Identifying Crime Hot Spots in Space and Time:

An Analysis of Dallas Police Department Reports of Criminal Activity

Janis L. Schubert

University of Texas at Dallas

GIS Workshop Summer 2004

August 1, 2004

Abstract

The allocation of limited police resources in major metropolitan areas requires extensive information and planning to optimize effectiveness in the reduction of crime and to increase the safety of residents. One important technique is the identification of hot spots, or locations with unusually high crime rates. The identification of hot spots in time can also be very important, allowing for the strategic deployment of resources at times when they can make the greatest difference. A visual approach to locating temporal hot spots within certain categories of crime was presented. This approach was also demonstrated in locating temporal hot spots within spatial hot spots.

Keywords

hot spots; GIS; crime reduction; crime forecasting

Introduction

As cities have grown larger and more mobile and their populations more unstable, it has become increasingly necessary for law enforcement departments to utilize sophisticated technology to fight crime and to increase the effectiveness of the available resources. If analysts are able to predict the time and place of hot spots, or high crime areas, law enforcement personnel can be better prepared to either reduce the intensity of the hot spot or be able to more quickly mobilize officers to handle the increased crime rate.

Obvious high crime areas have long been identified based on location, but with changes in the way urban residents live and work, a temporal component becomes more important. Areas that are filled with white-collar workers during the normal workday may change dramatically in the evenings or on weekends. Suburban neighborhoods may be all but deserted during the middle of the day. Areas with convention centers or sports arenas may be packed with people one day and empty the next.

For this project, extensive data from the Dallas Police Department is used to demonstrate a preliminary methodology to identify hot spot times for crime in the city of Dallas. This methodology is then used to determine hot spot times within spatial hot spots.

Problem Statement

The purpose of this research was to develop a preliminary methodology by which the timeframe of hot spots of crime within the city of Dallas could be identified.

Crime patterns are inherently complex and uncertain. They develop over time due to an intricate interaction between the target of the crime and the offender over space. If we are to better understand the behavior of offenders, we need models that incorporate both space and time. Geographic Information Systems (GIS) are well suited to presenting the spatial aspects of crime events and discovering spatial correlations. However, additional research and analysis is necessary to determine effective techniques for recognizing significant temporal correlations.

In this project, various types of crime were studied to develop a visual technique by which their distinctive temporal characteristics can be identified. This technique was then used to analyze events within specific spatial hot spots to identify peak times of criminal activity. The resulting information about the space and time of hot spots can then be of use for future predictive modeling.

Literature Review

Over the years, many theories have been presented in attempts to define and explain criminal activity. Some of these have focused on individual criminals while others examine the aggregate crime within an area. Although the public perception may be that crime is randomly distributed in space, extensive evidence now exists that it is not.

Sherman, Gartin & Buerger (1989) detailed the many early problems associated with the analysis and identification of high crime areas or hot spots. An inability to pinpoint the location and/or the time of some crimes added error into the analysis. Also, until recently, reliable, complete data was not generally available.

As ever smaller police departments make the transition to computerized systems, both the quantity and quality of crime data available for research has seen tremendous improvement. However, the overwhelming quantity of data and the difficulty associated with processing such multidimensional data presents new difficulties (Guo, Peuquet & Gahegan, 2003). Exponential growth in the power of computers and the development of sophisticated GIS systems has advanced this research tremendously and it is now possible to process large amounts of spatial data and determine trends where they may not have been previously visible.

Several techniques have been used to analyze crime data. One limitation that has arisen when techniques are applied to administrative areal units such as census tracts or police beats is called the Modifiable Areal Unit Problem (MAUP). This problem occurs when different results are produced depending on the selection of areas within which the data is aggregated (O'Sullivan & Unwin, 2003). This issue is of particular importance in the study of crime. Since crime does not necessarily occur in any particular relation to the land area or the residential population, analyzing a simple spatial concentration of crime can be more valuable (Ratcliffe, 2004). Studies have produced several methods of spatially identifying hot spots. Among these methods are the Spatial and Temporal Analysis of Crime (STAC), which produces ellipses which cluster crime points (Block, 1995), and methods using kernel density estimation (Levine, 1998).

Nelson, Bromley, & Thomas (2001) note that the most common feature of all categories of crime reported was the spatial concentration into a relatively small number of hot spots allowing the anticipation of high crime in these areas in the future. It was also noted that these concentrations peaked at fairly specific times on a regular basis. According to Liu & Brown (2003), the prevalent forecasting method is currently primarily spatial in nature, with the assumption that clusters of crime will persist in the near future. Researchers have noted that a system capable of utilizing existing information in real-time and able to predict where and when high rates of crime are expected would be of great value for police resource allocation (Corcoran, Wilson, & Ware, 2003). Also, a greater understanding of the patterns of crime can result in a better grasp on the causes, and lead to improvements in urban planning and design to reduce susceptibility to crime (Nelson, et. al., 2001).

One problem encountered when attempting to analyze the temporal qualities of the crime data involves the lack of detail in many police crime databases. In crimes such as burglary or theft, it can be impossible to pinpoint the actual time of occurrence and the possible window of opportunity may cover many hours. The Aoristic Temporal Analysis Method, a temporal weighting method, was developed to attempt to deal with this problem (Ratcliffe, 2002). Ratcliffe (2004) also proposed several general categories of temporal clustering based on prevalent patterns.

