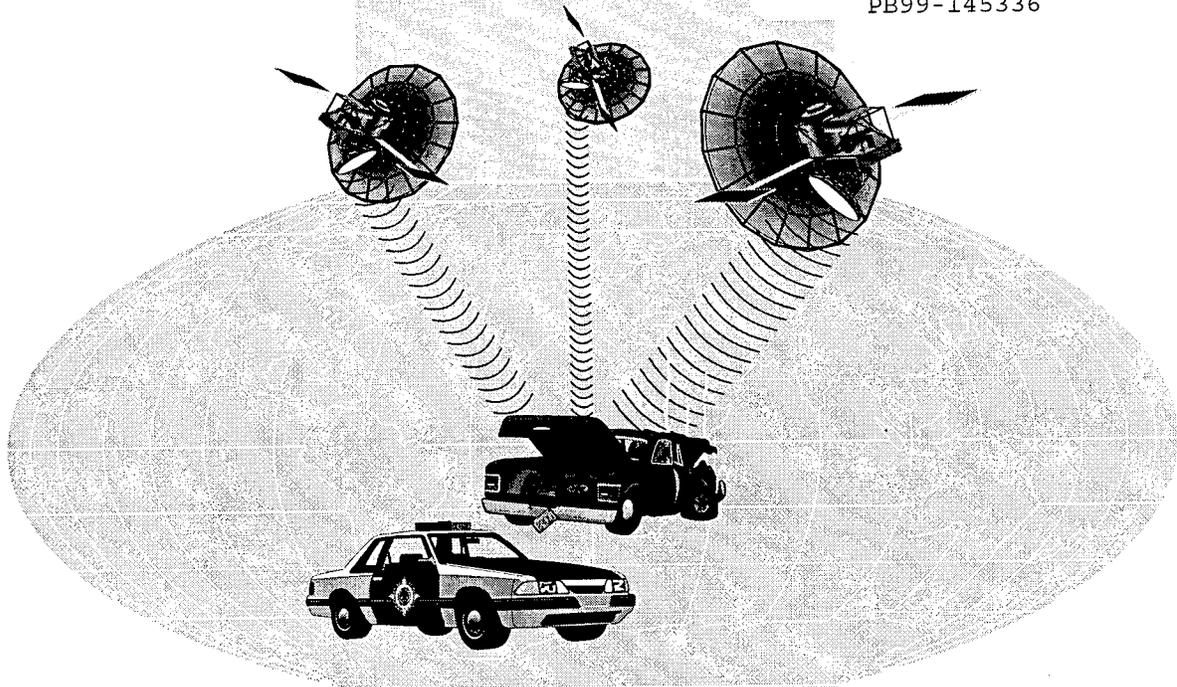


FINAL REPORT

# WHAT VALUE MAY GEOGRAPHIC INFORMATION SYSTEMS ADD TO THE ART OF IDENTIFYING CRASH COUNTERMEASURES?



PB99-145336



JOHN S. MILLER  
Senior Research Scientist

V · I · R · G · I · N · I · A



TRANSPORTATION RESEARCH COUNCIL

REPRODUCED BY  
U.S. DEPARTMENT OF COMMERCE  
NATIONAL TECHNICAL  
INFORMATION SERVICE  
SPRINGFIELD, VA 22161



**Standard Title Page - Report on State Project**

Report No. VTRC 99-R13	Report Date April 1999	No. Pages 40	Type Report: Final Period Covered: Jan. 1998-Jan. 1999	Project No.: 00021552  PB99-145336 GIS, GPS, crash data analysis, accident locations, traffic records, safety
Title and Subtitle: <i>What Value May Geographic Information Systems Add to the Art of Identifying Crash Countermeasures?</i>				
Authors: John S. Miller				
Performing Organization Name and Address:  Virginia Transportation Research Council 530 Edgemont Road Charlottesville, VA 22903				
Sponsoring Agencies' Name and Address  Virginia Department of Motor Vehicles      Virginia Department of Transportation 2200 W. Broad Street                              1401 E. Broad Street Richmond, VA 23220                              Richmond, VA 23219				
Supplementary Notes				
Abstract  Geographic Information Systems (GIS) can be employed to relate, organize, and analyze roadway and crash data, thereby facilitating crash countermeasure identification and evaluation. GIS cannot, however, replace the critical role of the local analyst as a problem solver who still needs to interpret results and recommend engineering, enforcement, or educational improvements.  A literature review illustrates many of the analysis capabilities of GIS. Case studies at the corridor and citywide level illustrate how GIS may be used as an instrument to identify potential crash countermeasures. Although the currently available crash and roadway data are not perfect, they are sufficient for applying some GIS-based analytical techniques. By learning how to use GIS now, analysts can be ready to take advantage of more extensive crash and roadway data sets that may become available in the future.  Using the PC-based Micro Traffic Records System (MTRS), a software package employed by law enforcement agencies in Virginia that records crashes at either a specific intersection or between two cross streets, it was possible to place approximately 82% of the MTRS crash locations within a GIS. Without crashes that were demarcated at "private property" locations, the placement rate climbs to an estimated 94% for intersection locations and 87% for midblock locations.				



**FINAL REPORT**

**WHAT VALUE MAY GEOGRAPHIC INFORMATION SYSTEMS ADD TO THE ART  
OF IDENTIFYING CRASH COUNTERMEASURES?**

**John S. Miller**  
**Senior Research Scientist**

(The opinions, findings, and conclusions expressed in this  
report are those of the author and not necessarily those of  
the sponsoring agencies)

Virginia Transportation Research Council  
(A Cooperative Organization Sponsored Jointly by the  
Virginia Department of Transportation and  
the University of Virginia)

Charlottesville, Virginia

April 1999  
VTRC 99-R13

PROTECTED UNDER INTERNATIONAL COPYRIGHT  
ALL RIGHTS RESERVED.  
NATIONAL TECHNICAL INFORMATION SERVICE  
U.S. DEPARTMENT OF COMMERCE

Copyright 1999 by the Virginia Department of Transportation

NTIS is authorized to reproduce and sell this report. Permission for further reproduction must be obtained from the copyright owner.

## ABSTRACT

Geographic Information Systems (GIS) can be employed to relate, organize, and analyze roadway and crash data, thereby facilitating crash countermeasure identification and evaluation. GIS cannot, however, replace the critical role of the local analyst as a problem solver who still needs to interpret results and recommend engineering, enforcement, or educational improvements.

A literature review illustrates many of the analysis capabilities of GIS. Case studies at the corridor and citywide level illustrate how GIS may be used as an instrument to identify potential crash countermeasures. Although the currently available crash and roadway data are not perfect, they are sufficient for applying some GIS-based analytical techniques. By learning how to use GIS now, analysts can be ready to take advantage of more extensive crash and roadway data sets that may become available in the future.

Using the PC-based Micro Traffic Records System (MTRS), a software package employed by law enforcement agencies in Virginia that records crashes at either a specific intersection or between two cross streets, it was possible to place approximately 82% of the MTRS crash locations within a GIS. Without crashes that were demarcated at "private property" locations, the placement rate climbs to an estimated 94% for intersection locations and 87% for midblock locations.



## FINAL REPORT

# WHAT VALUE MAY GEOGRAPHIC INFORMATION SYSTEMS ADD TO THE ART OF IDENTIFYING CRASH COUNTERMEASURES?

John S. Miller  
Senior Research Scientist

## INTRODUCTION

Considerable resources are being expended to make crash data accessible in some type of Geographic Information System (GIS). These efforts typically include the use of Global Positioning System (GPS) technology to mark the crash location, coupled with the recording of certain crash data elements (such as the number of vehicles involved or the weather conditions at the crash scene). In conjunction with a platform containing roadway elements, such as lane widths or the location of relevant safety hardware, these crash data may be used to identify problem areas and develop potential countermeasures.

It is tempting to state that the reasons for using a GIS are that (1) the structure of the database is documented, (2) all the data can be stored in a single database or successfully linked, (3) the database is user friendly, and (4) limitations of the existing data, such as the accuracy of the crash location methods, are well known. Yet these features are *possible*, to varying degrees, through the use of existing crash data systems that are not conventionally classified as "GIS."

The initial motivation for conducting this research began with a brief question: what advantages, beyond the four cited previously, are gained by analyzing crashes stored within a GIS? Consider, for example, local law enforcement agencies in Virginia, many that use the PC-based MTRS. With this system, if an officer wants to know the top ten high-frequency crash locations in that jurisdiction, he or she can obtain this information from a tabular database such as MTRS without ever using GIS. Therefore, besides being able to see a picture of crashes against a backdrop of roads, what else can we learn from the use of GIS? What additional benefits can GIS provide us with that do not already exist without its use?

## PURPOSE AND SCOPE

The purpose of this study is to ascertain which analytic capabilities of GIS can practically be applied to crash data evaluation. *Practically* means that one must consider the diverse environments in which law enforcement officers obtain crash data. *Crash data evaluation* refers to examining crash data for patterns that would lead to the development of engineering, enforcement, or education-related countermeasures. These crash data might reflect an entire city or county (a macroscopic level) or they might reflect a single corridor or intersection (a microscopic level).

Some persons classify GIS exclusively as an information technology. With this in mind, one gains benefits derived from being able to share data electronically with multiple users across multiple platforms. While this electronic aspect is certainly relevant, it is beyond the scope of this study. The fact that GIS allows linkage based on a common method of geographically locating features, however, is relevant to this work.

## METHODOLOGY

The case study approach was employed, where GIS-based analysis techniques were applied to real motor vehicle crash data. For the geographical area in Central Virginia, the analysis was conducted with two different sets of data, collected primarily during the years 1995-1998 from Virginia's accident report form, the FR-300P. At the macroscopic level, the possibility of simplifying the representation of crash data with respect to specific characteristics was examined for the city of Charlottesville, Virginia—this was the first data set. At the microscopic level, the advantages of being able to see crash locations for a specific corridor in Albemarle County, Virginia, were considered—this was the second data set.

To implement this case study approach in a sensible manner, it was necessary to understand what types of GIS analysis techniques have been practiced elsewhere. Also, it was necessary to acquire crash and roadway data that could be feasibly stored within a GIS. Therefore, four areas were examined in this study:

1. *A literature review of GIS applications that used crash data was conducted.* The review gave an indication of what GIS-based queries might be suitable for the case study areas.
2. *How to place MTRS data within a GIS was determined.* Crash data for the years 1990-1998 were obtained from the Charlottesville's MTRS database and were geocoded in a GIS.
3. *GIS-based queries were used to analyze crash data at the citywide (macroscopic) level.* In addition to the Charlottesville MTRS data, roadway data were obtained from four separate sources: the City of Charlottesville, the Thomas Jefferson Planning District Commission, the Virginia Department of Transportation (VDOT), and a private vendor.
4. *GIS-based queries were used to visualize crash data at the microscopic level.* Virginia FR-300P crash report forms were collected for specific sites of interest from the Albemarle County Police Department. Where possible, sites were visited in person in order to make more accurate recordings of crash locations. These crashes were plotted within a GIS.

