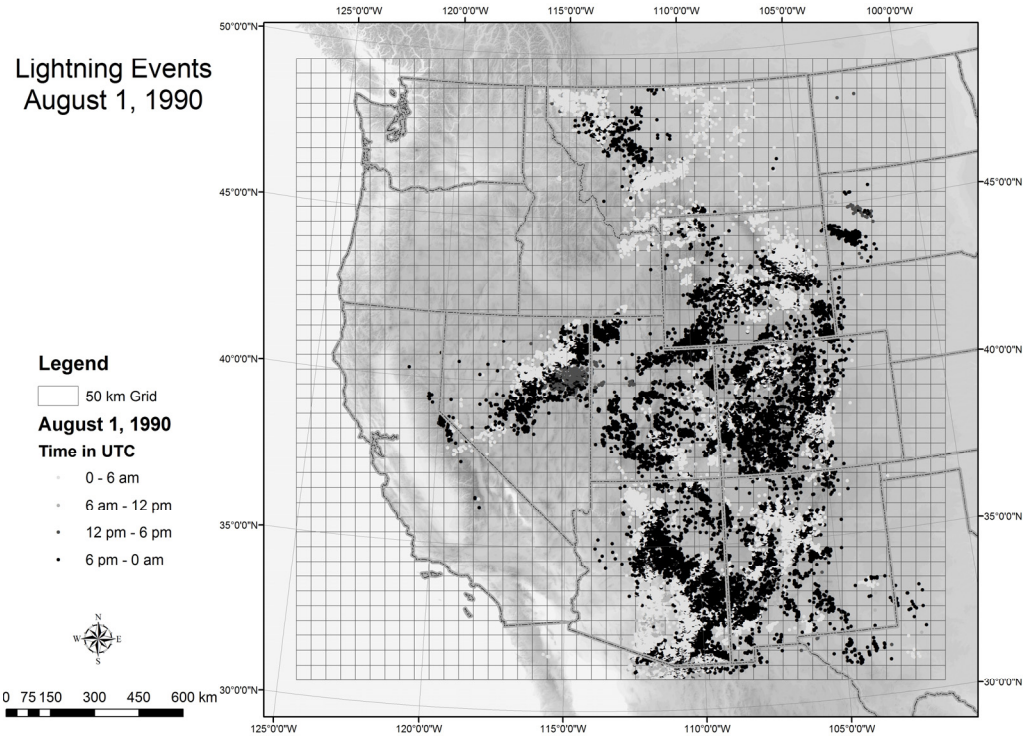

Figure 1. Spatial pattern of lightning occurrence associated with a given synoptic condition.

portion of the total rainfall events (Sheridan et al., 1997; Petersen and Rutledge, 1998; Tapia et al., 1999; Seity et al., 2001; Ezcurra et al., 2003). Spatial patterns of lightning were analyzed together with precipitation data for specific events, such as the 1993 flood in the U.S. Midwest (Kempf and Krider, 2003) and post-wildfire debris flows in Colorado (Underwood and Schultz, 2003, 2004). Lightning data may also be useful in the construction of global gridded precipitation datasets (Morales and Anagnostou, 2003; Chronis et al., 2004). In many studies with a spatial context, the lightning events were often overlaid with other layers of information. In predicting lightning occurrence in space and time, Diaz-Avalos et al. (2001) considered factors such as vegetation cover, elevation, and slope in the Blue Mountains area of Oregon. In a study covering different spatial scales, Dissing and Verbela (2003) concluded that various combinations of elevation and areal coverage of forest variables can be used to explain the variation in lightning strikes. In studies of wildfires, vegetation, topography and weather conditions have been used to predict fire behavior (Fowler and Asleson, 1984). On the other hand, the occurrence of lightning has been used to predict spatial pattern of wildfire occurrence and the associated vegetation changes from a long-term perspective (Griffin et al., 1983; Bergeron et al., 1997; Potter et al., 1998; Dissing and Verbyla, 2003). Another example of relating lightning occurrence to earth surface processes is the investigation of the possible enhancement effect of urbanization on lightning frequency, in which lightning frequency and density data must be overlaid accurately on the land use/land cover data in order to determine the spatial associations (Westcott, 1995; Steiger et al., 2002; Soriano and de Pablo, 2002; Naccarato et al., 2003; Pinto et al., 2004). Similarly, lightning occurrence downwind of forest fires has been investigated for possible enhancement by the smoke plumes (Smith et al., 2003), which again requires geographic coregistration of lightning and fire occurrence.

The purpose of this study is to investigate the potential of using density surfaces generated by GIS as a tool for both visual and quantitative analyses of the spatial pattern of lightning strikes. A density surface is a raster data format, with each cell containing a value to represent the density of point occurrence (Bailey and Gatrell, 1995), such as lightning strikes. The density is determined by counting the number of point occurrence within a given search radius or bandwidth. For regions with a high frequency of lightning strikes, density surfaces are especially useful in visually characterizing long-term patterns of lightning occurrence because a map of all lightning events as points during an extended period, such as a few weeks, would simply result in a solid mass without any discernable patterns (Figure 2). Another advantage of density surfaces lies in the nature of the raster data model, which has a simple data structure and is easy to use in biophysical modeling of spatial processes in which overlay of different data layers is often necessary (Lo and Yeung, 2002). One question we attempt to answer is whether density surfaces generated using different levels of smoothing as represented by the

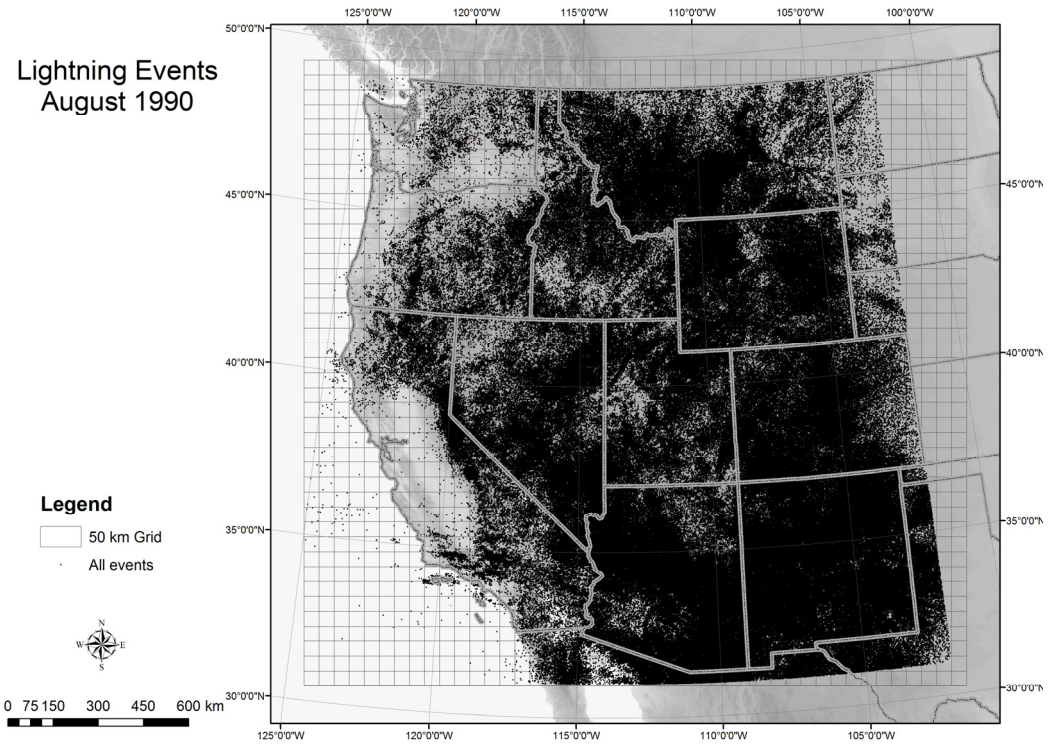

Figure 2. The study area – western United States, and all lightning events during August 1990.

search radius or bandwidth could satisfy the accuracy requirements of studies of lightning at different spatial scales. We will examine current practices of using GIS in lightning analysis and investigate the influence of various parameters in constructing the density surfaces, especially the influence of search radius or bandwidth, using the lightning strikes during August 1990 across the entire western United States as a case study.