A crime data analysis model utilizing information on spatial location as well as time of day would have the advantage of being able to pinpoint more accurately when additional policing may be needed in an area. Also, if certain areas are noted as problematic primarily during holidays or specific weekdays, this information can then be used to make better predictions and to improve the allocation of resources. When utilized with real-time data, such a model could possibly help to identify crime trends as they develop in unexpected areas.

Research continues on the integration of temporal components with spatial data. Peuquet (2001) notes that extending the relational model to include data in the temporal dimension can be effective when the temporal data is linear. However, the common use of temporal “snapshots” leads to inefficiency and inflexibility. Future work will continue to develop the theoretical framework and techniques to deal with both linear and non-linear temporal data. Corcoran et. al. (2003) have demonstrated the use of artificial neural networks to facilitate predictive modeling. Any effective approach to modeling the temporal dimension must incorporate both absolute and relative views of space-time to adequately represent temporal dynamics in geographic information systems (Peuquet, 1994).

Data Description

The data for this project was collected by the Dallas Police Department as a part of their normal operating procedure. Detailed information on calls for service, arrests, crime reports, and offenders was available. The decision was made to use the crime reports to identify the hot spots.

Each record identified a single report of criminal activity that was filed. The data fields used to identify the hot spots spatially and temporally were the starting date and time of the event, ending date and time of the event, address of the event, and the event type code, which identified the type of crime that was committed.

This data was provided on several CDs and covered a period of five calendar years, 1998-2002. The entire collection of data was then loaded into a single Oracle database from which all the information could be accessed.

Methods or Analysis

To make the data useful for our purposes, the first thing that needed to be done was to geocode the addresses of the events. This was neither a straightforward nor simple task. The Dallas Police Department (DPD) does not record addresses in a consistent or generally accepted format. In many cases, all blanks were removed from the address data, leading to confusion as to where the street name ended and the street type (if included) began. Others were recorded as intersections only, again with all blanks removed. In some cases, the address was a shopping center name and suite number, and in others an apartment number preceded the address. Directional designations (N/S or E/W) were missing in most cases. The greatest difficulty involved the 1998 data where in most cases the first two letters of the street names were removed and the remaining name recorded with no blanks. For example, 125 OAK ST became 125KST. In many cases, it was impossible to determine the correct address due to multiple possibilities.

Pre-processing of the addresses was performed using a custom application developed at UT-Dallas by Timothy Bray and Karen Hayslett-McCall (Bray & Hayslett-McCall, 2004). This application used both automated and interactive approaches to correct as many formatting problems as possible. Decisions were recorded in a database for use with future sets of data.

The data was then brought into ArcMap for geocoding. Often the zip code was not available so it was not used when determining the success of geocoding matches. With the crime reports data used (1999-2002) the final geocoding success rate on 1,223,107 records was approx. 91%.

To determine hot spots spatially, there are many different techniques available. Several were analyzed and it was determined that the Spatial and Temporal Analysis of Crime (STAC) (Block, 1995) technique implemented in the *CrimeStat II* software (Levine, 2002) produced results that were the most useful in focusing in on the hot spots. Several window sizes were used with a size of .25 mi. producing the best results.

Identifying the temporal distribution of crimes required the determination of the time the event occurred. In some crimes, such as assault, this is fairly straightforward, since most victims know with certainty when it occurred. But in other crimes, usually those without witnesses, it may be only possible to place the event within a span of time which can cover several hours, a day, or even longer. To handle these situations, the Aoristic Temporal Analysis Method (Ratcliffe, 2002) was used. This method spreads the influence of an event across the entire span of possible times. Events with exact occurrence times will thus have more weight in the distribution than those with indeterminate times. The event start date and time and the event end date and time were used to determine the span of time involved in hours. A weight of one per event is divided by the number of hours spanned to produce a weight per hour. Once the weight per hour elapsed has been calculated, events can be selected and summed to produce a wide variety of temporal distributions.

For this project, four types of crime were selected for analysis: burglary of a motor vehicle, unauthorized use of a motor vehicle, assault, and vandalism. For each category except vandalism, temporal distributions of hour x day and hour x weekday were produced for the year 2002. Distributions for vandalism were produced using 2001 data when it was discovered that the 2002 data was incomplete. These distributions were then recorded in a table with X as the day of the year or weekday, Y as the hour, and Z as the accumulated weights for this X and Y.

The table of X, Y, and Z values was imported into ArcMap and used to create a points shapefile. The shapefile was then interpolated to raster using kriging to produce a surface representing the temporal distribution. The resulting grid was loaded into ArcScene where it could be rendered in three dimensions and manipulated as needed. Temporal hot spots appear on this surface as peaks. Other trends, including seasonal ones, can be identified on the resulting surface.

Results

Currently many police departments use software such as *CrimeStat II* (Levine, 2002) to identify hot spots spatially. A shapefile is produced which contains ellipses which delineate the hot spots. As an example, reports of motor vehicle burglaries for 2002 were analyzed and the spatial hot spots identified. The static map in Figure 1 was produced which shows the four major hot spots for this type of crime within the City of Dallas. No information is provided as to the timeframe of these events.

{Figure 1 goes here}

When the aoristic analysis is used to produce a temporal distribution for the events within the largest of the four hot spots, this distribution is expressed in the bar chart in Figure 2. Although this method is adequate to identify fairly broad or obvious trends, less obvious trends may be hard to spot.

{Figure 2 goes here}

Utilizing the day/weekday and hour as X and Y values for a surface, and expressing the total of events as Z, a temporal “map” is produced. Figure 3 shows the two-dimensional surface which represents the distribution of motor vehicle burglaries over a span of approximately three months. A general idea of the temporal characteristics begins to emerge, but it is still not very easy to see.