## RESULTS AND DISCUSSION

### Task 1 Results: Literature Review of GIS Applications that Use Crash Data

Although a GIS often serves as both a database and a source of maps, these aspects alone do not fully explain its capabilities when used by persons who are trying to quickly understand large amounts of data. Goh captures this best when he states that "...the fundamental difference of a GIS from any other information systems is that it has the knowledge of how events and features are geographically located"<sup>1</sup> (p. 80). That is, there is a geographical relationship between the various types of data that may be incorporated into a database. An example is a common locating system both for intersection-related crashes and traffic signals. This capability of GIS to relate various types of data in a meaningful way becomes important as one moves beyond using a GIS to simply create a pin map of crash locations.

The potential of using GIS to store, query, and analyze crashes as well as their potential root causes has been widely documented. A quick review of the literature suggests several key areas where the use of GIS can help accomplish certain types of analyses. In these articles, simply being able to use a software package that manipulates geographic information is not significant. Instead, the articles illustrate how GIS is an instrument that helps one better understand crash data and use those data to make decisions that potentially will enhance roadway safety. Although there are numerous ways to describe these interdependent analysis methods, it is logical to delineate them into the following categories: simplification of data, creation of collision diagrams, spatial queries, network applications, integration issues, and alternative methods for pinpointing crash locations.

#### **Simplification of Data**

One of the most common uses of GIS is to visually digest a large amount of information quickly, such as a map of high accident crash locations. Another use of GIS, as suggested by Crespo Del Río et al., is to use a graphic that outlines the location and extent of poor quality pavement sections.<sup>2</sup> Mohle and Long demonstrated that the reasons for using a GIS are to create collision diagrams (which simplify the presentation of crash information at a specific site) as well as to "accurately spot accident trends" (p. 29).<sup>3</sup> A North Carolina study illustrated how one of the uses of GIS related to its ability to display crash sites. In this study, a "sliding scale" was used, whereby a segment of a specific length along a roadway was dynamically moved until that segment met a threshold, such as a minimum number of crashes or crashes of a particular severity (p. 469).<sup>4</sup> Thus, the value of this process ensured in a systematic manner that all possible hazardous sites were identified. Conversely, if the segments had been predetermined, it could have been possible that two adjacent segments with crashes close to the common boundary would not have been properly identified as hazardous. The threshold for what constitutes "hazardous" can be varied, which can simplify the presentation of the data for a large area.

Hovenden et al. also have used a GIS to present large amounts of data in a concise manner.<sup>5</sup> Similar to what is conventionally done with origin-destination paths for urban transportation planning, the authors displayed a map where a road segment's crash history was reflected in the width of the segment. Wider segments, of course, implied a greater risk. The paper also outlined the use of GIS to find the "worst" 1.5-km section of a road by dynamically moving startpoints along a route. This usage of GIS was similar to the North Carolina research in that it also avoided problems that might have been masked through the arbitrary definition of roadway segments.

In 1992 the Georgia Institute of Technology developed a prototype GIS for transportation that included an accident records component.<sup>6</sup> At least three key types of queries were outlined that could not have been done with a conventional link-node crash records system. The first of these queries involved a map that described crashes by some type of severity category. The advantage of this usage of GIS was that it provided a graphical representation of crash locations. However, the authors acknowledged that a listing of high-accident intersections was available from conventional means. The Georgia Tech study also outlined two other benefits that go beyond the presentation of data—spatial analysis and collision diagrams. Spatial queries, such as the ability to find all crashes within a specified distance of an intersection, became easier to accomplish using GIS. One can quickly envision why such a capability might be useful. For example, in some cases, the geometrics of this situation might have been that all crashes that occurred within 100 m of an intersection were intersection-related, whereas in another situation it might have been that one was searching only for those crashes that occurred within 25 m of the intersection.

### **Creation of Collision Diagrams**

The ability to easily reproduce a detailed collision diagram was offered by the Georgia Tech study. When data were available, one could "zoom" to the location of interest to view crashes in relation to the roadway geometry and safety hardware. A problem statement submitted to the AASHTO Standing Committee on Research noted that, at the microscopic level, "Collision diagrams offer one of the few means by which designers today display crash history graphically."<sup>7</sup> The committee's narrative pointed out that while the combined capabilities of CADD and GIS are increasingly offering the potential to use visual information, the challenge is to present meaningful interpretations rather than overwhelming the user. This suggests a microscopic component to GIS capabilities that may be applicable to presenting meaningful interpretations that do not overwhelm the user, depending on the precision of identifying both the crash locations and the roadway geometry features against which crashes are assessed.

One can expand this current potential of GIS to reproduce detailed collision diagrams to future uses of GIS as we consider what defines the crash location, which is a key element of any detailed collision diagram. The U.S. Department of Transportation's Guideline for Minimum Uniform Crash Criteria (MUCC) defines numerous data elements that should be collected at the scene of the crash. Within this Guideline (and with respect to the data element C5), it is stated that the crash roadway location is the "Exact location on the roadway indicating where the crash occurred"<sup>8</sup> (p. 23). For example, in some cases a law enforcement officer might define element

C5 as where the vehicle came to rest after the crash occurred, whereas another officer might define this as the point where a vehicle began a skid or left the roadway. Although such disparities might not be significant at the macroscopic level, it could be of benefit to the person studying the crash history at a specific site, such as at an intersection. In this sense, providing the officer's collision diagram in a GIS would allow the analyst, if necessary, to ensure that all crash locations were represented uniformly. An example of this uniformity would be using the first point of impact consistently as the crash location.

## **Spatial Queries**

The task of studying crashes in a GIS may be represented as a spatial analysis problem. In other disciplines, studying spatial relationships among data is a frequent occurrence. For example, one school district compared the centroids of building permits to the centroids of the locations of school students in order to discern demographic trends.<sup>9</sup> With crash data, a possible extension could be to observe the movement of crash locations as a function of time. The North Carolina DOT reported two chief safety-related benefits of GIS—(1) the integration of data from other sources that can be facilitated by GIS's geographical linkage capabilities and (2) the results of queries can be viewed in a spatial format rather than only in a tabular format.<sup>10</sup> This report indicated that “. . . it is hard to estimate what impact either of these advantages will have on conduction [of] accident analysis” (p. 5)—as pointed out from the website. The report also stated that one of these applications of GIS was to select route segments that had high concentrations of truck crashes. GIS graphical representations of crashes in these segments could then facilitate study of these representations in greater detail.

A safety-related short course implied two key benefits that may be gleaned from a visual representation of crash locations.<sup>11</sup> The first is an understanding of any clustering of high accident locations (HALs). For example, one may determine whether several locations are in geographic proximity to each other, thereby facilitating specific countermeasures, such as selective enforcement or reduced speed limits. The second benefit is subtler, yet of equal importance. Visual patterns may be used to discern geographic relationships based on select variables, such as the driver's age. For example, no discernable patterns may be apparent when looking at the crash locations. However, after limiting examination to those crashes that occurred Friday and Saturday evenings between 9 P.M. and 6 A.M. that involve drivers under 24 years of age for instance, certain types of problem areas could be identified. An extension of this enforcement-based approach would be to include potentially relevant geometric characteristics, such as short yellow signal timings in a progression of streets. While these concepts were developed without GIS per se in mind, it appears that the lessons regarding spatial analysis are certainly transferable to GIS.

While acknowledging that identification of high-accident sites can be accomplished with GIS, Austin et al. bluntly state that other types of inquiry make better use of GIS's potential.<sup>12</sup> Two aspects of GIS usage were proposed that go beyond data presentation per se. This first aspect is an error-checking scheme. The features coded by the officer can be compared to the features stored in the roadway database. For example, if the officer codes the speed limit of a route on a crash report that is different from what is recorded in the roadway database, this

clearly indicates that there is a discrepancy. The second aspect of GIS usage is to identify high accident regions or zones, as opposed to identifying only specific intersections or segments. This allows the analyst to categorize areas by land use and compare how they affect the number of crashes. Two spatial capabilities were also suggested. One concerns the safety of children walking to school, including selection of sites within proximity of a neighborhood school and the evaluation of the safety of routes where children walk. The second application was risk analysis based on where persons live (such as the number of injuries per 1,000 persons in population).

Panchanathan and Faghri concurred with Austin et al.'s perspective on the additional potential uses of GIS. They believed that the key advantages of GIS included capabilities to do spatial and network analysis and integration of data. One example of this additional potential of GIS was the use of buffers to capture items within a certain distance of one another, such as at-grade rail crossings within a quarter mile of another at-grade rail crossing.<sup>13</sup> Affum and Taylor exemplified the use of buffers in their consideration of identifying hazardous locations based on land use. For example, they explained that accidents involving children were always found within 1 km (0.6 mi) of schools.<sup>14</sup> They also outlined a ready-to-use GIS application that integrated some traditional methods of data reduction. For example, they used an automated display of high-accident locations in addition to using the spatial query capabilities of GIS.

Identifying traffic crashes that may have been caused by an earlier incident is a specific analysis capability that is greatly facilitated by the use of GIS, according to Raub.<sup>15</sup> By examining crashes that meet both a distance criterion of being within 1600 m of an event—as well as a time criterion of occurring within 15 minutes of the event—one can determine whether two crashes are causally linked. Clearly, this type of spatial analysis query would be restricted with an older system that did not have flexible geographic capabilities. However, such a query could be performed with an older system if only the events along a specified route are being examined.

One can go beyond visual inspection of crash locations in order to add rigor to how trends suggested by a map of crash locations are assessed. Choi and Park suggested a couple of simple yet innovative methods for looking at crash locations.<sup>16</sup> The first method was to divide the study area into a grid. For these grid cells, the number of crashes could be regressed to the segment length and/or number of intersections. Provided that a statistically sound relationship and high  $R^2$  value existed, a useful extension of this analysis would be to identify grids where the number of predicted crashes was lower than the actual crashes. The authors also computed a “coefficient of localization” where it could be ascertained whether some type or category of crash tended to be spread throughout the study area or was concentrated in specific locations. (This coefficient essentially stratified crashes by zone and type and used a method of computation very similar to that in determining a chi-square score.) Again, grid cells could have been used to accomplish this analysis. This usage of grid cells provides a mathematical way of quantifying trends that might not be directly observable.

Kim et al. stated that GIS could be used as an exploratory tool, especially for identifying patterns in crashes.<sup>17</sup> Examples were given where a spatial analysis could have been quickly accomplished in the GIS. Techniques for this type of analysis included assessing how spatial crash patterns vary by time of day, day of week, injury level, and seeing how crash location

patterns changed when the selection criteria are varied, based on factors such as speed or alcohol use.

## **Network Applications**

GIS can facilitate analysis for network routing applications as well as for obtaining characteristics about specific routes that have already been selected. For example, Souleyrette and Sathisan used GIS to characterize the risk of certain routes that were used for moving “high-level radioactive material.”<sup>18</sup> Estimates of resident population, visitor population, and ecologically sensitive areas (such as wildlife refuges, wilderness areas, and water surfaces) were obtained within a specified distance from the route that was used for shipments. The authors observed that these data, all taken from disparate sources such as census databases (for residential population) and commercial information (e.g., hotel locations to give visitor estimates) could then be used as inputs for various risk estimation models. Thus, while GIS was not used for all of the computations, it facilitated the application of software already having that specialty. Patel and Horowitz used the capabilities of GIS in a different manner. They selected the “best” route that should have been taken regarding the shipping of hazardous materials.<sup>19</sup> These authors discussed how to select a route that minimized risk based on population and environmental considerations, such as wind direction, should a spill occur. This procedure was facilitated by having network capabilities (e.g., finding the shortest path) and supplemental data (e.g., population data) integrated into the same platform.