2. Data and Methods

The lightning data of August 1990 used in this study were obtained by the National Lightning Detection Network (NLDN), a ground-based detection system established during the 1980s. There are over 100 sensors across the conterminous United States with an accuracy of 0.5 ~ 1 km (Cummins et al., 1998; Underwood and Shultz, 2004). This network mainly monitors cloud-to-ground lightning activities and is currently managed by Vaisala located in Tucson, Arizona (www.vaisala.com). The dataset contains the following fields: latitude and longitude of lightning events, date and time, polarity, frequency when there were multiple strikes at a given location, and amplitude. Since the main focus of this study is the lightning occurrence frequency, each event was treated as a single point without considering the multiplicity. The lightning data were not screened for errors or to separate positive and negative polarities since the main purpose of this study is to explore the methods of analysis. For the month of August 1990, there were a total of 666,811 lightning events in the western United States in the spatial realm of 31 ~ 45°N and 103 ~ 125°W.

Lightning studies have employed several ways to summarize lightning density over space. One approach of integrating other spatial data with lightning data is by boundaries of spatial units such as counties (Stallins, 2004) or watersheds. The results are step-like statistical surfaces similar to choropleth maps (Dent, 1999), which may not offer sufficient spatial resolution for modeling purposes as other spatial data layer may be presented as continuous fields. Another commonly used method is to use grids of a fixed size, such as 5 km (Steiger et al., 2002), and present the number of lightning strikes in each cell, essentially constructing a 2-dimension histogram (Silverman, 1986). In an investigation of the relationship between lightning occurrence and cloud-top temperature, Molinie and Jacobson (2004) used variable grid sizes from 10 × 10 km to 20 × 20 km across the conterminous United States corresponding to the footprints of the GOES-8 sensor. An alternative, but similar method is to use the grid of latitude and longitude to create grid cells (Pinto et al., 2004), although it may not be appropriate for large areas because the physical dimension of the latitude/longitude grid is not constant at different latitudes. Both approaches result in a surface presentation of the lightning density. These surfaces, however, may not be smooth (unless the grid size is small relative to the spatial extent of the study area) and there could be abrupt changes from one cell to the next depending on where the cell boundary is located.

A density surface, on the other hand, is a raster presenta-

tion of point density distribution as a continuous smooth surface over space. Density surfaces are produced by counting the frequency of points within a moving disc of a given search radius. There are two approaches to presenting the density of points inside the disc. One is a simple summation of the number of points divided by the search area, which is assigned to the cell at the center of the search area (*Help Manual of ArcGIS* under *Point Density,* ESRI, Redlands, CA). Another approach, kernel bandwidth, uses a bivariate probability density function to attribute the density values inside the search area for a given location, with the highest value centered at the event and reaching zero at the margin of the search area (Silverman, 1986; Bailey and Gatrell, 1995). The intensity of a spatial process that generates the point distribution ($s_1 \ldots s_n$) at any location s can be estimated as:

$$\hat{\lambda}_\tau(s) = \frac{1}{\delta_\tau(s)} \sum_{i=1}^{n} \frac{1}{\tau^2} K[\frac{(s - s_i)}{\tau}] \quad (1)$$

where $\delta_\tau(s)$ is an edge correction factor to ensure that all computations are limited within the study region, and τ is the bandwidth or search radius that defines the disc of search area surrounding the location s (Bailey and Gatrell, 1995). K is the bivariate probability density function, also known as kernel function, with a radially symmetric unimodal shape. The variate of K is the distance between the location s and the point event s_i, normalized by the bandwidth τ. If the quartic kernel function as described by Silverman (1986, p. 76, Eq. 4.5) is used without considering the edge effect, the above equation can be approximated as:

$$\hat{\lambda}_\tau(s) = \sum_{h_i \le \tau} \frac{3}{\pi\tau^2} (1 - \frac{h_i^2}{\tau^2}) \quad (2)$$

where h_i is the distance between a location s and an observed point event s_i, which has an influence to the density value at point s within the bandwidth of τ (Bailey and Gatrell, 1995, p. 85). Figure 3 illustrates the profile of the quartic density kernel function for half of the search area based on a 5 km bandwidth. At any given cell location of the output raster, all overlapping density values are added up to form a continuous density surface. Generally speaking, the greater the bandwidth, the flatter and smoother the resulting density surface, but the more likely that the local features may be obscured (Bailey and Gatrell, 1995).

Although the density surface approach has advantages over the other methods, except for a rough estimate or rule-of-thumb based on the average point density, there are no clear guidelines on how to determine a proper bandwidth. For example, in ArcGIS, a leading GIS software package (ESRI, Redlands, CA), the default search radius or bandwidth is set arbitrarily as the smaller of the width and height of the point data extent divided by 30 (*Help Manual of ArcGIS* under *Point Density*). A number of estimates of the bandwidth exist based on the mean density of points in area of A (N/A) or mean distance between points as $A^{0.5}/N$, where N is the total number of point events. Bailey and Gatrell (1995) outlined an estimate of suitable bandwidth as $0.68\ N^{-0.2}\ A^{0.5}$. This implies that for lightning studies in a given region, different bandwidths or search radii might be necessary to achieve the best results as the number of lightning strikes fluctuates. The issue may be more relevant in GIS analysis since the operator may zoom in and out to various spatial extents to examine features and processes of different scales. In this case, variable bandwidths may not be feasible as the number of total point events changes, while a single fixed bandwidth based on the total number of points in the entire dataset may not work well at all scales. Williamson et al. (1999) proposed a method based on the spacing between a given number of nearest neighboring points. They claimed that their method was superior to the others. However, the user was still left with the need to determine a factor (k) for the smoothing effect.

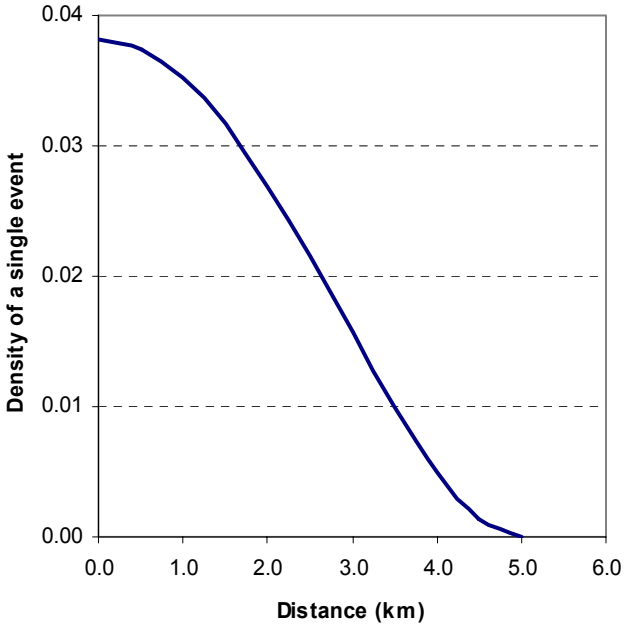

Figure 3. A profile of the quartic kernel density function (Silverman, 1986, p.76 Eq. 4.5; Bailey and Gatrell, 1995, p. 85) of half of the search area for a 5 km bandwidth.

We used ArcGIS to analyze all lightning events in August of 1990, focusing on two approaches to generating lightning densities. For the first approach, we used fixed grid cell sizes of 1, 5, 10, 20, 40, and 50 km and counted lightning events within each cell, as described by Steiger et al. (2002). A Visual BASIC script (Nicholas, 2003) was used to generate "fishnet" grid layers of 1 km, 5 km, 10 km, 20 km, 40 km, and 50 km in size (Figure 4). These grid sizes can be related to meteorological conditions of various scales: 1 ~ 5 km to represent micro-scale distribution patterns, such as those influenced by local topography, 10 ~ 20 km grids to represent meso-scale or regional patterns, and grids of 40 ~ 50 km to repre-