{Figure 3 goes here}

With the capabilities of ArcScene, our temporal “map” is rendered in three dimensions. As we can see in Figure 4, distinct trends are clearly visible. Not only can daily and weekly cycles be easily identified, but also seasonal differences are visible.

{Figure 4 goes here}

Distinct patterns can be seen when surfaces are produced for the different categories of crime. Comparing Figures 5 – 7 with Figure 4 demonstrates the unique characteristics of each different type of crime. In motor vehicle burglaries, the highest activity is in the very early morning hours, and there is increased activity in the months leading up to Christmas. Unauthorized use of a motor vehicle shows less seasonality, higher activity in both early morning and late night hours, and some increased activity during the summer. Assault crimes are heavily weighted in the evening hours, much more frequent in the warmer months, and the characteristic and regular ripples at the edges show a strong weekly cycle. Vandalism shows the regular rippling of a weekly cycle, but very little in the way of seasonal changes. One thing to note about this surface is the apparent area of greatly increased activity at the beginning of the year. Seeing this visually would immediately prompt a review of the data to determine if this is a real phenomenon. It is most likely that it is a reporting anomaly and the visual cues make it very noticeable.

{Figures 5 – 7 go here}

Although some types of crime show more distinctive rippling than others, indicating a stronger correlation between the day of the week and the number of events, all the categories of crime studied showed distinctive patterns over the course of a week. To better identify these patterns, the same process as before was used to produce the temporal "map" by weekday. Figures 8 – 11 show the temporal "maps" for the four categories studied. Not only do these surfaces clearly identify weekday trends, but they also show the hot spot times of those days. Such a tool could be very useful in identifying the most effective times to increase police presence in high crime areas.

{Figures 8 – 11 go here}

Finally, the temporal characteristics of the largest of the motor vehicle burglary spatial hot spots in Figure 1 were analyzed. The temporal "map" in Figure 12 was produced by this spatial hot spot. It is very similar to the one produced using all the motor vehicle burglary events in 2002. However, a greater "bulge" of activity around lunchtime may point to unusual conditions in the area that modify the characteristic distribution for this type of crime. In this case, we know that this hot spot is in downtown Dallas and it is possible that the increased burglaries at lunchtime are the result of an increase in the number of vehicles parked on the street or in open lots at restaurants. This could suggest that increased vigilance by police around parking lots at this time might be recommended.

{Figure 12 goes here}

Conclusions and Future Research

When temporal information about a specific type of crime is presented visually, distinct patterns emerge. These patterns can be studied for clues about cyclical variations and to identify the time of day, the week day, or the date of increased criminal activity. Presenting the information visually in three dimensions leads to a more intuitive determination of temporal hot spots.

The identification of these patterns of activity can be very useful in understanding the timeframe in which spatial hot spots of crime exist, and in planning police activities to ameliorate the conditions leading to increased crime both in general and within hot spots which have been previously defined spatially. Events within specific hot spots can be analyzed to determine if they fit the typical pattern or if they produce a unique pattern which may point to an unusual and possibly modifiable problem which is encouraging the increased activity.

The study of the temporal patterns of activity can also be of use on other levels. Understanding the distinctive temporal patterns of specific types of crime could aid researchers in the understanding of these types of crime and may help to suggest areas of future study. A first step to defining temporal differences in activity is clearly identifying these differences. This technique shows these differences on a general level.

When working with crime data, accuracy will always be an issue. Often reports rely on a victim's memory, conflicting reports of witnesses or bystanders, or in the case of crimes with no witnesses, the best that can be done is to define a probable time span during which the event occurred. This all leads to inexactness in the temporal distribution. However, even with some inexactness, spatial analysis produces valid and useful results and trends can be determined. The same is true for temporal analysis. The aoristic approach reduces the effects of indeterminate events on the resulting distribution without removing them entirely. But it is important that records be maintained in as accurate a fashion as possible to increase their value as a useful predictive tool.

Temporal analysis of the type demonstrated in this study is most valid for types of crime where the time of occurrence is known or the probable timeframe is fairly small. Greater accuracy is then seen in the temporal "map" and it thus becomes more useful in predicting future hot spot times.

This study suggests a number of areas that would benefit from further research. Temporal information needs to be fully integrated with spatial information to produce the most useful hot spot analysis. This would be far more valid than the current two-step approach of either spatial analysis followed by temporal analysis or vice versa. Research is already progressing on the development of temporal topology and this will be extremely useful in the area of hot spot analysis when it becomes available.

Until that time, there is a need for tools that can facilitate the type of analysis presented in this study. These tools would integrate the STAC spatial analysis with the temporal analysis and produce spatial and temporal maps in realtime. Techniques also need to be developed to animate the progression of hot spots in space and time. These techniques could then be used for predictive modeling and simulations and lead to a more intuitive grasp on the identification of hot spots in the future.

Current hot spot algorithms require a critical mass of data to produce valid results. For this reason, a full year's data is often analyzed. Because of this, any temporal variations are lost. It would be very useful to develop hot spot spatial analysis methods which do not require as large a critical mass of data, and ideally ones that did not use static data. One possibility would be developing a process similar to the roving window approach used in STAC with the addition of a temporal element.