Further evidence of the networking capabilities was offered by Mefford. He used GIS to graphically display the “shortest and safest” bicycle routes from suburban areas to the Central Business District (CBD). While this type of analysis could have been done without GIS, a key feature of this graphic capability was the ability to envision the impacts of improvements quickly. For example, a high-risk link could have been redesigned. Subsequently, one could observe whether this improvement affected which route was the least risky for cyclists.<sup>20</sup>

Finally, Austin et al. suggested a pedestrian-oriented application that combined the network, display, and integration capabilities of GIS.<sup>21</sup> A paper survey of routes that children used when they walked to school was coded in a GIS and then checked against accident rates of specific street segments in order to identify where school crossing guards should have been located. Parents were then asked to identify dangerous locations, which the authors then compared with locations having the worst crash history. Although the authors acknowledged that exposure limitations were a problem, since not all routes had equal numbers of pedestrians using them, it was possible to compare what parents thought were the worst locations to those locations where the greatest number of crashes occurred. These network, display, and data integration capabilities can thus facilitate public outreach efforts by better educating parents and school personnel as to which locations should be avoided by children without adult supervision.

## Integration Issues

Although not directly focused on analysis capabilities, some articles from the literature review articulated why it is important to assess GIS capabilities before moving forward with its implementation. In a pilot effort by FHWA where laptop computers, GPS receivers, and GIS software were used to record crash locations, one result was that officials learned about the overall capabilities and needs of other agencies with whom they were cooperating.<sup>22</sup> The understanding of how multiple agencies function is especially crucial when considering the diversity of those individuals and agencies involved in highway safety analysis. The crash data are collected, stored, and analyzed by persons with different missions, even if these individuals are in the same agency. Since GIS data collection requires a substantial investment, it is useful for persons to know in advance what the return on that investment will be. In this case, it was beneficial for both data collectors and analysts to know what a GIS can accomplish that could not be achieved without such a system. In short, integration of agencies' missions should be considered.

One of the key benefits of GIS, as outlined in the study by the Georgia Institute of Technology, is the potential of GIS to enhance data integration. This study correlated crash rates with poorly maintained roadways in terms of signs and safety hardware. Not only is a pictorial representation of crashes a benefit, but additionally, the capability to relate these data to other data sets (such as in maintenance information) is of interest. This data integration has value where data sets are being updated. For example, an agency or corporation that maintains a list of utility pole locations has information that would be of interest to the person who keeps the roadway database current.

Aultman-Hall and Hall illustrated this capability of database integration when they estimated crash risk exposure for bicyclists.<sup>23</sup> The authors first surveyed cyclists regarding their previous crashes as well as their commute routes, then they digitized the cyclists' routes onto coverages of a roadway network. Finally, the authors related these routes to infrastructure information stored within the GIS coverage. This coverage included such information as distance or road type (paved, unpaved, cut-through, sidewalk, and so forth). Another example of linking crash data to pavement data was presented graphically by Siegel and Yang. They visualized the number of crashes at specific locations, in contrast to pavement conditions at the same locations.<sup>24</sup> Although Saccomanno et al. discussed the use of GIS for accident risk modeling, it is clear from their work that a key contribution of a GIS is that it can possess the capability to link three disparate databases: roadway geometrical data, traffic volumes, and police accident report data.<sup>25</sup> The importance of traffic volumes as a normalizing factor was emphasized again by Affum and Taylor, who indicated that failure to incorporate these volumes could result in too much attention being given to high volume roads.<sup>26</sup> While there are cases where these databases may already be linked through prior planning and interagency coordination, clearly the use of GIS gives one the flexibility to integrate information when such planning has not taken place or when an unforeseen need for a certain type of data arises.

In summarizing NCHRP Project 20-27(2), Opiela pointed out that GIS could be used to enhance integration of data from different sources, especially when technological limitations would otherwise hinder the transfer of data between agencies or even functional units within the

same agency.<sup>27</sup> This ability of GIS to enhance integration of data from different sources, in spite of technological barriers, is potentially relevant to crash data analysis, where it is probable that the roadway network, the location of safety hardware, and crash data will come from diverse sources. A practical application of this data integration would be where the state department of transportation maintains guardrail location data, but where updated subdivision location information comes from another source, such as the county planning department or private sector.

Lamm et al. illustrated how GIS integration capabilities were fundamental to a project even when the focus of that project was not GIS.<sup>28</sup> In their work, design elements for various roadway sections were assessed using rather complex relationships. For example, in one type of analysis, the accident rate for various sections was regressed to the operating speed and degree of road curvature. The suitability of GIS in this case was that roadway inventory data could be rapidly incorporated. Thus, the focus of their work was not on GIS, but instead, on a type of analysis for which GIS provided a convenient platform.

### **Alternative Methods for Pinpointing the Crash Location**

Of course, a benefit related to GIS integration issues is the locating of crashes themselves with other technologies, such as GPS receivers. This refers to defining the spot where the individual crash occurs, and not the various analysis methods of groups of crashes that have been discussed previously.

While the location accuracy of data from disparate sources will not be perfect, it may often be good enough to facilitate analysis using GIS. For example, the North Carolina Department of Transportation found that GIS and GPS coordinates of intersections had an average difference of approximately 6.3 m, slightly more than an average car length.<sup>29</sup> For a macroscopic level analysis, this implies sufficient accuracy to locate crashes. However, for creation of a detailed collision diagram, this may not be sufficient. Even without technologies such as GPS, GIS can often be used to pull crash locations from another source provided there is a common method of locating the crashes. This does not always work perfectly, since it must be ensured that both data sources store crash locations in a common manner or appropriate adjustments must be made. In a 1994 study at Kansas State University, researchers were able to locate approximately 44% of the intersection-related crashes when converting from a legacy system to a GIS for a single county. A large portion of the problem was that the naming convention for intersection locations was sensitive to the order. For example, the authors noted that a crash at 8<sup>th</sup> & 10<sup>th</sup> streets was not the same as a crash at 10<sup>th</sup> and 8<sup>th</sup> streets.<sup>30</sup>

Work completed by Timmons Associates in November 1997 for VDOT included geocoding Albemarle County (but not the City of Charlottesville) crash data based on their locations as recorded in VDOT's Highway Traffic Records Information System (HTRIS).<sup>31</sup> That effort resulted in 760 out of 1633 Albemarle County crashes successfully being located for a single year where routes were defined within the street coverage. Appendices C and D of the consultant's report indicate that a higher match rate could be achieved if additional time is spent making the GIS roadway data more compatible with HTRIS (or vice-versa). Steps that need to

be taken include adding nodes, ensuring that each GIS roadway segment corresponds to exactly one HTRIS link, and ensuring that all data sets are updated. This consultant also pointed out several features available for analysis within a GIS, including spatial and logical queries and the ability to integrate data from different sources. Examples given in the sample application included being able to select all crashes that occurred during rush hour, displaying of crashes against traffic volumes on the roadway, selection of crashes where alcohol or bad weather was a factor, and network applications for routing oversized vehicles.

### **Lessons Learned from the Literature Review**

Not all of the benefits cited in GIS-related articles should necessarily be ascribed solely to GIS. For example, it has been stated that presentation tools such as pie charts of crashes by type are useful. These tools may be easier to use with some of the GIS software packages. However, this use of presentation tools may not necessarily be exclusive to GIS. Link-node based crash data can just as easily be exported to a worksheet to create similar charts. However, since GIS software packages are constantly being upgraded, they are more likely to have many of the newer presentation capabilities.

There is a degree of overlap among the benefits of GIS as a tool to study crash countermeasures. For example, being able to see the data visually may for one analyst be useful as a means of understanding a large amount of information, whereas another analyst might use that same GIS capability to discern trends that otherwise would not be apparent. Initially, these GIS applications focused on a graphical display of crash locations; later, these GIS applications became integrated with statistical techniques. It is important to highlight that in many instances the value of the GIS analysis is not only its ability to provide visual representation, but additionally, it is GIS's ability to either organize the data in a different manner than has been done previously or to integrate the crash data with another data set from a different source.

While the literature illustrates what GIS can help accomplish, it is of interest to us to assess the feasibility of using GIS to identify potential problem areas and suitable crash countermeasures in a given location with Virginia data.

### **Task 2 Results: Determining How to Place MTRS Data Within a GIS**

In order for the crash locations to have any meaning within a GIS, these locations must be referenced in some way to the roadway data. There are several options for accomplishing this task, each option having its advantages and drawbacks. One of the most flexible methods is to obtain at each crash site the latitude and longitude coordinates with a GPS receiver. These coordinates can be directly placed in the GIS if the coverage is stored in decimal degrees. Conversely, these coordinates also may be easily converted to the appropriate cartographic projection. There is a data collection cost associated with the use of GPS receivers. While that cost should decrease dramatically as the technology improves, this cost reduction still does not solve the problem of missing an accurate location for historical data. This problem has led

agencies to look for additional methods for referencing their crash locations to the roadway network.

One approach is simply to denote crashes at the closest street intersection. The City of La Mesa, California found that approximately 80% of the city's crashes were intersection-related. For these crashes, this city felt it was acceptable to denote the crash location as the nearest street intersection.<sup>3</sup> While this method may not differentiate between crashes that are, say, 15 m to 31 m (50 ft to 100 ft) from an intersection, this method can be semi-automated fairly easily, quickly giving the analyst a rough pictorial representation of the crash history throughout the area, as well as enhancing some of GIS's data manipulation capabilities.

For crashes that are not intersection-related, there are several options. The narrative on the crash report form can be used to pinpoint the crash location; a table of aliases that associates specific locations with an XY coordinate can be created; or a street segment can be defined as a route feature whereby a point at a fixed distance along the route can be located. An example of this would be taking the point at 2 km from the start of Rt. 1. (Methods for locating crashes are discussed in the literature; for example, O'Neill and Harper outlined an algorithm appropriate for covering street segments from a route-based system to a GIS platform.<sup>32</sup>) Such approaches require a working knowledge of how a route-based system, such as HTRIS, stores location data. Alternatively, if data are stored in a CADD format, known locations on the CADD drawing may be associated with real-world coordinates, as pointed out by Pawlovich and Souleyrette.<sup>33</sup>

The City of Charlottesville uses an MTRS program for storing, querying, and reporting crash results. At one point, MTRS was in use in over 100 jurisdictions within Virginia as well as in other states. One reason we are interested in this software is that in Virginia incorporated cities such as Charlottesville maintain their own streets. Although a computer system managed by VDOT, in conjunction with the Virginia Department of Motor Vehicles (VDMV) and the Virginia State Police, stores all reportable crashes in the Commonwealth, this system currently contains only some (but not all) of the streets that are not maintained by VDOT. Instead, these streets are sometimes contained in systems that are maintained by incorporated cities. Hence, it is possible that (barring a change in practice) MTRS data may represent the only source of roadway-specific location data should a crash occur on a city street that is not stored within VDOT's system.

MTRS crash locations are stored in one of two possible formats. One option is that the crash will simply be defined as "intersection-related." This means that a crash is denoted as having occurred at 5<sup>th</sup> Street and Main Street. The second option is that a crash may be stored as a midblock location. This midblock location is essentially a street segment between two cross streets. An example would be a crash on Main Street between 5<sup>th</sup> and 6<sup>th</sup> streets.