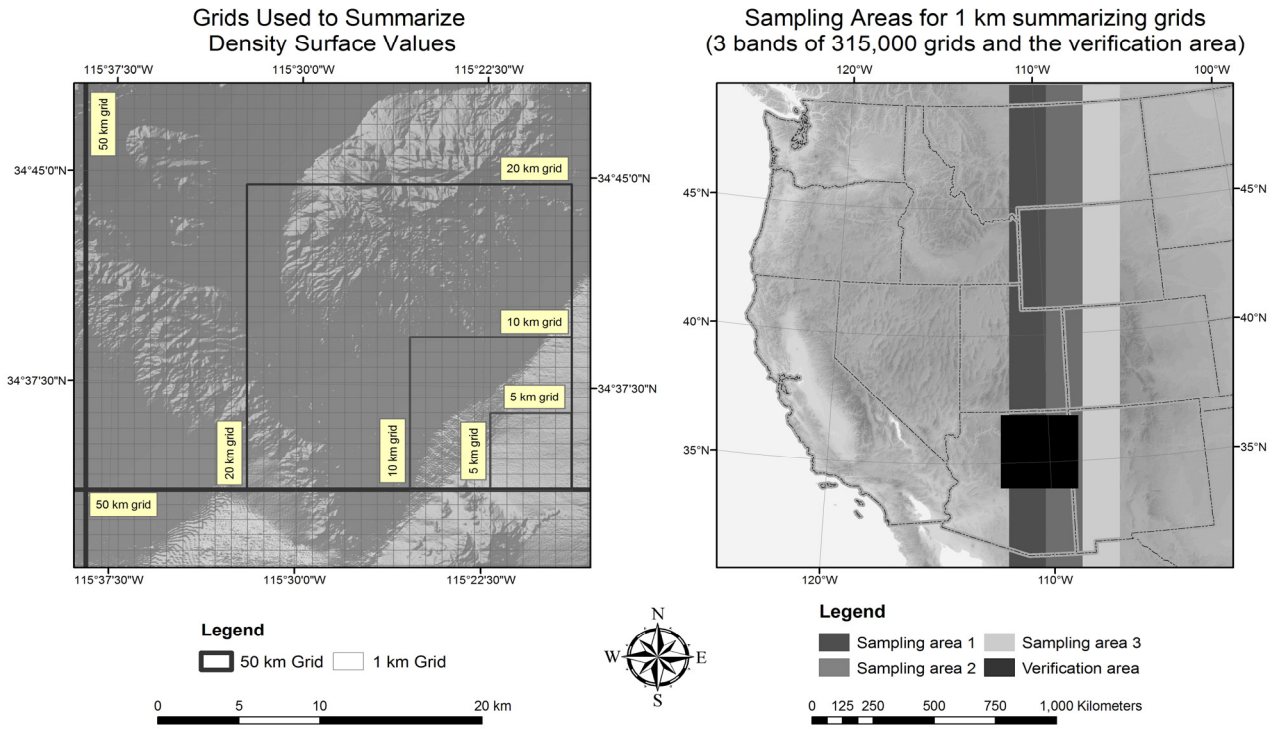

Figure 4. Fishnet grids of 1 km to 50 km in size used as the summarizing grids (a), three bands of 1-km grids of 315,000 cells used in analyzing the performance of density surfaces (b), and an area of 99,225 km^2 in northeastern Arizona for verification using the Zonal Statistics function of ArcGIS.

sent synoptic and planetary patterns, such as those related to upper-air ridge and trough locations and general circulation patterns. The fishnet grid layers output from the script are comprised of polygons, which were used to count the lightning events inside each grid cell using the spatial join function of the GIS software.

For the second approach, kernel bandwidths of 1, 5, 10, 20 and 50 km were used to generate raster lightning density surfaces, while maintaining a constant output cell size of 1 km. In producing raster point density surfaces, the kernel bandwidth option was chosen over the simple search area option because of the smoothed visual effect (Figure 5). Additionally, the underlying assumption of the kernel bandwidth estimation is that the event occurrence is probabilistic, with the recorded location having the highest probability (Silverman, 1986), which fits well to the characteristics of lightning occurrence and the accuracy of the lightning data (0.5 ~ 1 km) (Cummins et al., 1989; Underwood and Schultz, 2004). The same set of fishnet grid polygons were then used to summarize raster values of the five density surfaces using the Zonal Statistics function. The density results from both approaches were compared using regression analysis, with the estimated number of events from the raster density surfaces regressed against the actual number of events counted by the fishnet grids.

For analysis based on 1 km grids, the Zonal Statistics function could not be directly employed because we encountered a limit to the maximum number of the vectors zones allowed. Therefore, three fishnet polygon layers, each with 315,000 1-km cells running N-S across and study region, were generated for the areas with high lightning occurrence (Figure 4). The raster density surfaces were first converted to a point data layer using a Visual BASIC script (Rathert, 2005), with the points located at the center of the raster cells assigned with the cell density values. Then the point layer was spatially joined to the fishnet grid polygons to transfer the raster cell values to the fishnet polygon layers, which were also used to count the number of lightning events from the original lightning point data. Additionally, an area of 99,225 km^2 was selected in a region of relatively high lightning spatial variability (north-eastern Arizona) to be used for verification (Figure 4) using the Zonal Statistics function.

In regression analysis for the comparison between the density values obtained from the raster density surfaces and those from the fishnet grids of different sizes, we used the ratio of the search radius or kernel bandwidth (KB), to the summarizing grid size as the criterion to standardize the effect of grid size. For example, a value of 1.0 means the search radius is the same as the summarizing grid size ranging from 1 km to 50 km when the lightning events were counted. A ratio smaller than 1.0 represents the bandwidth being smaller than the size of the summarizing grid cells, and vice versa. The coefficient of determination (R^2) and the standard error of estima-

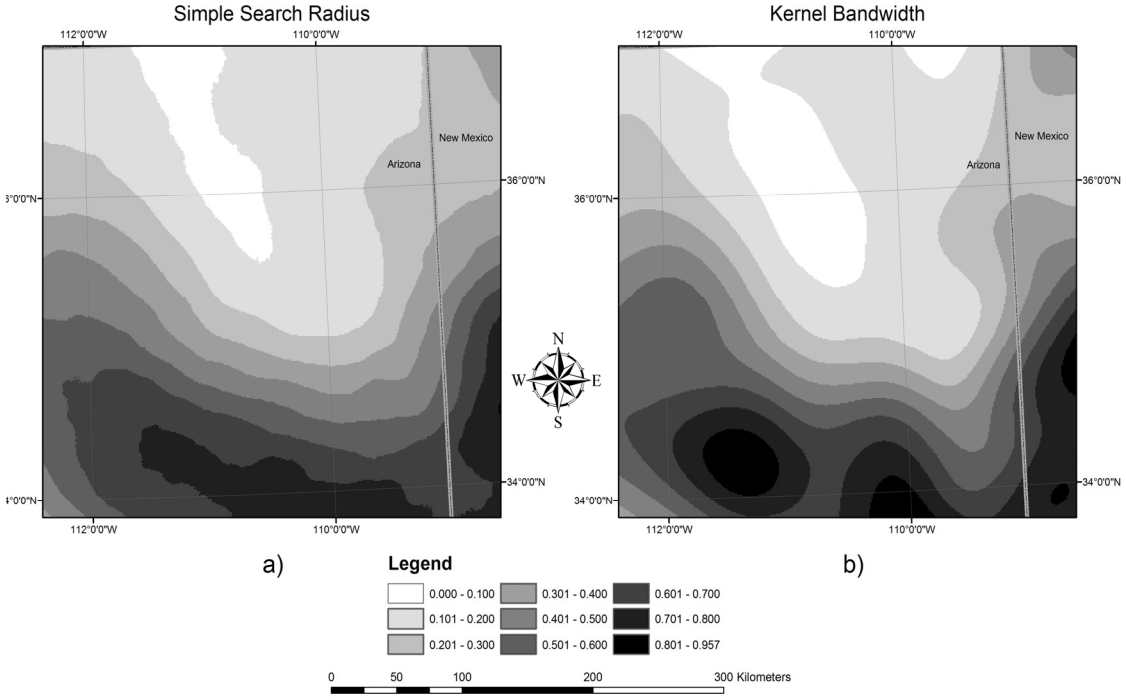

Figure 5. Visual comparison between the simple search radius (a) and kernel estimation (b) approaches in generating density surfaces. Both were based on the ArcGIS default search radius of 68.2 km and generated for the entire study region, but zoomed to the verification area in northeastern Arizona.

tion of regression were used as the measures of performance of the raster density surfaces to predict lightning distribution patterns.