In this study, ArcScene was used to render the temporal “maps” in three dimensions. However, ArcScene has very limited capabilities for the manipulation of this data. Improved and integrated graphics capabilities need to be developed to better facilitate the use of these techniques and to improve the ability to see the resulting trends intuitively.

The study of crime data and the behavior of criminals is a field that is growing at a phenomenal rate at this time. In the past decade, many police departments have installed and begun using sophisticated computer systems to maintain their information. The massive amount of data produced has previously been a stumbling block for researchers due to the time and equipment required to access and utilize it. Use of the resulting information produced from this data often required computer and statistical skills that are not commonly held by those personnel who need these results. However, current advances in computing speed and capabilities, and the continuing downward trends in cost have made it increasingly possible to put this wealth of data to use. With advances in computer graphics capabilities, techniques that are more visual and intuitive are possible which opens up this information to users who are neither statistical specialists nor computer professionals. The future in the study of crime data for policing purposes is the development of tools utilizing realtime data that are intuitive, easy to use, and quick to produce. This would open up this wealth of information to every police department and greatly assist in the understanding of patterns of criminal activity.

References

Block, C. 1995. STAC hot-spot areas: A statistical tool for law enforcement decisions. In Block, C. R., Dabdoub, M., & Fregly, S. (Eds.), *Crime analysis through computer mapping.* Washington, DC: Police Executive Research Forum, p. 20036.

Bray, Timothy, and Hayslett-McCall, Karen. 2004. Custom software application for address pre-processing on City of Dallas crime data.

City of Dallas Police Department. Crime data 1998-2002.

Corcoran, Jonathan J., Wilson, Ian D., and Ware, J. Andrew. 2003. Predicting the geo-temporal variations of crime and disorder. *International Journal of Forecasting* **19**: 623-634.

Guo, Diansheng, Peuquet, Donna J., and Gahegan, Mark. 2003. ICEAGE: Interactive clustering and exploration of large and high-dimensional geodata. *GeoInformatica* **7:3** 229-253.

Levine, N. 1998. “Hot Spot” analysis using *CrimeStat* kernel density interpolation. *Presentation at the Annual Meeting of the Academy of Criminal Justice Sciences, Albuquerque, NM, March 10-14, 1998.*

Levine, Ned. 2002. *CrimeStat II: A Spatial Statistics Program for the Analysis of Crime Incident Locations (version 2.0).* Ned Levine & Associates: Houston, TX / National Institute of Justice: Washington, DC.

Liu, Hua, and Brown, Donald E. 2003. Criminal incident prediction using a point-pattern-based density model. *International Journal of Forecasting* **19**: 603-622.

Nelson, A.L., Bromley, R.D.F., and Thomas, C.J. 2001. Identifying micro-spatial and temporal patterns of violent crime and disorder in the British city centre. *Applied Geography* **21**: 249-274.

O'Sullivan, David, and Unwin, David J. 2003. *Geographic Information Analysis.* Hoboken, New Jersey: John Wiley & Sons, Inc.

Peuquet, Donna J. 1994. It's about time: A conceptual framework for the representation of temporal dynamics in geographic information systems. *Annals of the Association of American Geographers* **84:3** 441-461.

Peuquet, Donna J. 2001. Making space for time: Issues in space-time data representation. *GeoInformatica* **5:1** 11-32.

Ratcliffe, J. H. 2002. Aoristic signatures and the temporal analysis of high volume crime patterns. *Journal of Quantitative Criminology* **18**: 23-43.

Ratcliffe, J. H. 2004. The hotspot Matrix: A Framework for the spatio-temporal targeting of crime reduction. *Police Practice and Research* **5**: 5-23.

Sherman, Lawrence W., Gartin, Patrick R., and Buerger, Michael E. 1989. Hot spots of predatory crime: Routine activities and the criminology of place. *Criminology* **27**: 27-55.

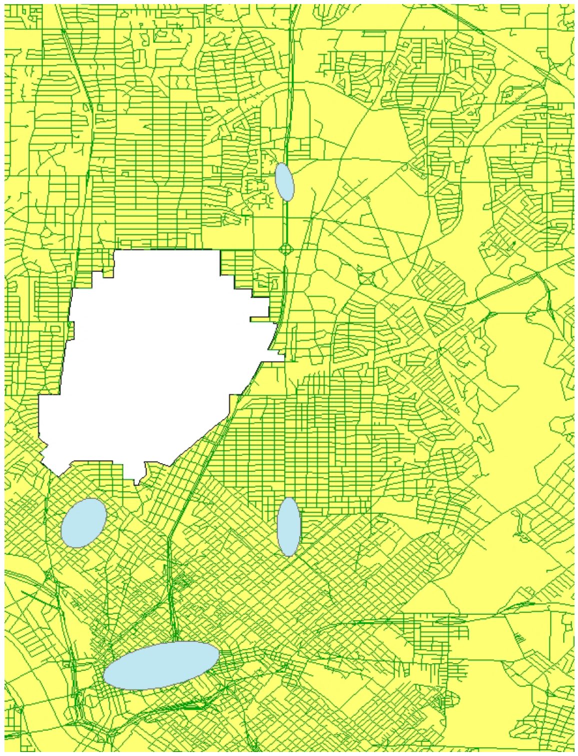

Figure 1: Section of North Central Dallas with

with motor vehicle burglary hot spots identified.