Geocoding the first type of crash is a relatively simple process, although it can be somewhat tedious. A summary of the method that we employed was to use the MTRS street.dat file to create a lookup table between segment numbers and street names in MTRS. Then, we imported the MTRS data into ESRI's PC ArcView and geocoded the intersection crashes directly. Although the initial results were not favorable, we found that by associating some common places not found in the street network (e.g., Barracks Road Shopping Center) with an

XY coordinate, we were able to successfully match 82% of the intersection-related crashes. Therefore, of the 7,868 intersection crashes, 6,441 could be matched. One of the problems with this method was that some of the MTRS links were simply defined as “private property.” Had these crashes not been coded in this manner, the success rate for intersection crashes would have been estimated to be 94%. Note that if a landmark, such as a shopping center, is too large to be useful as a single location, then one can split this landmark into multiple locations, such as the north, south, east, and west quadrants of the shopping center. In short, the accuracy of the crash location in the GIS is only as good as its source data – in this case the MTRS database.

Midblock locations, the second type of crash, were more complicated. For each crash, two intersections had to be geocoded. The most workable solution was to average the XY locations of the intersections and use that average as the midblock location. This method allowed rough identification of the crash locations for macroscopic types of analysis. However, accuracy was not sufficient for a collision diagram representation. For these data, out of the 5,330 midblock locations, a total of 4,328 (81%) could be located; this is a merging of the error rates from geocoding the neighboring left and right intersections. Had there been no crash locations designated as “private property,” we estimate the success rate would have been approximately 87%.

Therefore, out of the 13,198 crashes for the Charlottesville area between 1989 and part of 1998, approximately 82% could be located with either an intersection or a midblock location. Because other localities may be interested, a procedure that does not require much programming is shown in the Appendix. However, many geocoding options are available.

### **Task 3 Results: Using GIS-Based Queries to Analyze Crash Data at the Macroscopic Level**

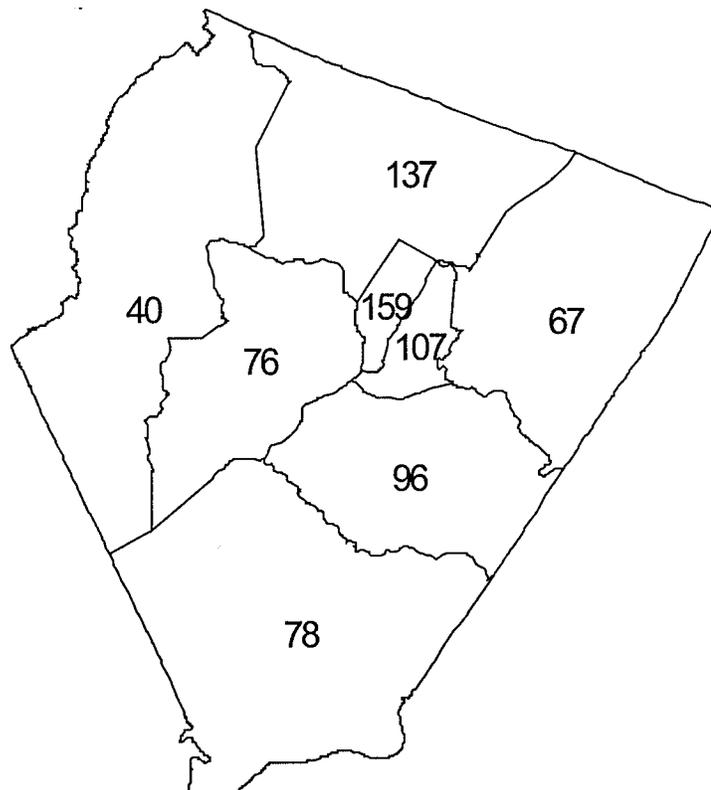
To investigate how well spatial queries may be employed at a citywide (macroscopic) level, three types of sample spatial queries are shown. Next, a demonstration of how to use GIS as part of a more comprehensive analysis of crash data is presented. The reason for this approach is that the sample spatial queries demonstrate how one may apply techniques directly from the literature. These techniques are not tailored to the needs of the study area; however, they do illustrate the salient features of GIS. The section titled “analysis of crash data at the macroscopic level,” however, is a more realistic application. This section uses GIS—not for the sake of using GIS—but instead, as a single tool that is part of a comprehensive safety study.

Crash data were obtained directly from the local police departments, as they could be geocoded using the methods outlined in the previous section. An alternative, of course, would have been to use VDOT’s HTRIS data. However, the geocoding rate was higher with the MTRS data and the precision, though not perfect, was sufficient enough to allow for some demonstrations of GIS applications.

## Sample Spatial Queries

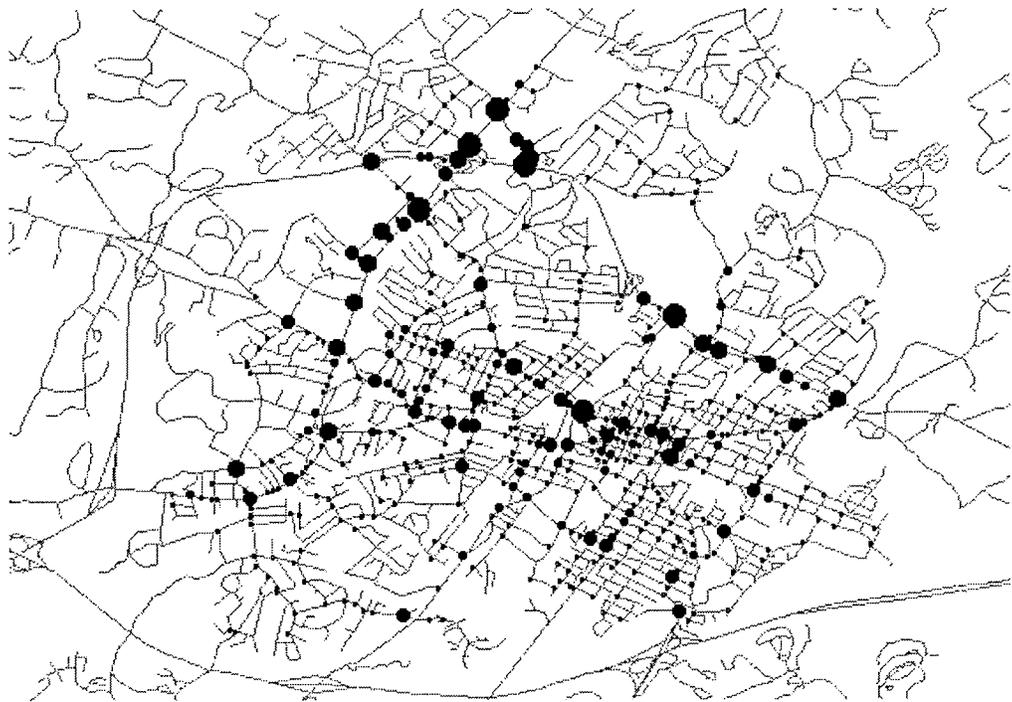
Three example applications illustrate types of GIS queries: data simplification, simple spatial analysis, and grid-based spatial analysis.

Figures 1 and 2 illustrate how GIS may be used to simplify large amounts of data. Figure 1 summarizes Albemarle County crashes by sector, whereas Figure 2 highlights Charlottesville intersection crashes. Since some intersection crashes in Figure 2 are overlaid exactly on top of one another, we need to have a means of seeing how many crashes occurred at the exact same location. A similar application done with Albemarle County crashes was slightly different in that the crashes were not overlaid, but were still in proximity of one another (e.g., crash 1 might be 70 m from a landmark and crash 2 might be 90 m from the same landmark). Thus, we were able to select all crashes that were within a specified distance (50 m) of an intersection and

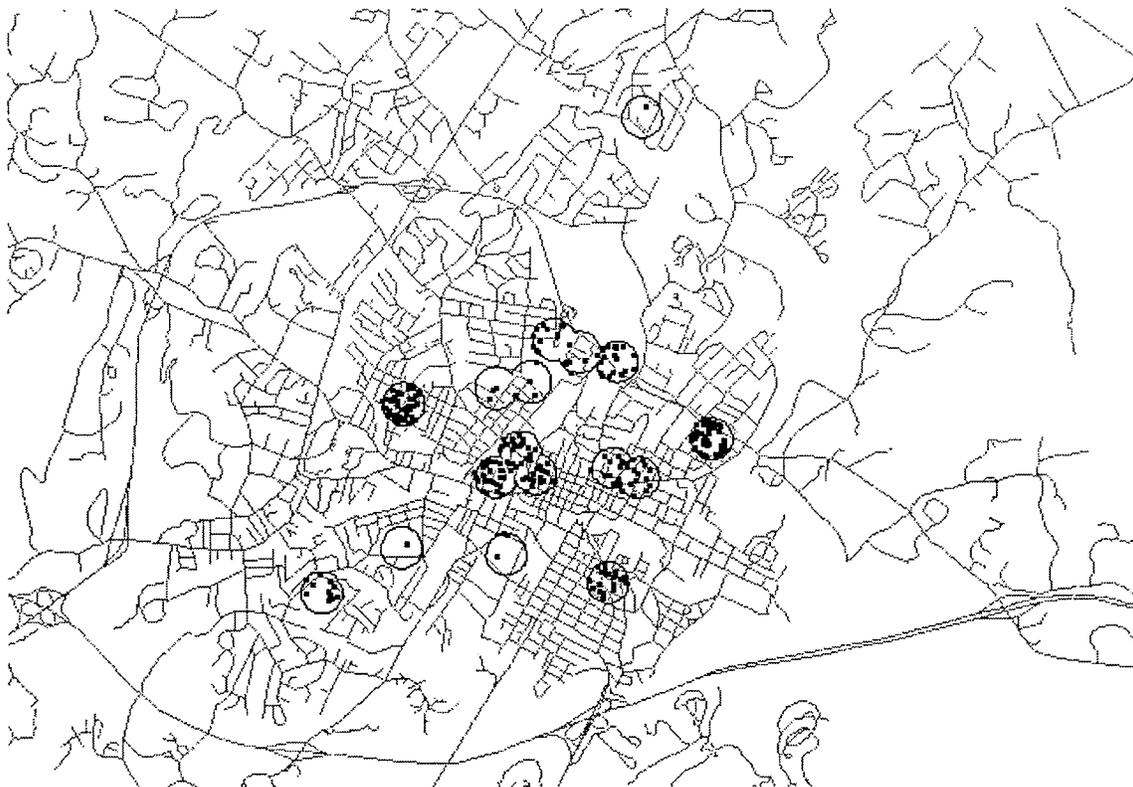


**Figure 1. Albemarle County Crashes by Sector, Sample Application**

represent them accordingly. The Charlottesville crash locations were geocoded by the author from the MTRS data obtained by the City of Charlottesville Police Department. The Albemarle crash locations were obtained from Timmons Associates, who had already geocoded the locations using VDOT HTRIS data. The Albemarle County sector map was obtained from the Albemarle County Police Department. While these applications may facilitate the presentation of data, they do not fully harness the analysis potential of GIS.