3. Results and Discussion

3.1. Visual Evaluation

The lightning strike data were projected to an Albers equal area projection with parameters listed in Table 1. The resulting spatial extent was measured as approximately 2150 km in width (X) and 2070 km in height (Y). Table 2 contains general statistics of the lightning strikes during August of 1990. The mean density was calculated by dividing the total estimated area (A) using the spatial extent height and width into the total number of lightning events (N). The mean distance between lightning events was determined as $A^{0.5}/N$. Figure 6, zooming into the verification area in northeastern Arizona for a visual comparison, shows the raster density surfaces using the default bandwidth of 68.4 km in ArcGIS, and that based on the Bailey and Gratrell (1995) bandwidth estimate of 98.2 km. Both raster surfaces were generated for the entire study region with an output cell size of 1 km. There is no doubt that the evaluation of visual presentation can be subjective. However, one can make a general assessment whether a map, based on an appropriate bandwidth at a specific bandwidth, renders good visual results. Both maps gave reasonably good presentations of regional patterns of lightning occurrence at the synoptic scale, with very smoothed density surfaces (Figure 6). When compared with the original point pattern in Figure 2, the density surfaces clearly demonstrated the advantage of displaying the spatial pattern of lightning occurrence during the month. The smaller bandwidth (ArcGIS default) rendered slightly more details than the map based on the greater bandwidth.

Table 1. Parameters of the Albers Equal Area Projection Used in this Study

	Projection Parameters
Projection Name	Albers Equal Area Conic
Central Meridian	114°W
Reference Latitude	20°N
1st Parallel	35°N
2nd Parallel	45°N
False Easting	0.00 meter
False Northing	0.00 meter

The following examples at the local/micro-scale and regional scales (Figure 7) illustrate the influence of bandwidths on the patterns displayed on density surfaces, as bandwidth increased from 1 km to 20 km. Again, all density surfaces were generated using 1-km cell size for the purpose of comparison in the following section. The maps focus on an area with relatively high lightning density in northeastern Arizona. It can be seen that the density surfaces became increasingly

smoothed as the bandwidth increased and, concurrently, local details were diminished in the process.

Table 2. General Characteristics of Lightning Strikes in the Western United States in August 1990

	General Statistics
Total Lightning Events	666,811
Spatial Extent Height	2070 km
Spatial Extent Width	2150 km
Total Area	4,450,555 km^2
Mean Lightning Density	0.1498/km^2
Mean Distance	2.58 km
ArcGIS Default Bandwidth	68.4 km
Bailey & Gatrell Bandwidth	98.2 km

3.2. Accuracy Assessment Using Regression Analysis

In regression analysis to assess the performance of the density surfaces, the dependent variable is the lightning density values obtained by counting the number of lighting events in grids of different sizes, while the independent variable is the lightning density obtained by summarizing the density surfaces using the same set of summarizing grids. The R^2 value is a measure of the goodness-of-fit of the point pattern to the regression line or the prediction power of the regression equation. In this case, the R^2 value represents the quality of prediction of the actual occurrence of lightning events by a raster density surface within a specific grid system. For each summarizing fishnet grid size 5 km or larger (5, 10, 20, 40, and 50 km), a regression model was produced for a given kernel bandwidth (KB). For example, when 10 km fishnet grids were used, the R^2 values of the regression models decreased from 0.999 to 0.993, 0.980, 0.932, and 0.832 for bandwidths of 1 km, 5 km, 10 km, 20 km, and 50 km, respectively (Table 3). For summarizing fishnet grids of 1 km in size, four different regression models were produced using results from three bands of 315,000 fishnet grids and the verification area of 99,225 km^2 in northeastern Arizona. The other metric used in evaluation is the standardized errors of estimates of regression analysis, further normalized by the mean density and half of the data range for cross-bandwidth comparisons.

When the R^2 values and the normalized standard errors are plotted against the ratio of bandwidth/grid size (Figure 8), it is clear that the density surfaces rendered accurate results as long as the bandwidth is smaller than or equal to the grid size used to summarize the surface values. When the bandwidth is greater than the grid size, the density surfaces become less and less effective as the bandwidth to grid size ratio increases. For summarizing grids of 5 km or greater, R^2 values dropped to below 0.9 as the bandwidth/grid size ratio increased to 2.0 or higher. Similarly, the standard errors/mean ratio reached 0.5, while the standard errors normalized by half of the data range reached approximately 0.04. When the 1 km fishnet grids were used to summarize the density surfaces, however,

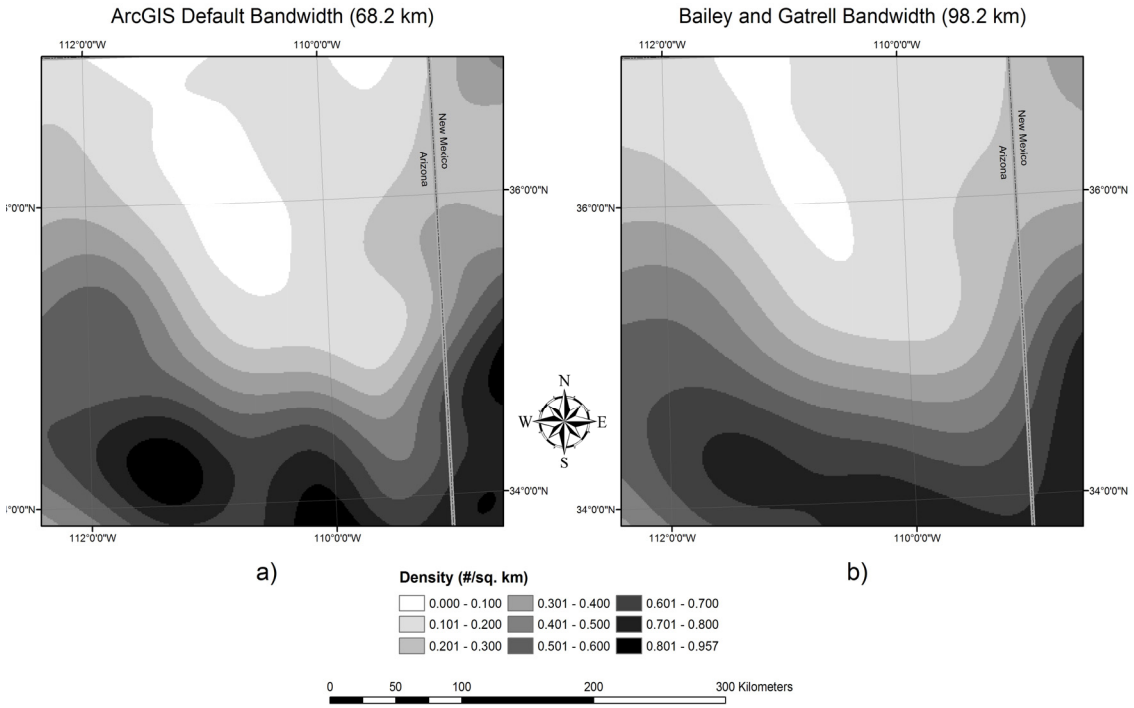

Figure 6. Visual comparison between density surfaces generated using the ArcGIS default kernel bandwidth (68.2 km) (a) and bandwidth estimated using the Bailey and Gatrell (1995) method (b). Both were generated for the entire study region, but zoomed to the verification area in northeastern Arizona.