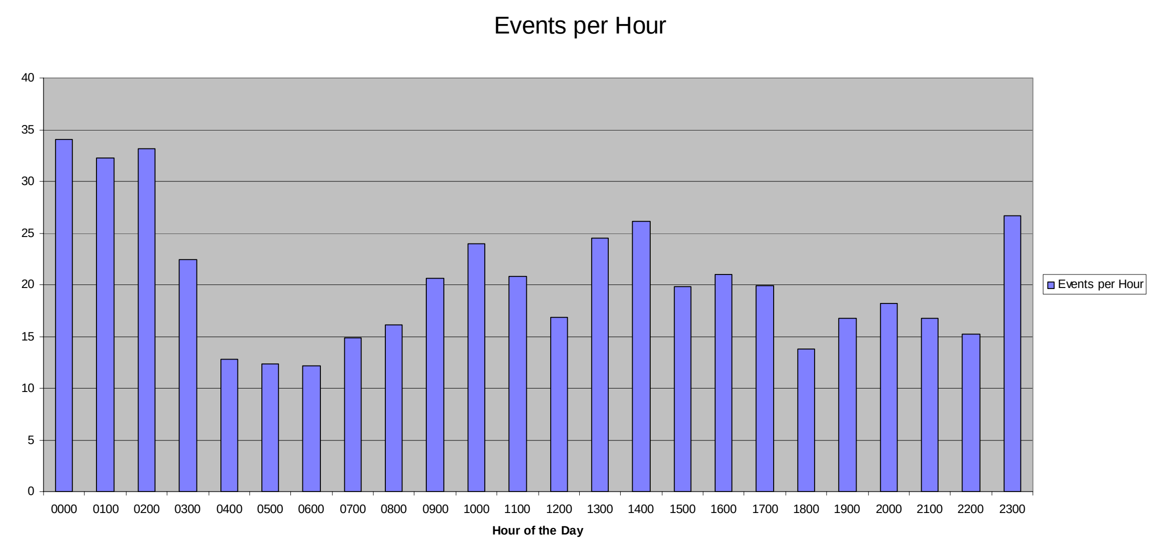

Figure 2: Temporal distribution for motor vehicle burglaries in largest hot spot

in Figure 1.

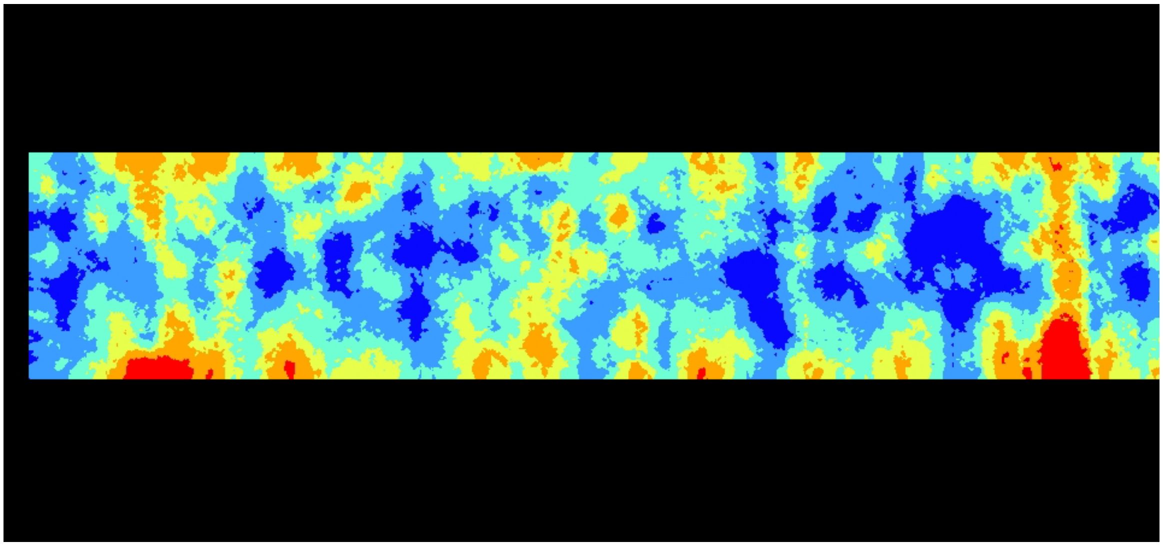

Figure 3: Two-dimensional temporal "map" of motor vehicle burglaries in 2002, January through March approximately.

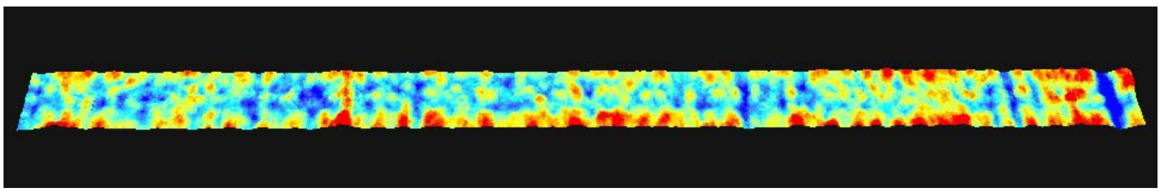

Figure 4: Motor vehicle burglary – 2002.

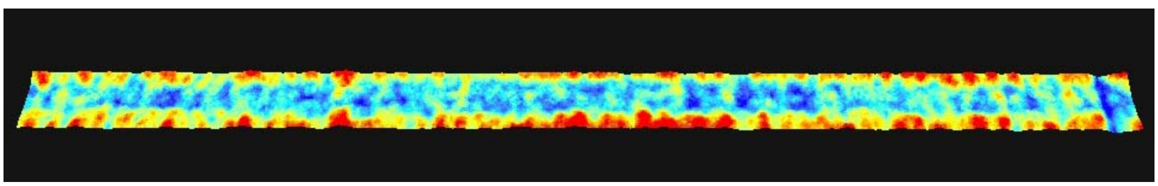

Figure 5: Unauthorized use of a motor vehicle – 2002.