**Figure 2. Charlottesville Intersection Crash Location, 1989-1998**



**Figure 3. Charlottesville Crashes within 200 m of Schools**

The second type of application, shown in Figure 3, is a simple spatial analysis. Crashes in this figure occurred within 1/8th of a mile of schools (200 m). In the first draft of this figure, darker colors implied higher crash densities for certain spots. We can normalize these crash numbers by the size of the buffer area; however, if two schools were close together, we might have a total area larger than 1/8th of a mile (200 m). Areas that continued to have a high number of crashes, even when accounting for differently sized geographical areas, may have exhibited problem locations. Additionally, instead of using colors, the number of dots are proportionate to the number of crashes divided by the area.

The third type of application was based on a suggestion from Choi involving the use of grids. Each grid cell indicates the total length of street miles within the grid. We also can obtain the total number of crashes within each grid. If we look at a picture of only the number of crashes, then we can obtain one perspective of where the problem areas were located. However, we can also look at a picture that accounts for the number of street miles. In this case, crashes that occurred in 1994 or later, divided by the number of street miles, are shown in Figure 4; darker shading indicates more crashes per street mile. Note, of course, that one could normalize by using other variables, such as population density or vehicle miles traveled. For future work, this grid-based procedure offers a quick way to develop some statistical applications, such as studying which variables have the greatest influence on crash severity in urban areas.

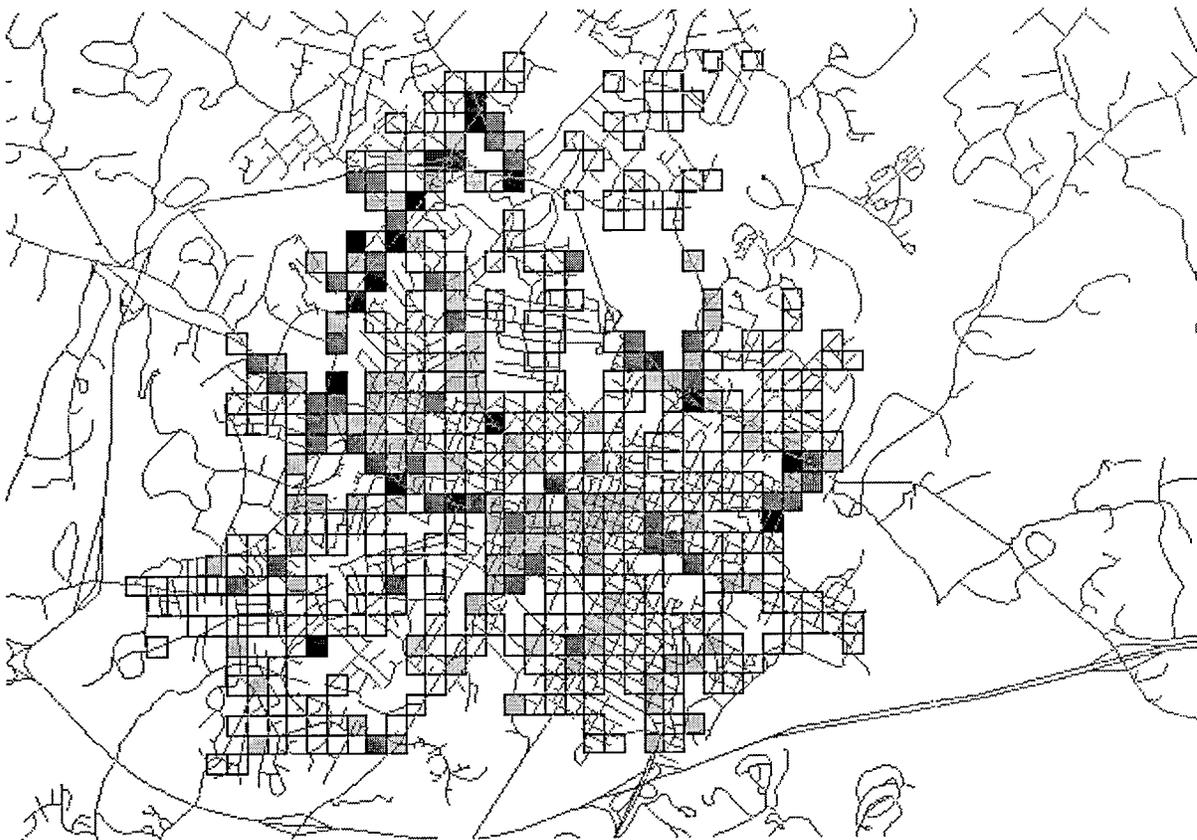


Figure 4. Grid-Based Theme: Crashes Divided by Street Miles

## Analysis of Crash Data at the Macroscopic Level

It is a normal practice to consider problem areas by identifying high crash locations based on count or severity. What can GIS add to this analysis?

One way to consider these data is to focus on alcohol and driver error. The decision was initially made to look at alcohol-related crashes, since alcohol involvement is viewed as a factor that can be addressed through public education and enforcement and is therefore within the purview of law enforcement agencies. *This reflects a fundamental concept behind crash records, i.e., that the ultimate goal is not to conduct analysis, but instead, to take actions that will reduce crash frequency or severity.* Consequently, one could, for example, focus on geometric design aspects of two-lane roads if the analysis was being done for traffic engineers, since construction and maintenance are within the purview of the engineers.

The analysis was ultimately expanded to include driver error, in part because it was something officers could control, but also, because initial analysis of the alcohol data showed that alcohol crashes were less of a problem than driver-error crashes for this specific area and time period. This reflects a second key element, i.e., that there will be occasions where *data analysis suggests that the problems being studied initially are not as severe as subsequent problems that are discovered.*

As the next two subsections illustrate, a temporal analysis is still useful for understanding the magnitude of the problem and is a necessary precursor to the third subsection, which uses a spatial analysis to find patterns within the crash locations.

### *Temporal Analysis of Crashes that Involve Alcohol*

Key variables are defined for Virginia's FR-300P crash report form, as reported by MTRS for the City of Charlottesville. Alcohol involvement was given for each crash as one of seven possible scores that were evaluated by the officer: 0 was unknown, 1 was no alcohol involvement, 2 was drunk, 3 was drinking but impaired, 4 was drinking but not impaired, 5 was drinking but unknown, and 6 was other. Driver action had 37 possible codes. Notably, 1 indicated no driver error, whereas 2 through 37 indicated some type of error, such as excessive speed. Day of week was also available (e.g., 1 for Monday, 2 for Tuesday, etc.) as was the two-hour block during which the crash occurred (1 was for a crash between Midnight and 2 A.M., 2 was for a crash between 2 A.M. and 4 A.M., etc.) Note that there could be some variation to these definitions by jurisdiction. While these jurisdictions all used the same FR-300P, these jurisdictions may elect to record information in different schemes within MTRS.

To accomplish the analysis, intersection and midblock crash locations were first stratified by alcohol involvement as well as by weekend, time of day, and year. Previous observations of an adjacent law enforcement agency suggested three shifts: midnight until approximately 7 A.M., 7 A.M. until 3:30 P.M., and 3:30 P.M. until midnight. Therefore, crash times were blocked into these three categories as well as by day of the week (Monday through Thursday, Friday, or Saturday and Sunday). Alcohol involvement of some type (codes 2-6) was classified against no

alcohol involvement. Alcohol-related crashes for the past three years were found to be distributed as shown in Table 1.

**Table 1: Crashes with Alcohol Involvement, 1995 –1998\***

	Mon.	Tues.	Wed.	Thurs.	Fri.	Sat.	Sun.	Total
Midnight – 8 A.M.	8	5	9	16	9	38	24	109
8 A.M. – 4 P.M.	3	1		3	4	8	5	24
4 P.M. – Midnight	12	11	8	17	25	28	15	116
Total	23	17	17	36	38	74	44	249

\*Tables 1 and 3 covered the time period from 1995 through part of 1998.

Table 1 suggests that Saturday had a large proportion of alcohol-related crashes compared to other days of the week, whereas the 4 P.M. to Midnight and the Midnight to 8 A.M. periods accounted for a similar numbers of crashes. If individual cells are examined, it can be seen that two sets of consecutive periods (Friday night to Saturday morning and Saturday night to Sunday morning) accounted for the four largest cells of crashes. A similar trend was observed in Table 2 if one examines crashes from 1989 to 1994. The consecutive weekend periods again had a large number of crashes.

**Table 2: Crashes with Alcohol Involvement, 1989 – 1994**

	Mon.	Tues.	Wed.	Thurs.	Fri.	Sat.	Sun.	Total
Unknown			1		1	1		3
Midnight–8 A.M.	4	16	20	11	18	42	56	167
8 A.M.–4 P.M.	5	8	4	6	13	16	10	62
4 P.M.–Midnight	29	15	15	35	48	59	19	220
Total	38	39	40	52	80	118	85	452

A superficial examination such as this leaves out the possibility that there might be other factors that explain why the dual Friday and Saturday time periods have so many alcohol-related crashes. For example, it may simply be the case that more crashes occurred during that time period regardless of alcohol involvement. Table 3, which shows crashes that occurred between 1995 and 1998 where alcohol was not involved, suggests that this was not the case. We can see that the weekday period from 8 A.M. to 4 P.M. represented the time of greater crash concentrations. Smaller blocks of time could be selected should this be necessary—for example, the two-hour period when the most weekday crashes occur, which was from 4 P.M. to 6 P.M.

**Table 3: Crashes without Alcohol Involvement, 1995 –1998\***

	Unknown	Mon	Tues	Wed	Thurs	Fri	Sat	Sun	Total
Unknown	5	5	3	2	2	4	2	1	24
Midnight–8 A.M.		40	35	43	43	68	69	46	344
8 A.M.–4 P.M.		339	324	286	333	416	261	147	2106
4 P.M.–Midnight		219	201	204	226	303	164	116	1433
Total	5	603	563	535	604	791	496	310	3907

It is not necessary to have a GIS in order to do this type of reporting; any type of database that allows queries is acceptable for a temporal analysis. In fact, many persons would argue that a crash records database is not needed in order to pinpoint Friday and Saturday evenings as a problem time period for alcohol involvement in crashes. This does confirm, however, that if one is interested in reducing alcohol-related crashes, then Friday and Saturday evenings are appropriate time periods for targeting special enforcement activities.

*Temporal Analysis of Crashes that Involve Driver Error*

The data in Table 3 suggest that midday crashes (in shaded area), in addition to those that are alcohol-related, need to be considered. Yet, it is interesting to note that most crashes listed some type of driver error. Of those crashes that occurred between 1995 and 1998, only 14% did not have some type of driver error associated with them. It would be of interest in future work to determine if the crashes that did not have driver error had some common characteristics.

Additional analysis not presented here illustrates the types of driver actions involved in crashes, such as “disregarded stoplight” or “followed too close.” Initially we considered targeting specific types of violations. However, since the presence of an officer was required to issue a citation regardless of the violation type, we simply focused on driver errors. Future work could, however, associate specific types of violations with specific types of countermeasures. For example, if a certain area experiences a high number of failures to stop at a red light, this might suggest the need for technologies that focus on red light running as opposed to having officers checking for speed violations.