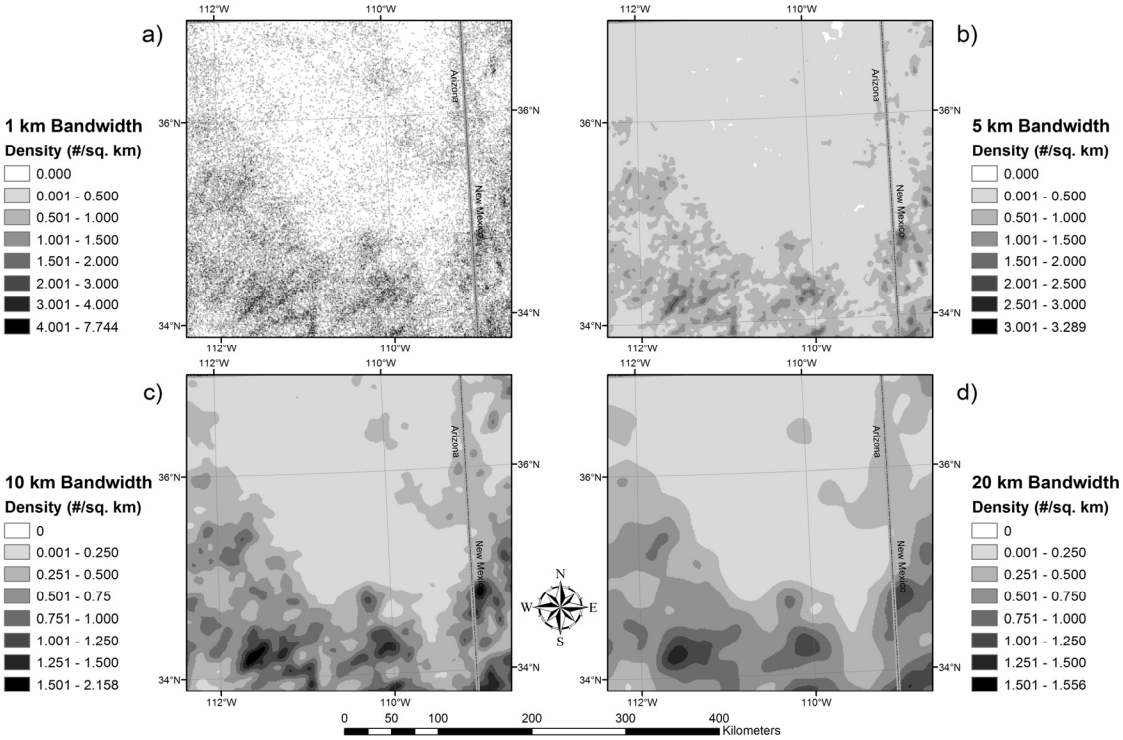

Figure 7. Density surfaces generated using bandwidths of 1 km (a), 5 km (b), 10 km (c), and 20 km (d) for micro and regional scales, focused on the verification area in northeastern Arizona. The density surface of 50 km is not presented as it is similar to that using the bandwidth of 68.2 km.

the regression analysis produced dramatically worse performance for all density surfaces of bandwidths greater than 1 km, although the results of the 1-km summarizing fishnet grids were very consistent in the three different sampling bands. This was further confirmed by the results for the 99,225 km^2 verification area.

Since the results of the 1-km summarizing grids differ so much from the grids of 5 km or greater, we analyzed the relationships between the performance metrics and the bandwidth/grid size ratio separately. For plotting purposes, $1/R^2$ was used for graphing instead of the original R^2 values. The best curve-fitting function (linear, logarithmic, power, or exponential) was determined for each metric (Figure 9, Table 4). It should be pointed out that due to small sample sizes; the confidence regions of the estimated parameters were not calculated. Nevertheless, based on these equations (Table 4), a bandwidth can be estimated for a given summarizing grid size and a desirable performance metric, or vice versa. For example, if the desirable R^2 is 0.95, then based on the equation $1/R^2 = 0.0418 \times$ Bandwidth/Grid Size + 0.9928 for summarizing grids 5 km or larger, a Bandwidth/Grid Size ratio of 1.431 would be needed to guarantee good quality outcomes. In the case of a regional study using 10 km summarizing grids, a bandwidth of 14.3 km or smaller should be considered. In the case of a synoptic level study using 50 km summarizing grids, a bandwidth of 71.6 km or smaller should be considered. A similar approach can be applied to requirements based on the standardized error of estimates. For example, a Bandwidth/Grid Size ratio of 1.431 will render a SE/Mean ratio of 0.227, and a SE/0.5 of data range ratio of 0.034.

3.3 Examples of Topographic Analysis of Lightning Distribution

To illustrate potential applications of density surfaces derived from different bandwidths, we performed analysis on the relationship between lightning frequency and topographic variables, such as elevation and slope aspect. It has been a common belief that lightning strikes occur on high topographic locations. However, previous studies on the relationship between lightning occurrence and topography were inconclusive (Fowler and Asleson, 1984; Potter et al., 1998; Dissing and Verbyla, 2003). Based on the general rule identified in the previous section, we matched the summarizing grid size with the bandwidth of the density surface. In other words, a bandwidth/grid size ratio of 1.0 was maintained for all following analyses to ensure good performance of the raster density surfaces. The original 1-km resolution DEM data were not used directly because of the poor results from the 1-km summarizing grids. The DEM data were first resampled to 5 km, 10 km, 20 km, and 50 km resolutions using the cubic convolution algorithm (Lillisand and Kiefer, 1999) to match the bandwidths and represent the spatial scales of analysis, shifting

Table 3. Results of Regression Analysis in Assessing the Performance of Density Surfaces of Different Kernel Bandwidth (KB) against the Actual Counted Lightning Frequency Using Different Summarizing Fishnet Grid Sizes

Model	R^2	Std. Error of the Estimate	Summarizing Grid Size (km)	KB/Grid Size
1 km KB	0.9948	0.402	5	0.20
5 km KB	0.9553	1.176	5	1.00
10 km KB	0.8929	1.822	5	2.00
20 km KB	0.8189	2.368	5	4.00
50 km KB	0.7208	2.941	5	10.00
1 km KB	0.9992	0.574	10	0.10
5 km KB	0.9935	1.670	10	0.50
10 km KB	0.9801	2.917	10	1.00
20 km KB	0.9325	5.374	10	2.00
50 km KB	0.8323	8.469	10	5.00
1 km KB	0.9999	0.841	20	0.05
5 km KB	0.9990	2.420	20	0.25
10 km KB	0.9970	4.319	20	0.50
20 km KB	0.9847	9.688	20	1.00
50 km KB	0.9145	22.932	20	2.50
1 km KB	1.0000	1.251	40	0.025
5 km KB	0.9999	3.366	40	0.125
10 km KB	0.9996	6.073	40	0.25
20 km KB	0.9978	13.844	40	0.5
50 km KB	0.9762	45.755	40	1.25
1 km KB	1.0000	1.393	50	0.02
5 km KB	0.9999	3.755	50	0.1
10 km KB	0.9998	6.854	50	0.2
20 km KB	0.9988	15.815	50	0.4
50 km KB	0.9854	55.331	50	1

* Results for 1 km summarizing grids were not included due to poor performance and multiple sampling areas, but are presented in Figures 8 and 9.

Table 4. Regression Equations between Metrics of Performance and the Bandwidth/Grid Size Ratio

Summarizing Grid Size	Dependent Variable	Regression Equation	R^2
1 km	$1/R^2$	= 1.3233 Ln(Bandwidth/Grid Size) + 1.4718	0.9608
	SE/Mean	= 0.2616 Ln(Bandwidth/Grid Size) + 1.1048	0.7193
	SE/(0.5 Data Range)	= 0.0701 $(\text{Bandwidth/Grid Size})^{0.1958}$	0.6356
≥ 5 km	$1/R^2$	= 0.0418 (Bandwidth/Grid size) + 0.9928	0.9606
	SE/Mean	= 0.1631 $(\text{Bandwidth/Grid Size})^{0.9243}$	0.9169
	SE/(0.5 Data Range)	= 0.0255 $(\text{Bandwidth/Grid Size})^{0.7891}$	0.9517

from local/micro, regional, to synoptic scales. Then the DEM cells were used as the zones to summarize the lightning density surface of the matching bandwidth. DEM cells of the same elevation (as integers) were aggregated so that for a given elevation value, there is a set of summary statistics of the lightning density surface. It should be pointed out that the spatial resolution of the DEM data and the accuracy of the lightning data do not allow an accurate account of the relationship between lightning occurrence and topography at the micro scale.

The mean density and total events (calculated as the sum of all cell values multiplied by the cell area) corresponding to each elevation value are presented in Figures 10 and 11. It is beyond the scope of this study to examine the mechanisms

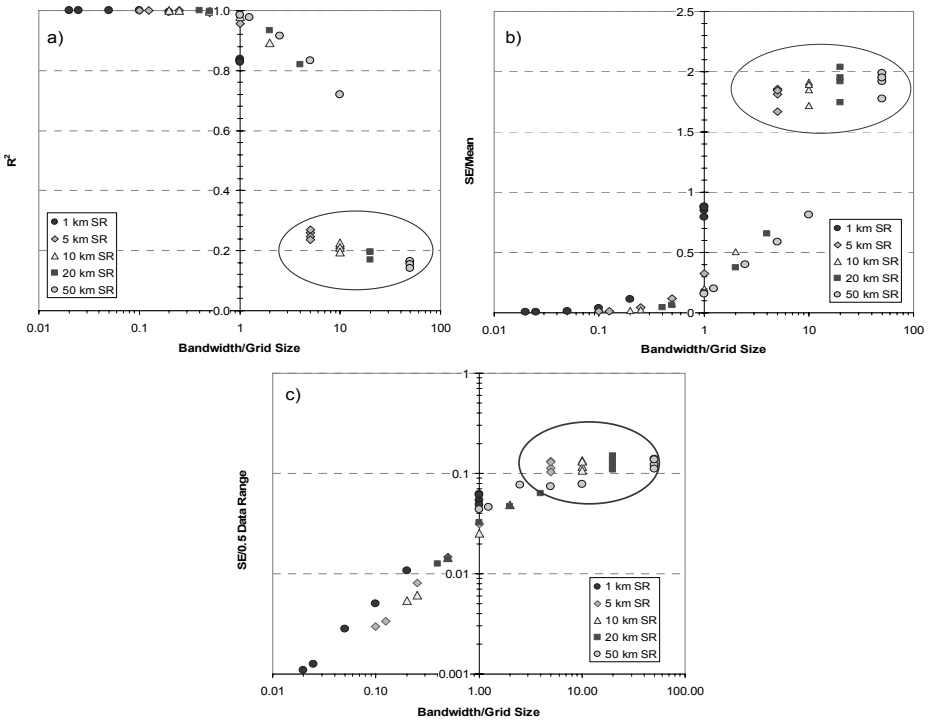

Figure 8. Performance metrics: R^2 (a), SE/mean (b), and SE/0.5 data range (c), plotted against the ratio of bandwidth to summarizing grid size.