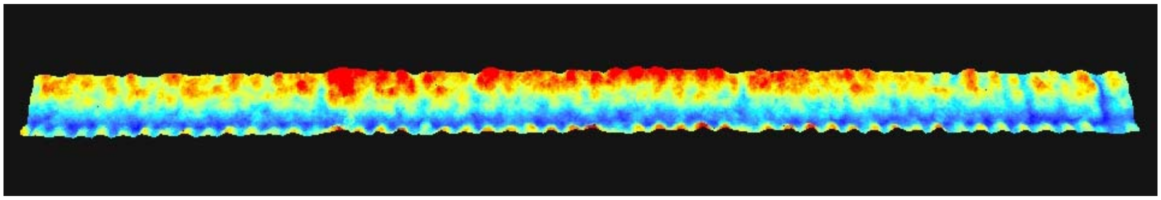

Figure 6: Assault – 2002.

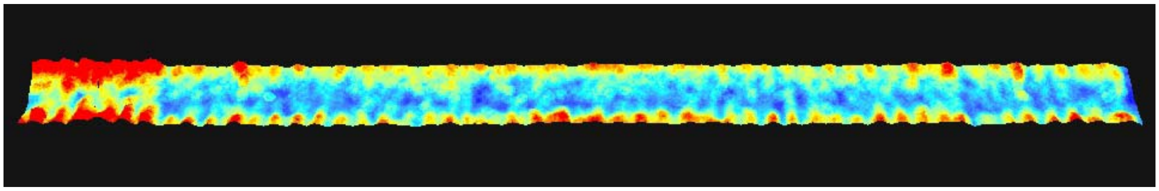

Figure 7: Vandalism – 2001.

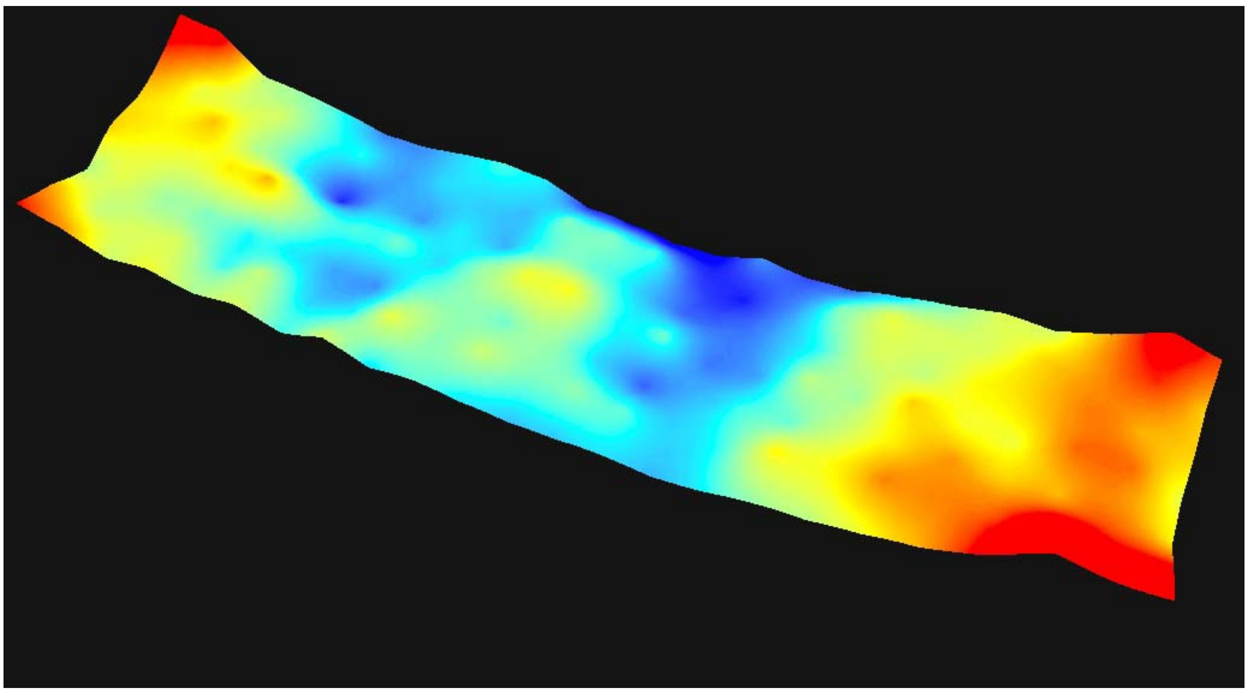

Figure 8: Motor vehicle burglary by weekday – 2002.