Based on these temporal analyses, one can consider two courses of action:

- Setting up sobriety checkpoints or increased patrols for the two weekend periods (8 P.M. Friday to 4 A.M. Saturday and then 8 P.M. Saturday to 4 A.M. Sunday). This would address alcohol-related crashes through targeting high-crash areas.
- Setting up speed checkpoints for select weekday shifts (8 A.M. to 4 P.M.). This would address crashes that are caused by driver error.

### *Using GIS to Consider Selective Enforcement Countermeasures*

Where might officers establish these checkpoints or patrols? This requires a spatial analysis, where GIS can be applied.

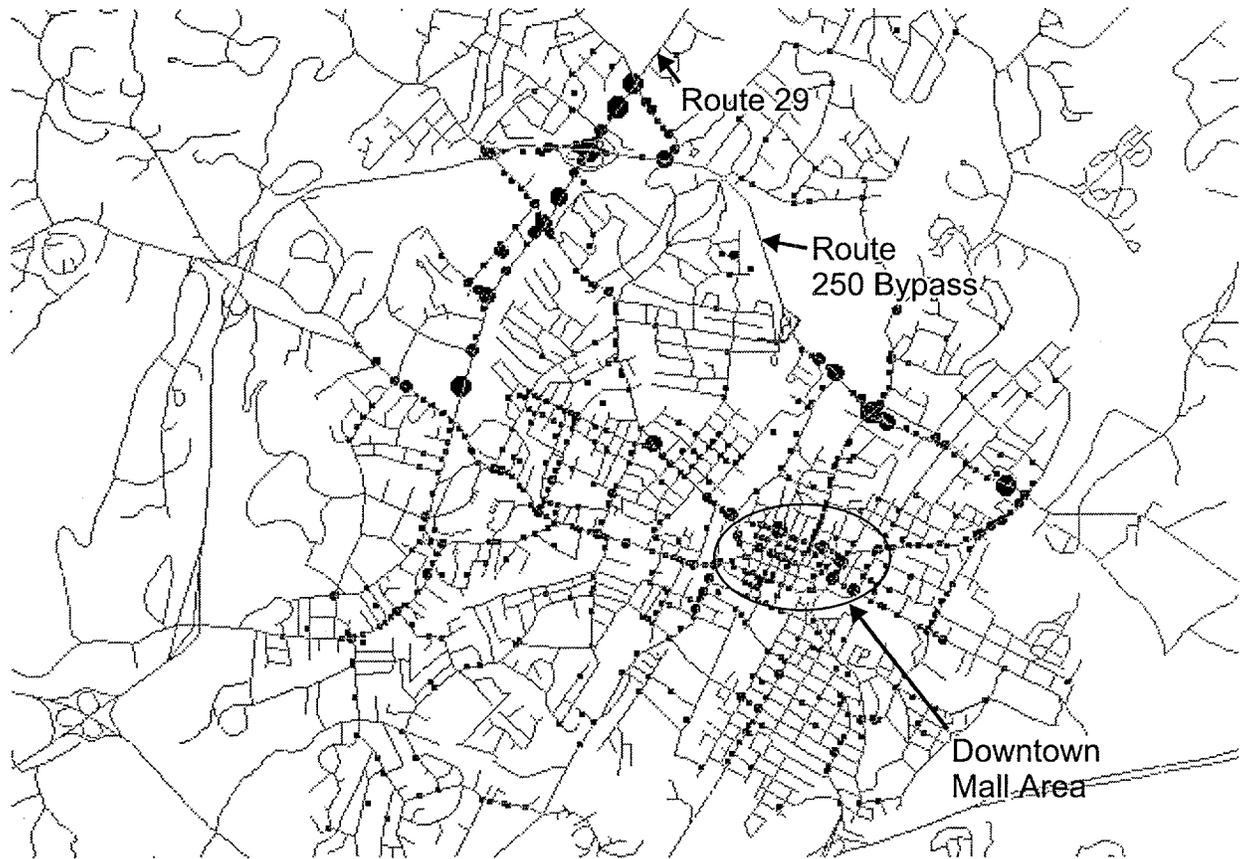
Sites grouped by the number of alcohol-related crashes over the period 1995 to 1998 could be identified. Figure 5 shows the crashes that occurred Friday night, Saturday morning, Saturday night, or Sunday morning, with the largest circles indicating 3 crashes with alcohol involvement. Based on the crash history, a Friday and Saturday night patrol would be beneficial nearer the center of the map; the northern boundary of town had fewer alcohol-related crashes for this time period.



**Figure 5. Alcohol Related Crashes Friday and Saturday Evenings, 8 P.M. to 4 A.M.**

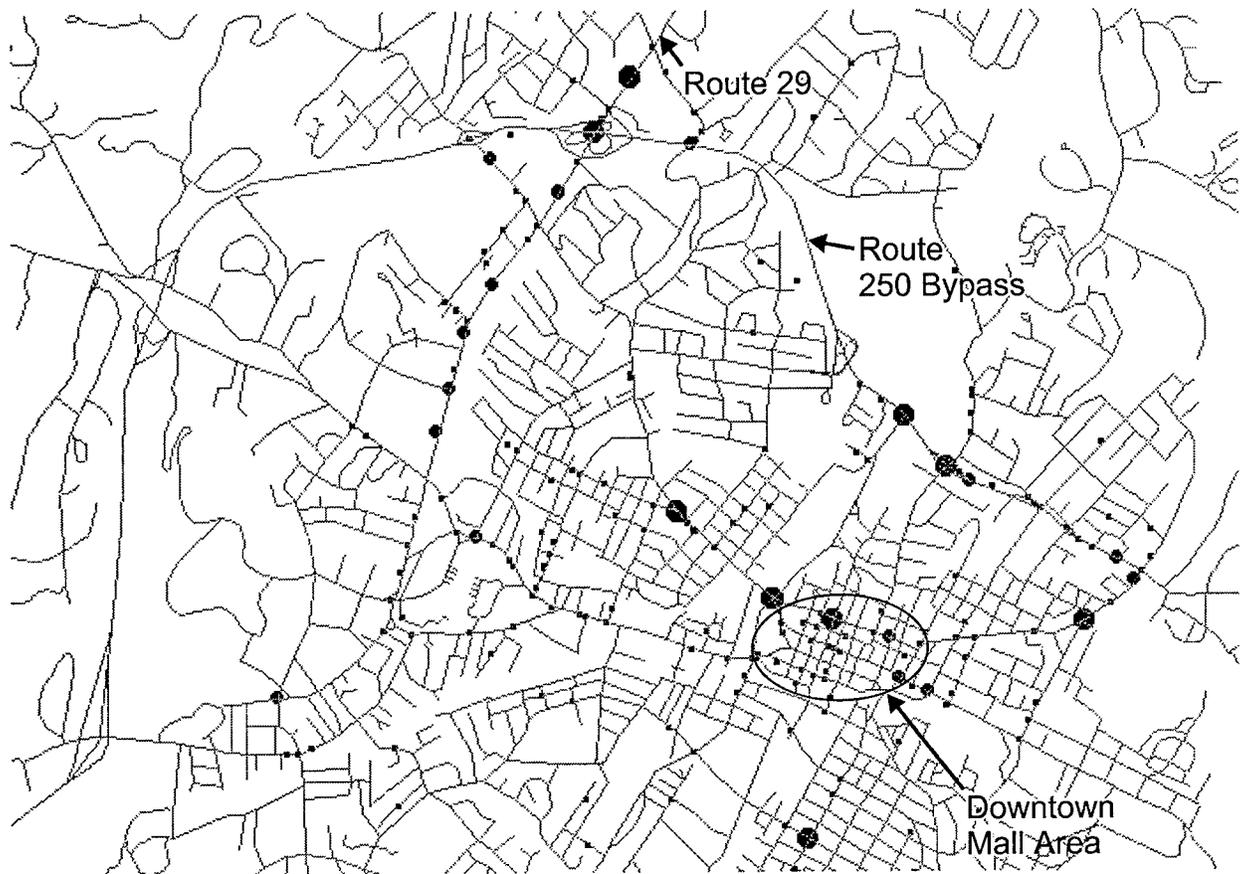
The weakness of these sobriety checkpoints is that they target a small number of crashes. We can expand the role of selective enforcement to include crashes that involve driver error on weekdays between 8 A.M. and 4 P.M., which often corresponds to an officer's day shift. We are free to include alcohol-related crashes or leave them out. There were a total of 1709 crashes during this period, 11 of which had some type of alcohol involvement, as denoted by the FR-

300P variable “alcohol involvement.” Figure 6 depicts a map of all crashes during this period. Although common knowledge suggests that the main divided arterial serving the town (Route 29) would have numerous crashes, it is also apparent that large numbers of crashes occurred in the areas of the Downtown Mall and the eastern edge of the Route 250 Bypass, which is a limited access highway.



**Figure 6. Weekly Crashes Between 8 A.M. and 4 P.M.**

How weekday crash locations vary by location and time is of interest. Figure 7 shows that the Downtown Mall area was about as problematic as the Route 29 area between 10 A.M. and noon. However, by the evening rush hour between 4 P.M. and 6 P.M., Route 29 and the Route 250 Bypass had become more problematic. In establishing a patrol, it would therefore be prudent to use this information to determine how many officers should cover certain areas—depending on historic crash rates by time of day and location.