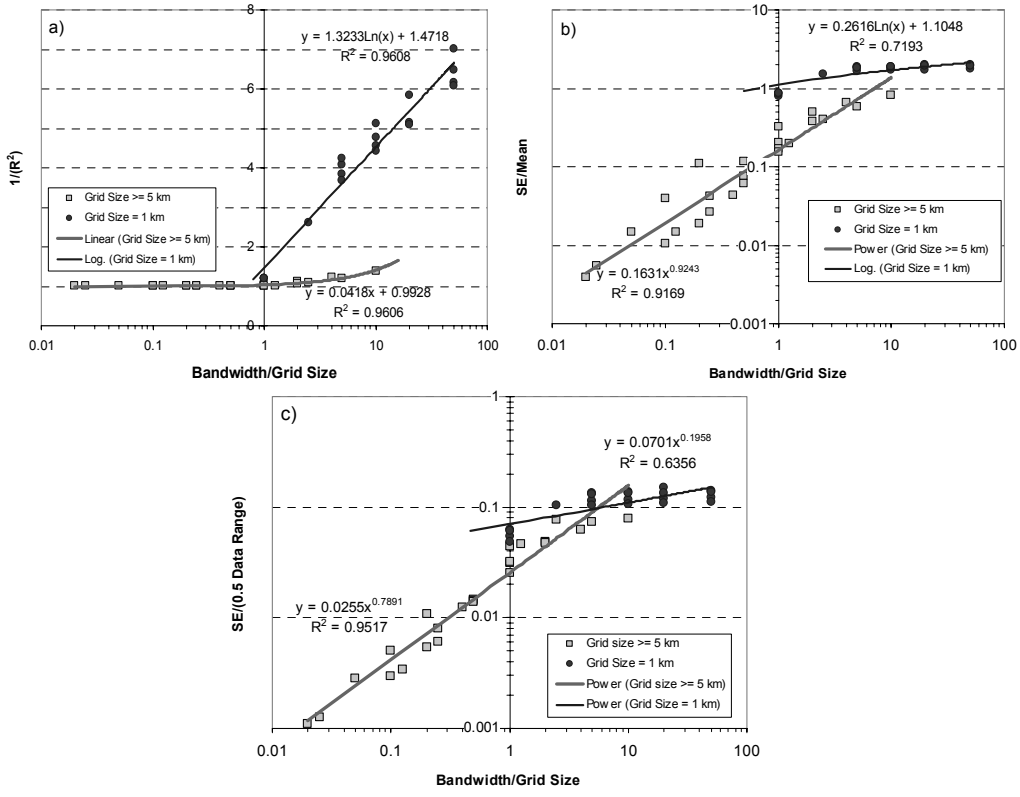

Figure 9. Result of curve-fitting for the performance metrics: R^2 (a), SE/mean (b), and SE/0.5 data range (c), plotted against the ratio of bandwidth to summarizing grid size.

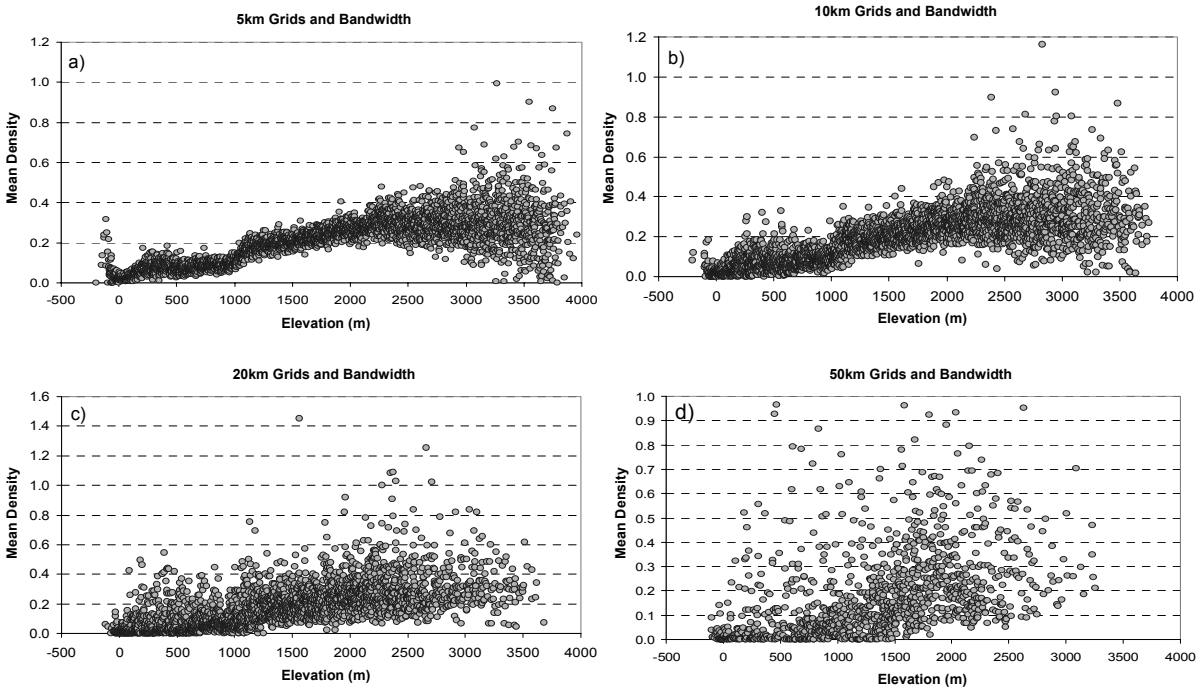

Figure 10. Mean lightning density by elevation at different spatial scales: 5 km (a), 10 km (b), 20 km (c), and 50 km (d).