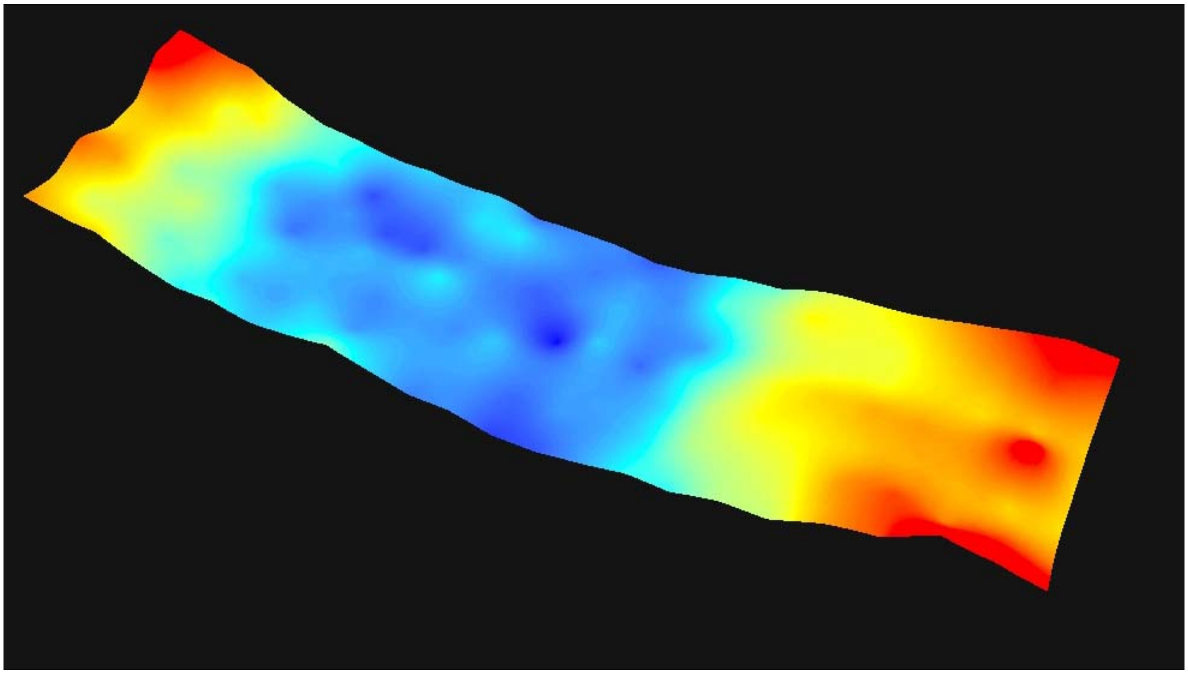

Figure 9: Unauthorized use of a motor vehicle by weekday – 2002.

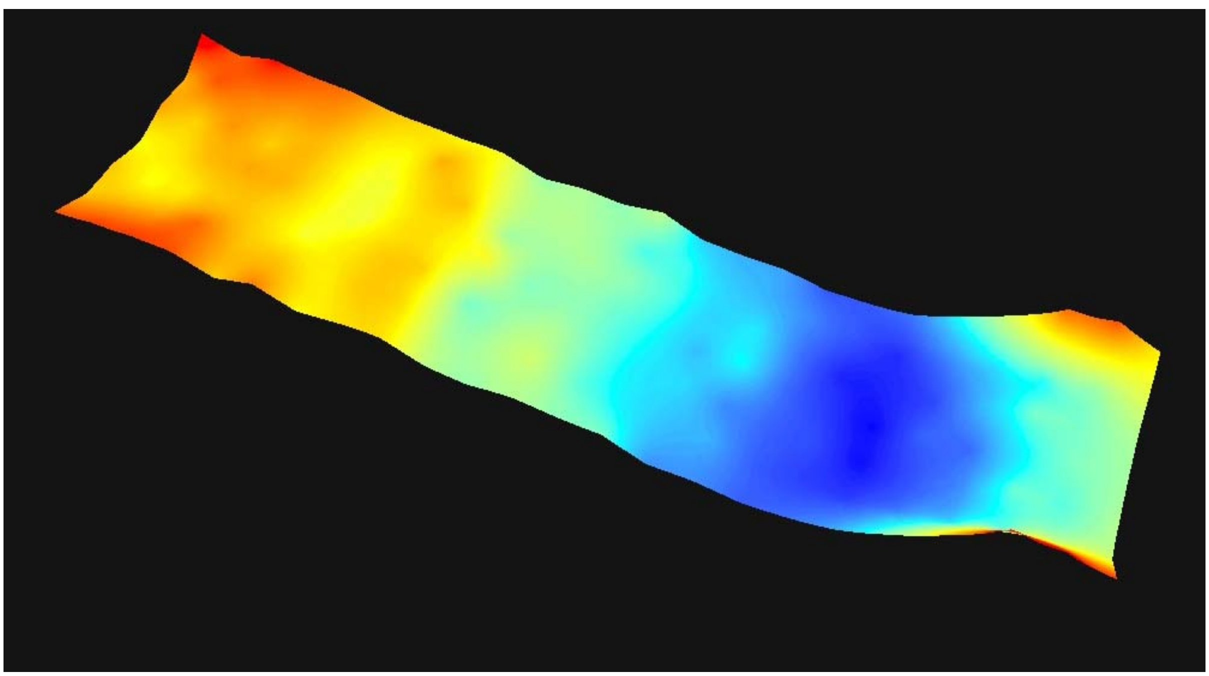

Figure 10: Assault by weekday – 2002.