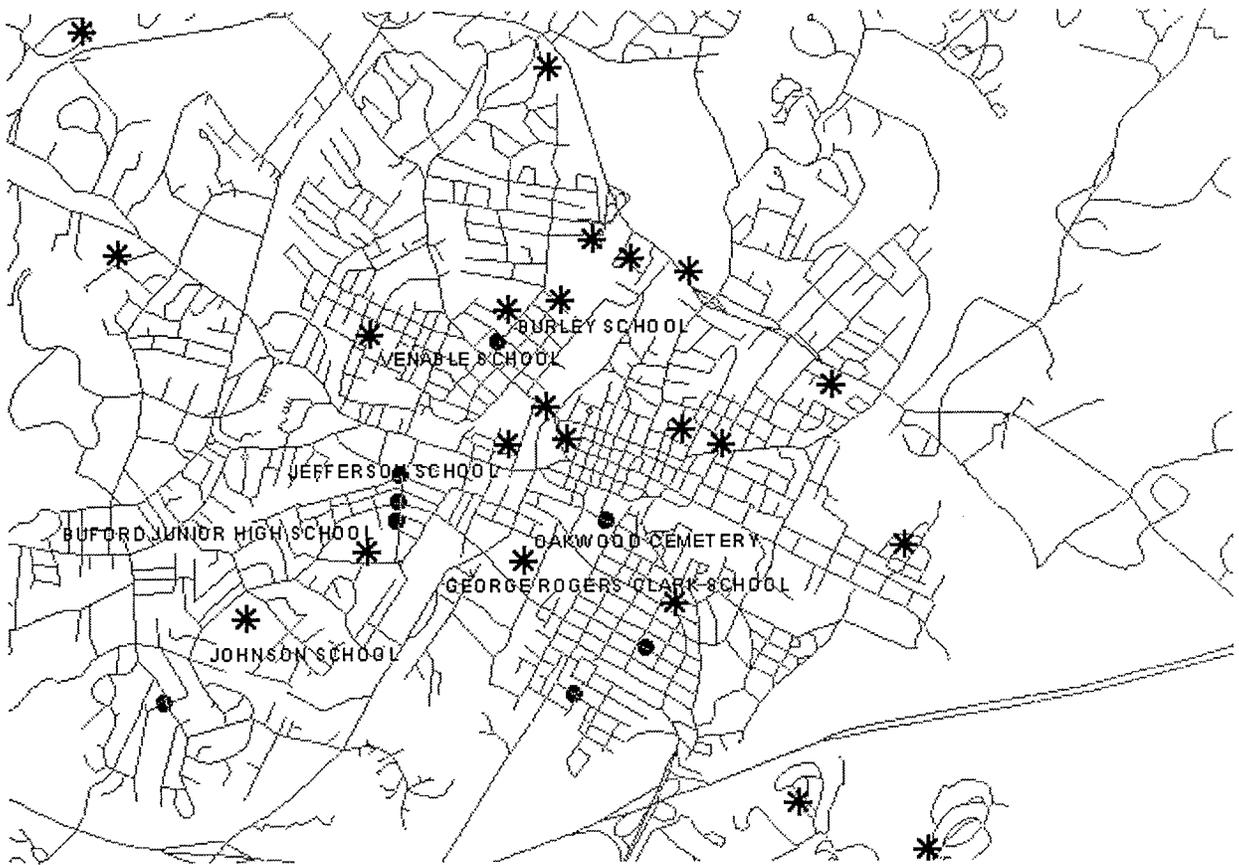


**Figure 7. Weekday Crashes between 10 A.M. and Noon**

### **Using GIS to Consider Other Criteria: Injuries, Pedestrian Involvement, and Proximity to Schools**

Crash injuries, and not just numbers, can also be evaluated. During the 10 A.M. to noon rush hour, these crash injuries highlighted the problem of some areas on the eastern side of the city. This could suggest that EMS access time might be worth assessing. Other data can be incorporated into this evaluation as well, such as crash and land use type. For example, knowing pedestrian crashes in the vicinity of schools during times that children are walking to and from school might be of interest in this evaluation process.

As an illustration, crashes between 6 A.M. and 10 A.M. as well as between 2 P.M. and 4 P.M. that involved pedestrians may be selected. However, only 23 crashes fit this criterion, with only 8 involving pedestrians under the age of 18. A map of these 8 crashes, as shown in Figure 8, may be related to a map of schools and other institutions in the area and suggests a few areas where one might target engineering efforts or attention from school guards. Especially troubling are the 3 pedestrian crashes in close proximity to Buford Middle School (formerly Buford Junior High) and Jefferson School. Since no roadway defects were noted, this suggests a site that might be worth examining in closer detail.



**Figure 8. School-Hour Crashes with Pedestrians under Age 18**

One problem with only examining crash sites is that there lacks a standard against which to compare these sites. For example, from 1995 on there were 96 pedestrian crashes in the Charlottesville area. While one can readily ascertain whether a site was lit or unlit, one does not know from that information alone whether having lighting at all sites would have reduced the numbers of crashes in which pedestrians were involved. We can, however, ascertain potential problem areas by looking at a map of all pedestrian crashes, about one third of which did not occur during daylight. Lighting is not necessarily a problem, for in some instances the officer had coded the location as being dark but lit. This suggests, however, that a visit to these locations should be conducted in order to determine whether improving the lighting situation is a worthwhile endeavor. Similarly, a grouping of a few pedestrian crashes in one area (e.g., southwest of the Downtown mall) suggests a site visit might be worthwhile.

*Linkage with Other Data Sets*

The inclusion of school locations highlights an important feature of GIS—being able to link data sets from disparate sources. Other types of data that could be employed to refine these analyses include utility pole locations (to verify lighting concerns), transportation planning data

(to identify areas of high vehicle or pedestrian traffic), and census data (to anticipate high growth areas where traffic safety may become a pivotal issue, such as the opening of establishments that would attract children).

### **Summary of Macroscopic Applications in the Charlottesville Area**

The types of queries shown in this study regarding the Charlottesville data are concerned primarily with the display capabilities of GIS as well as simple spatial queries, including those queries that took advantage of geography (e.g., all crashes that occurred within a certain distance of a landmark) and those queries that benefited from other data sets (e.g., school locations). Conversely, the grid-based applications suggest further potential for using GIS capabilities to categorize data in a manner that would be very difficult with only a link-node system.

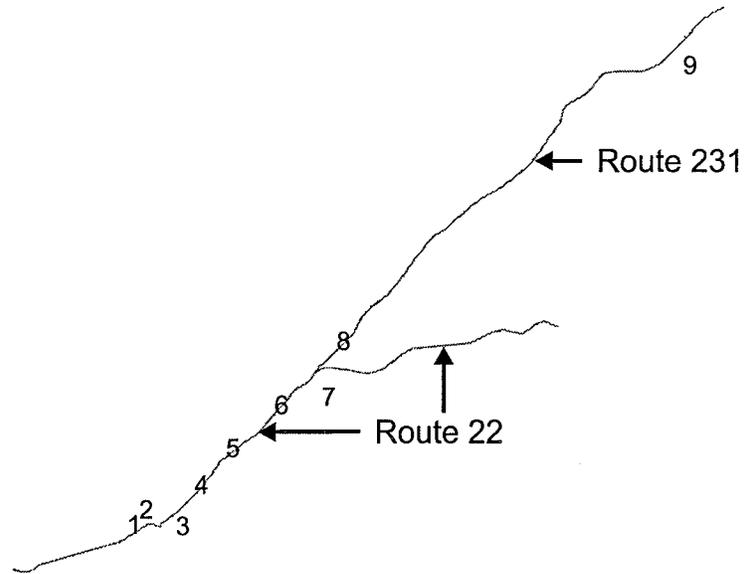
The salient feature of these examinations is that GIS can facilitate data analysis and help the user ask questions. However, GIS does not replace some of the basic problem solving that must be undertaken. GIS will not, for example, replace the need for the law enforcement officer, administrator, or traffic engineer to make an estimate as to how much resources should be directed toward preventing one type of crash over another.

Note that these figures are for comparison purposes only. The symbols can be adjusted to reflect the number of crashes. Furthermore, these data reflect only the level of precision available from the FR-300P report form: the two hour window, for example, may not be precise enough if one is trying to target morning crashes between 8:30 A.M. and 9:30 A.M.

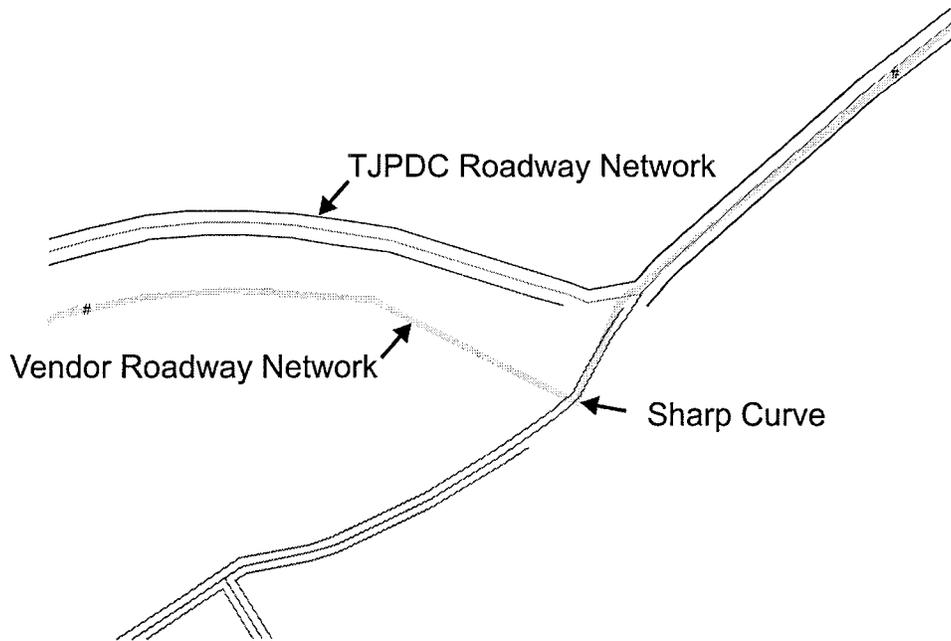
### **Task 4 Results: Use GIS-Based Queries to Visualize Crash Data at the Microscopic Level**

At the microscopic level, the question arises as to whether GIS can be useful for analyzing a particular corridor. Problematic sites in Albemarle County were identified for further study through conversations with that county's Police Department. Figure 9 shows nine possible sites along Routes 22 and 231. These are two-lane roads with many no-passing zones, where each "site" has at least four crashes within 500 ft of one another. (The latter portion of the Appendix indicates how these crashes were defined as a route theme within a GIS and how the site visits were conducted.)

Key variables may be considered, such as traffic control, weather and surface conditions, roadway defects, lighting, and collision type. As an example, site 2 was selected, where seven crashes had occurred: three involved injuries and five occurred during good weather; no roadway defects were cited except for "slick pavement." All of these crashes occurred during daylight. Five of the crashes had been rear-end collisions. One of these was a sideswipe crash, and another involved a fixed object. Examination of Figure 10, however, suggests that a sharp curvature in the roadway may have been a contributing factor, since several crashes occurred near the curve. Two of the crashes at the site showed that the officer had reported the driver's vision as being obscured.



**Figure 9. Problematic Crash Sites Along Route 22 and Route 231**



**Figure 10. Problems at Site 2 Suggesting a Sharp Curve**

Figure 10 is a reminder that there will be some error inherent in the roadway network. Two different sources of roadway data are shown in the figure. The Thomas Jefferson Planning District Commission data showed actual lanes, whereas data from a private vendor indicated

centerlines only. Acquisition of another data set from VDOT still did not overlay perfectly with these data sets. Additionally, two differentially corrected positions obtained with GPS receivers were approximately 18 m (60 ft) from where one would have expected them to appear relative to the roadway. This disparity could be significant if one is evaluating sites where such an offset would affect how one viewed crash causal factors, such as the location of a driveway in proximity to an intersection. In short, there are two sources of error that should be kept in mind—the roadway network itself, and the positioning of the crash on that network.

This type of query also permits one to identify problems that might not become evident through examination of reports alone. For example, at site 4, two of the five crash reports showed the driver's vision was obscured as a result of the hill crest. This implies, but certainly does not prove, that the crest might be worth further examination as a potential causal factor for the remaining crashes.

In addition to looking at problematic sites, one can also ascertain whether a visual inspection of key characteristics provides additional insight into the problem of crash causation. In this case, one might consider the problem of crashes hitting a fixed object off the side of the road. Here we see that slightly less than 1/3rd of the crashes (42 out of 142 records) involved a fixed object. The sections of road where these fixed-object collisions occurred may be grouped and on a trial basis the worst sections of road can be selected for studying the possibility of enlarging the clear zone if necessary. A more refined example arises when one restricts the query to sideswipe collisions; eight of the crashes fit this category. Further examination suggests a specific spot that may be targeted for an assessment of the roadway alignment.

It should be noted that not all queries will prove fruitful. For example, the author had hypothesized that there might be a pattern between surface defects and crash locations. In this case, we queried for all crashes along the two routes that involved surfaces that were defective. This particular query did not show a strong relationship between defective road surfaces and crash locations. There were either too many other variables or not enough exposure data to make use of that information for this small sample.