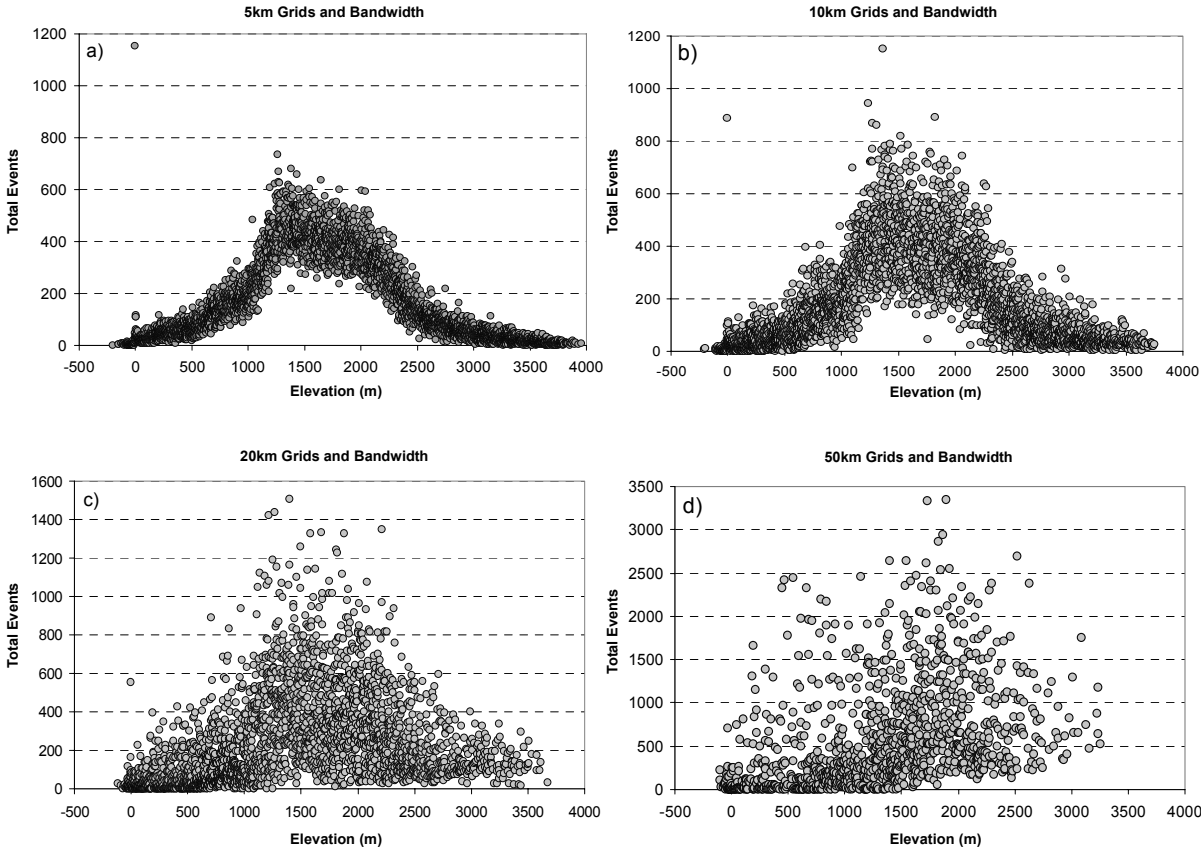

Figure 11. Total lightning events by elevation at different spatial scales: 5 km (a), 10 km (b), 20 km (c), and 50 km (d).

that determine the relationship between lightning occurrence and topography, but it can be clearly seen that the relationship between lightning density and elevation changes as the scale changes from the micro to synoptic level, apparently influenced by factors acting at different spatial scales. For example, at the micro and regional levels as represented by bandwidths of 5 and 10 km, there was a consistent trend of increase in mean lightning density with increasing elevation up to 3,000 m and then it was obscured by large variability at higher elevations. With the bandwidth of 20 km, the trend was still visible, and then it all but disappeared at the synoptic scale represented by the 50 km bandwidth. For total lightning events at various elevation values, there was a peak between 1000 ~ 1500 m, best defined by 5 km bandwidth, but visible for bandwidths of 10 and 20 km. Again, this peak disappeared for the 50 km bandwidth, suggesting that at the synoptic scale, topographic influences on lightning distribution patterns may be diminishing. It should be pointed out that the distribution of total lightning events by elevation is strongly influenced by the areal hypsography distribution in the study area and, therefore, may not be as revealing as the pattern of mean density by elevation. Similarly, we calculated slope aspects using the resampled DEMs and then divided the slope aspects into the eight major directions (N, NE, E, SE, S, SW, W, and NW) and flat (mostly water surfaces). For each aspect direction, the mean lightning densities were summarized (Figure 12). It can be seen that results obtained from the density surfaces of 20 and 50 km bandwidths were different from those from the smaller bandwidths. The distribution of lightning strikes on slopes of different aspects was relatively uniform at the micro-scale, with only a minor peak of lightning occurrence over the east-facing slopes based on the 5 ~ 10 km bandwidths and DEM resolutions. For 20 ~ 50 km bandwidths and DEM resolutions, however, the peak occurrence was over the south- and southwest-facing terrains and there was low occurrence over the north-facing slopes. Although we used only one month of data and the relationships between lightning occurrence and topography may not be spatially and temporarily robust, our results suggest that such relationships could vary with spatial scales and may have reflected the influence of different atmospheric and surface processes.

4. Conclusions

In this study we examined the method and accuracy of using GIS generated raster density surfaces to visually and quantitatively present the spatial distribution patterns of lightning occurrence, using the western United States as an example. Besides clear visual advantage of using density surfaces to present lightning data, we determined that satisfactory performance can be achieved as long as the kernel bandwidth matches the summarizing grid size at a given spatial scale. In other words, the bandwidth or search radius used to generate the density surface should be determined according to the needs of analysis. In this case, a ratio of bandwidth/grid size of 1.431 or lower renders good results with R^2 values higher than 0.95 and standard error/mean ratios lower than 0.25 in regression analysis of density-surface predicted values against the actual counted values for grids of 5 km to 50 km in size. However, there is probably no need to use a search radius smaller than half of the summarizing grid size. Once the above requirement is met, density surfaces performed very well both visually and quantitatively. This also suggests that density surfaces can be a means for data assimilation and as a form of model input.

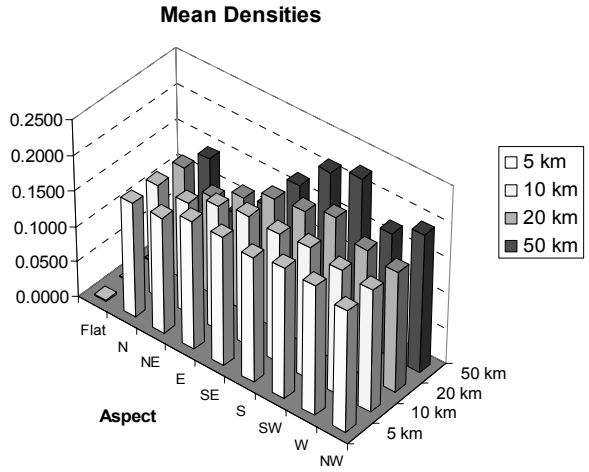

Figure 12. Lightning distribution by major slope aspects at different spatial scales: 5 km, 10 km, 20 km, and 50 km.

As grid size approaches the spatial resolution of the density surface (1 km) and the accuracy of the lightning data (0.5 ~ 1 km), the performance of density surfaces dropped drastically. Even the best performance of the 1-km summarizing grids was much worse than that of other grid sizes (R^2 = 0.82 ~ 0.83, SE/mean = 0.8 ~ 2.0). Therefore, we separated the 1 km summarizing grids from the rest in our regression analysis of performance for the raster density surfaces. Using these regression equations, a kernel bandwidth can be estimated once the spatial scale and a desirable performance level are determined.

We used raster density surfaces generated using kernel bandwidths from 5 km to 50 km to analyze the relationships between lightning occurrence in August 1990 and topographic factors (elevation and slope aspect) as examples. Our results suggested that such relationships, if valid, may change with the spatial scale, as the processes reflected by these relationships may be different at micro, regional, and synoptic levels. Through these examples, the potential of density surfaces in studying spatial patterns of lightning occurrence was clearly demonstrated. It is the intention of the authors to further investigate the spatial patterns of lightning events in the western United States in relation to a number of factors, such as vegetation, topography, and elevation (Potter et al., 1998; Dissing and Verbyla, 2003). Results of this study demonstrated that density surfaces can be a valuable analytical tool for this purpose.

Acknowledgments. This study was in part supported by a grant from the University of San Diego (FRG 04-05). The authors would

like to thank the anonymous reviewer who provided constructive comments and suggestions.

References

Bailey, T.C. and Gatrell, A.C. (1995). *Interactive Spatial Data Analysis*, Longman, London, pp. 413.

Bergeron, Y., Leduc, A. and Li, T.X. (1997). Explaining the distribution of *Pinus spp.* in a Canadian boreal insular landscape. *J. Vegetat. Sci*., 8(1), 37-44.

Berz, G., Kron, W., Loster, T., Rauch, E., Schimetschek, J., Schmieder, J., Siebert, A., Smolka, A. and Wirtz, A. (2001). World map of natural hazards-A global view of the distribution and intensity of significant exposures. *Nat. Hazards*, 23, 443-465.

Blong, R. (1997). A geography of natural perils. *Aust. Geogr*., 28(1), 7-27.

Coates, L., Blong, R. and Siciliano, F. (1993). Lightning fatalities in Australia, 1824-1991. *Nat. Hazards*, 8(3), 217-233.

Chronis, T.G., Anagnostou, E.N. and Dinku T. (2004). High-frequency estimation of rainfall from thunderstorms via satellite infrared and a long-range lightning network in Europe. *Q. J. Royal Meteorol. Soc*., 130 (599, Part B), 1555-1574.