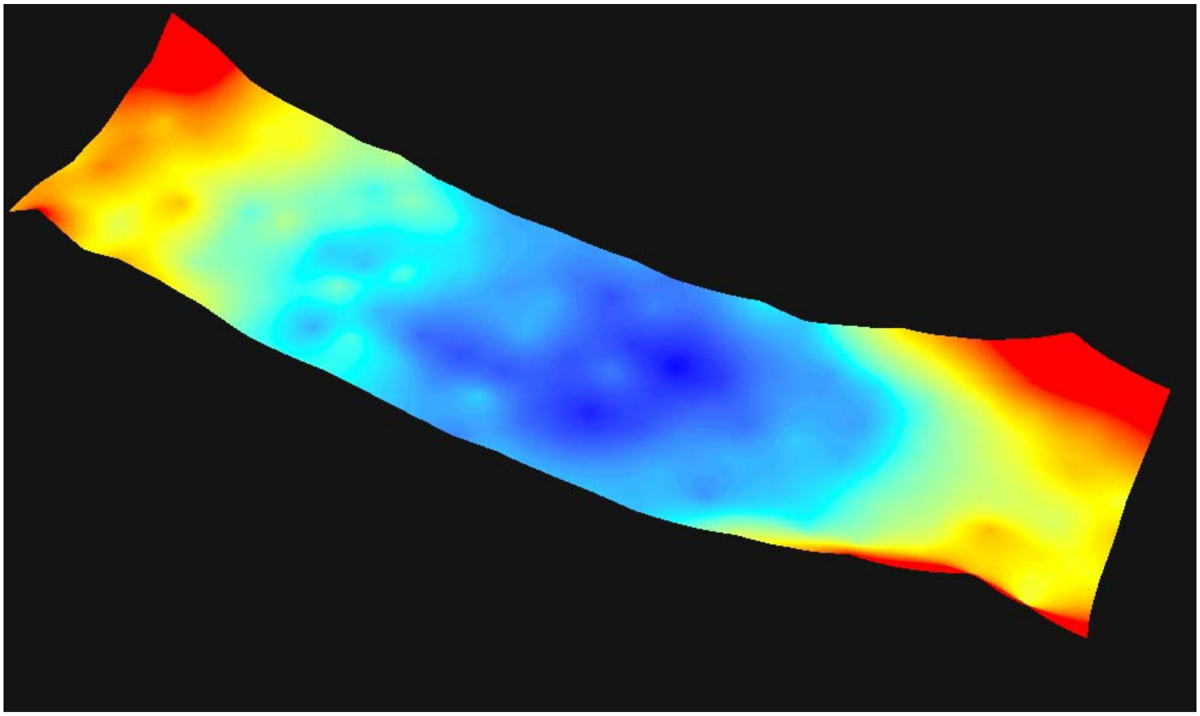

Figure 11: Vandalism by weekday – 2001.

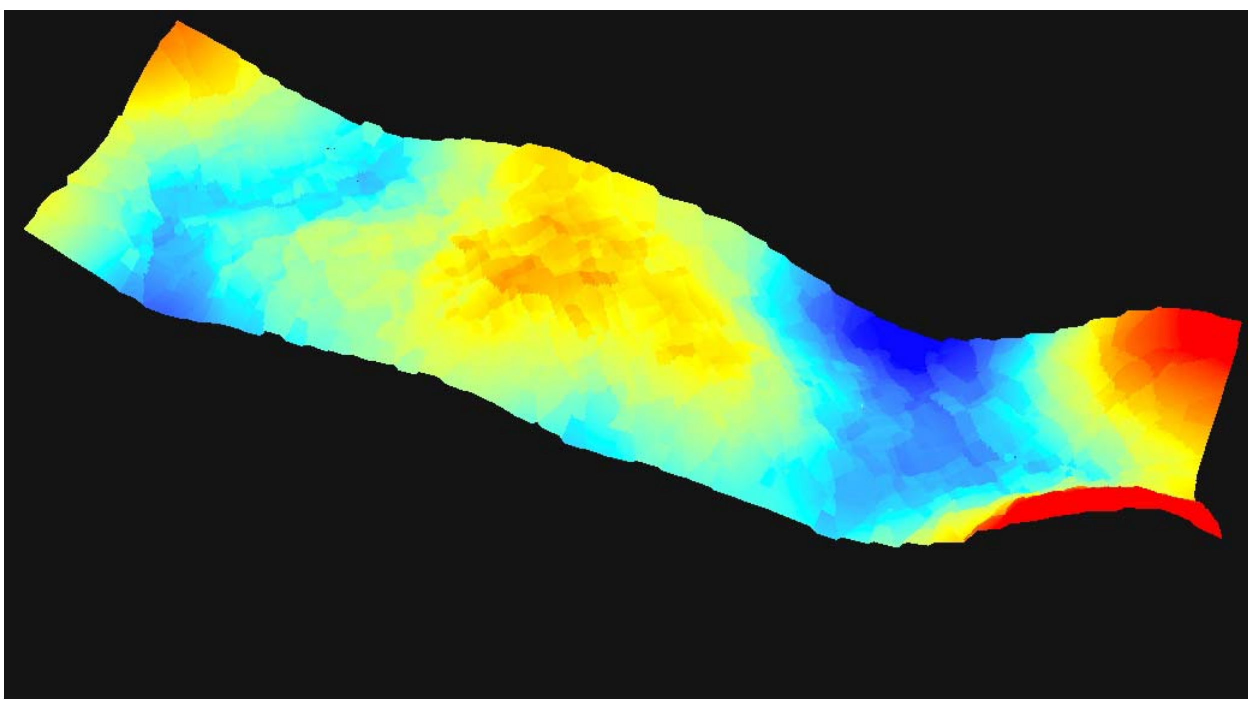

Figure 12: Burglary of a motor vehicle within largest hot spot in Figure 1 by weekday.