Cummins, K.L., Murphy, M.J., Bardo, E.A., Hiscox, W.L., Pyle, R.B. and Pifer, A.E. (1998). A combined TOA/MDF technology upgrade of the U.S. National Lightning Detection Network. *J. Geophys. Res*., 103 (D8), 9035-9044.

Dent, B.D. (1999). *Cartography-Thematic Map Design*, 5th Edition, WCB/McGraw-Hill, Boston, pp. 417.

Diaz-Avalos, C., Peterson, D.L., Alvarado, E., Ferguson, S.A. and Besag, J.E. (2001). Space-time modeling of lighning-caused ignitions in the Blue Mountains, Oregon. *Can. J. For. Res*., 31, 1575-1593.

Dissing, D. and Verbyla, D.L. (2003). Spatial patterns of lightning strikes in interior Alaska and their relations to elevation and vegetation. *Can. J. For. Res*., 33, 770-782.

Elsom, D.M. (1996). Surviving being struck by lightning: A preliminary assessment of the risk of lightning injuries and death in the British Isles. *J. Meteorol*., 21(209), 197-206.

Elsom, D.M. (2001). Death and injuries caused by lightning in the United Kingdom: Analyses of two databases. *Atmos. Res*., 56, 325-334.

Ezcurra, A., Areitio, J. and Herrero, I. (2003). Relationships between cloud-to-ground lightning and surface rainfall during 1992-1996 in the Spanish Basque Country area. *Atmos. Res*., 61(3), 239-250.

Fowler, P.M. and Asleson, D.O. (1984). The location of lightning-caused wildland fires, northern Idaho. *Phys. Geogr*., 5(3), 240-252.

Gedalof, Z., Peterson, D.L. and Mantua, N.J. (2005). Atmospheric, climatic, and ecological controls on extreme wildfire years in the Northwestern United States. *Ecol. Appl*., 15, 154-174.

Griffin, G.F., Price, N.F. and Portlock, H.F. (1983). Wildfires in the central Australian rangelands, 1970-1980. *J. Environ. Manage*., 17(4), 311-323.

Hardy, C.C. (2005). Wildland fire hazard and risk: problems, definitions, and context. *For. Ecol. Manage*., 211, 73-82.

Kempf, N.M. and Krider, E.P. (2003). Cloud-to-ground lightning and surface rainfall during the great flood of 1993. *Mon. Weather Rev*., 131(6), 1140-1149.

Lillesand, R.W. and Kiefer, T.M. (1999). *Remote Sensing and Image Interpretation*, 4th Edition, Wiley and Sons, New York, pp. 724.

Livingston, E.S., Nielsen-Gammon, J.W. and Orville, R.E. (1996). A climatology, synoptic assessment, and thermodynamic evaluation for cloud-to-ground lightning in Georgia: a study for the 1996 Summer Olympics. *Bull.-Am. Meteorol. Soc*., 77(7), 1483-1495.

Lo, C.P. and Yeung, A.K.W. (2002). *Concepts and Techniques of Geographic Information Systems*, Prentice Hall, Upper Saddle River, New Jersey, pp. 492.

López, R.E., Holle, R.L., Watson, A.I. and Skindlov, J. (1997). Spatial and temporal distributions of lightning over Arizona from a power utility perspective. *J. Appl. Meteorol*., 36, 825-831.

Molinie, G. and Jacobson, A.R. (2004). Cloud-to-ground lightning and cloud top brightness temperature over the contiguous United States. *J. Geophys. Res*., 109, D13106, doi: 10.1029/2003JD003593.

Morales, C.A. and Anagnostou, E.N. (2003). Extending the capabilities of high-frequency rainfall estimation from geostationary-based satellite infrared via a network of long-range lightning observations. *J. Hydrometeorol*., 4(2), 141-159.

Naccarato, K.P., Pinto, O.Jr. and Pinto, I.R.C.A. (2003). Evidence of thermal and aerosol effects on the cloud-to-ground lightning density and polarity over large urban areas of Southeastern Brazil. *Geophys. Res. Lett*., 30(13), 1674, doi:10.1029/2003GL017496

Nicholas, R. (2006). Create a grid polygon shapefile (FISHNET), website of ESRI Support Center. http://arcsripts.esri.com/details.asp?dbid=12807

Petersen, W.A. and Rutledge, S.A. (1998). On the relationship between cloud-to-ground lightning and convective rainfall. *J. Geophys. Res. D: Atmos*., 103(D12), 14025-14040.

Pinto, I.R.C.A., Pinto, O.Jr., Gomes, M.A.S.S. and Ferreira, N.J. (2004). Urban effect on the characteristics of cloud-to-ground lightning over Belo Horizonte-Brazil. *Ann. Geophysicae*, 22, 697-700.

Potter, D.U., Gosz, J.R., Molles, M.C. Jr. and Scuderi, L.A. (1998). Lightning, precipitation and vegetation at landscape scale. *Landscape Ecol*., 13, 203-214.

Rathert, D. (2006). Raster to XYZ, website of ESRI Support Center. http://arcscripts.esri.com/details.asp?dbid=12622.

Rorig, M.L. and Ferguson, S.A. (1999). Characteristics of lightning and wildland fire ignition in the Pacific Northwest. *J. Appl. Meteorol*., 38, 1565-1575.

Rorig, M.L. and Ferguson, S.A. 2002. The 2000 fire season: lightning-caused fires. *J. Appl. Meteorol*., 41, 786-791.

Seity, Y., Soula, S. and Sauvageot, H. (2001). Lightning and precipitation relationship in coastal thunderstorms. *J. Geophys. Res*., 106(D19), 22801-22816, doi: 10.1029/2001JD900244.

Sheridan, S.C., Griffiths, J.F. and Orville, R.E. (1997). Warm season cloud-to-ground lightning-precipitation relationships in the south-central United States. *Weather Forecast.*, 12(3), Part 1, 449-458.

Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London, pp. 175.

Smith, J.A., Baker, M.B. and Weinman, J.A. (2003). Do forest fires affect lightning? *Q. J. Royal Meteorol. Soc*., 129(593), Part B, 2651-2670.

Soriano, L.R. and de Pablo, F. (2002). Effect of small urban areas in central Spain on the enhancement of cloud-to-ground lightning activity. *Atmos. Environ*., 36, 2809-2816.

Stallins, J.A. (2004). Characteristics of urban lightning hazards for Atlanta, Georgia. *Clim. Change*, 66, 137-150.

Steiger, S.M., Orville, R.E. and Huffines, G. (2002). Cloud-to-ground lightning characteristics over Houston, Texas: 1989-2000. *J. Geophys. Res*., 107(D11), 4117, doi: 10.1029/2001JD001142.

Stocks, B.J., Wotton, B.M., Amiro, B.D., Flannigan, M.D., Hirsch, K.G., Logan, K.A., Martell, D.L., Skinner, W.R., Mason, J.A., Todd, J.B. and Bosch, E.M. (2003). Large forest fires in Canada, 1959-1997. *J. Geophys. Res*., 108(D1), 8149, doi: 10.1029/2001JD000484.

Tapia, A., Smith, J.A. and Dixon, M. (1999). Estimation of convective rainfall from lightning observations. *J. Appl. Meteorol*., 37, 1497-1509.

Underwood, S.J. and Schultz, M.D. (2003). Cloud-to-ground lightning flashes and debris-flow-generating rainfall in the post wildfire environment: an exploratory study of the Mitchell Creek debris flow in western Colorado, summer 2002. *J. Geophys. Res*., 108(D18), 4567, doi: 10.1029/2003JD003636.

Underwood, S.J. and Schultz, M.D. (2004). Patterns of cloud-to-ground lightning and convective rainfall associated with postwildfire flash floods and debris flows in complex terrain of the western United States. *J. Hydrometeorol*., 5, 989-1003.

Watson, A.L. and Holle, R.L. (1996). An eight-year lightning climatology of the southeaster United States prepared for the 1996 Summer Olympics. *Bull.-Am. Meteorol. Soc*., 77(5), 883-890.

Westcott, N.E. (1995). Summertime cloud-to-ground lightning activity around major Midwestern urban areas. *J. Appl. Meteorol*., 34(7), 1633-1642.

Williamson, D., McLafferty, S., McGuire, P., Goldsmith, Y. and Mollenkopf, J. (1999)**.** A better method to smooth crime incidence data, *ArcUser*. http://www.ersi.com/news/arcuser/0199 /crimedata.html.