To further consider what the data suggest, photographs of sites 2 and 4 were taken in conjunction with an evaluation of the narrative associated with the crashes at the two sites. At site 2, it was strongly indicated that crashes were occurring when a vehicle would rear-end another vehicle waiting to make a turn from Route 22 on to another two-lane road (Route 731). In some instances, the turning vehicle had made an abrupt change in speed to make the turn. Complicating matters was the proximity of a nearby post office, where in at least in one case a driver argued that a tractor-trailer had blocked his view of the roadway. At site 4, four of the five crashes were rear-end collisions where the following vehicle struck the leader vehicle that was either turning or stopped to make a turn. Photos 1 and 2 suggests that site 2 suffers from a sharp horizontal curvature in the proximity of the post office. At site 4, the problem appeared to be more severe than what would be gleaned simply by looking at the cluster of crashes in the GIS. As shown in the photographs in Photos 3 and 4, there are two vertical curves where the sight distance is limited. In conjunction with high speed traffic and close vehicle following distances, these two vertical curves appear to be a reasonable factor in explaining most of the crashes at the site. In the short term, possible remedies include selective enforcement and

guardrail installation in selected portions. In the long term, the site might be a candidate for a vertical realignment; however, these two other options may be more effective and hence should be considered first. A resident of the area noted that there had been many more crashes than the five studied by the author. While these crashes may predate the 1995-1997 study period, they also were probably not reported, since the resident noted that many of the crashes involved a single vehicle running into a tree or fence with the driver then leaving the scene of the accident. In both these cases, views of the accident clusters using GIS software suggested spots that merited further study. However, use of GIS in this instance did not replace the necessary site visits.

A limitation to this microscopic analysis was the small sample size of reportable crashes. This is one problem that GIS cannot resolve—the fact that, according to the resident, many crashes had occurred at the site but had gone undetected. These unreported crashes complicated the analysis, but a site visit could clarify the situation.



**Photo 1. Site 2, Looking South**

## **EMERGING THEMES**

Several themes can be drawn from the process of conducting this research, the results that came out of this research, and consideration of future challenges that would arise when trying to apply this type of research on a larger scale in Virginia.



Photo 2. Site 2, Looking North



Photo 3. Site 4, Looking South



Photo 4. Site 4, Looking North

### **Multiple Sources of Useful Data**

Crash, roadway, topographic, population, and utility information may all be relevant for crash analysis, especially if spatial queries are considered. The fact that many data sources exist suggests that not everyone has the most up-to-date information. Depending on the application, this may not be a critical issue. The problem arises, though, when one entity purchases or makes substantial improvements to a data source when it was possible to obtain this information at a lower cost elsewhere. To avoid this problem in the future, especially as data become available, it would be in Virginia's collective interest to maintain a good information exchange among the affected agencies, such as local governments, VDOT, EMS, and state and local law enforcement.

### **Rudimentary Types of Analysis Currently Possible.**

The use of converted MTRS data in itself is most likely not going to provide crash locations that are accurate enough for a collision diagram. However, these same locations can be brought into a GIS format with minimal effort. It took approximately 40 hours with the Charlottesville MTRS data to achieve a match rate of approximately 82% once the procedure was understood and software difficulties were eliminated. With training and practice, the amount of time for geocoding may be reduced much further. Eliminating the "private property" convention as a crash location increases the match rate. As data sets improve, various types of microscopic analysis will become more feasible. It remains to be seen whether these detailed analyses will help reduce the overall crash rates.

## **Organizational Barriers to Implementation**

The proliferation of data sources and crash location technologies has led to growth in the number of organizations conducting pilot projects that use GIS or GPS. While this increased availability of data and technology represents an opportunity for better crash analysis, it also places a burden on practitioners, forcing a resource tradeoff between training and daily work. The diversity of safety/GIS players—ranging from analysts to information technology specialists to officers who collect the data—increases this challenge. Furthermore, these personnel are not housed within a single organization. Interdisciplinary groups such as Virginia's Safety Management System (SMS) Steering Committee, the GIS Technical Advisory Committee, and the Safety Research Advisory Committee all have a role in affecting how GIS can be used to augment safety programs. As a result, there needs to be some way of effectively sharing knowledge among these diverse personnel. Projects of common interest include VDOT and DMV's joint effort to determine crash locations with GPS receivers and VDOT's work to associate HTRIS crash locations with real-world coordinates.

In the course of conducting this study, for example, it became apparent that there are GIS-related efforts planned or underway within different functional units and organizations. These efforts, by themselves, are not unusual, since GIS is an instrument that many believe has tremendous potential. This does mean, however, that it is possible for one person doing GIS work not to know about significant developments in other parts of the organization unless steps are taken to consciously share information and data.

Another challenge to GIS implementation is the fact that while geocoding the MTRS data is feasible, it does not necessarily mean that this geocoding will become standard practice. The method shown in the Appendix is simple but somewhat tedious, since it requires the user to recall the intricacies of several software packages. It may be the case that law enforcement agencies would not have the time to test these different GIS geocoding routines themselves. It may therefore be in Virginia's interest to either develop a method that more efficiently geocodes MTRS data for inclusion into GIS, or at least make the methods shown in this study available to law enforcement agencies statewide. This training would encompass how to geocode MTRS data, how to employ other data sources (e.g., changing the projection for a roadway system), and strategies for keeping information updated, including knowing where to search for accurate roadway data. It is even possible that this type of training or feedback could result in recommendations for changes to future versions of MTRS, such as allowing an officer to enter a street address as a crash location.

### **Another Barrier to GIS Implementation—Choosing Correct Software and Hardware.**

Choosing the correct software and hardware is another barrier to GIS implementation because the techniques that organizations use to record the crash locations and their methods for analyzing crash data will evolve as the capabilities of GPS, GIS, and other technologies improve. For example, the author found that for the particular GIS software employed, two different

versions—one operating in a UNIX/workstation environment and one operating in a PC environment – were necessary for accomplishing some of the required tasks for this research, such as defining route-based coverages, creating and manipulating grids, or altering projections of coverages. Improvements to GIS and GPS products will occur, of course, even if the various persons in the traffic records community that work with traffic records do nothing, since these technologies benefit other industries besides traffic records. One contribution researchers should consider is to develop a better understanding of what analysts need for implementing GIS—whether it is better locating methods or determination of which types of spatial queries are the most helpful. This would help agencies decide how much of an investment they should make in this technology and over what time period that investment should occur.

As an example of such an investment, consider a future pilot project where the various methods for locating a crash are integrated. Depending on resources at the crash scene, it may be more effective for the officer to 1. use a GPS receiver to pinpoint a crash location, 2. identify this location on an electronic map, 3. enter an address or landmark that is tied to a specific location, or 4. measure the distance from a landmark. In such a project, a computer-based display might prompt the officer to manually pinpoint the crash location when this is easy to accomplish, yet to use a GPS receiver in other instances, such as on a rural roadway section with no landmarks. The results of this future pilot project could be translated by the SMS Executive Steering Committee into either *technical specifications*, e.g., “only purchase a certain type of receiver,” or *procedural specifications*, e.g., “record these type of landmarks at rural locations.”

### **Directions for Future Study**

With respect to performing a spatial analysis, we may now be where we were with statistical methods some time ago. We have capabilities to do various queries—not just crash representations, but tasks that involve grids, networks, and movement over time. However, we do not necessarily know how to integrate all of these queries in a logical manner. Probably the next step is to develop a systematic process for spatially studying crash data.

A systematic process already exists for applying statistical methods. Problem areas are identified, then simple models can be used (such as regression models) that may have potential. Then, if these models are found to be inappropriate, more advanced methods can be used. In this statistical approach, routine methods can be used for addressing pitfalls, such as collinearity or sampling size.

As is the case with statistical methods, we should begin to develop a systematic process for conducting spatial analyses. We initiate these spatial queries by listing high crash locations. Next, we move to more advanced methods, such as seeing how crash locations change in time and space, and we determine what variables influence these changes. There is probably a good deal more work to be done toward formalizing a robust method for conducting spatial queries. Just as statisticians or modelers know how to search numbers for patterns, similarly, visually oriented personnel should be able to assess representations of crashes for patterns. The spatial-oriented methodology will not be a one-size-fits-all solution, but it would be an appropriate way to capture GIS capabilities at the microscopic and macroscopic scale.

Training that illustrates how to use some of these spatial and statistical queries to evaluate motor vehicle crash patterns and then reduce these crashes has been offered previously. An excellent example is the Transportation Safety Training Center's 2-day course, "Traffic Analyst Institute." Some of the new GIS capabilities that are now available, such as linking traffic volumes with crash sites, associating pedestrian crashes with proximity to schools, or conducting spatial queries that show how traffic patterns shift during rush hour, could be part of this updated systematic process. The important aspect of such a synthesis is not the GIS capabilities alone, but instead, the thought process that underlies safety evaluations. GIS can help organize and manage the data sets that would be a resource for these evaluations.

## CONCLUSIONS

The first three conclusions relate directly to the purpose of this study, i.e., to ascertain what analytic capabilities offered by GIS may be helpful for crash data evaluation. The remaining conclusions arise as a result of the process of conducting this research and affect how one may deploy GIS for crash data analysis.

1. *At the macroscopic level, suggested benefits of GIS include not only being able to display the data to the user, but also being able to manipulate the data in a creative manner.* Analysis methods such as grid cell modeling, network applications, and risk computations are quite useful. For display purposes, some manipulations of crash data will be necessary in order to make the graphical views readable. These display capabilities can be readily accomplished with the use of techniques such as creating polygon coverages that encircle each intersection and then displaying the total number of crashes at each intersection. Of greater value is the ability to contrast these crash data with other data sets. For example, one can identify crashes in proximity of certain features, such as schools. The literature shows several applications where GIS is used as a method to accomplish types of analysis that would otherwise be difficult.
2. *At the corridor level, one may use a GIS to identify potential problem sites.* In the microscopic analysis, three GIS capabilities were helpful. First, the crash could be positioned exactly where it had occurred along a particular route (for example, 61m or 200 ft from a particular intersection). Second, how close certain crash locations were to one another can be visually ascertained. Third, other data, such as a sharp curve in the roadway, could easily be incorporated into the analysis. Even without a GIS, however, problem areas can still be identified with MTRS software, and crashes can be positioned exactly with HTRIS. Note also that in the case with the corridors studied, some effort was required to redefine the crash location in a convention that would be compatible with the GIS roadway network. (The error associated with the roadway network for this particular study appears to limit its use as a collision diagram, although this issue should be pursued in greater detail if there is a need to set standards for how future roadway data are obtained.)

3. *The literature suggests several innovative uses of GIS for crash analysis, where GIS can be used as an analytical tool rather than merely for its display capabilities.* The GIS applications conducted in this study did not result in definite identification of problem areas, much less recommendations of definite countermeasures. Instead, GIS appears more as an instrument that can help the analyst pinpoint locations that warrant greater study. The common theme between these efforts and those in the literature is that GIS alone cannot replace the need for some type of systematic evaluation of crashes.
4. *There are at least three computer-based methods for obtaining Virginia crash locations within a GIS:* recording the location with GPS receivers, converting from MTRS or other software used by law enforcement agencies, and converting from HTRIS using GIS route properties. In addition, one can ascertain the distance directly from the FR-300P and define a portion of a GIS coverage as a route theme in order to study a corridor in greater detail, as the author did in the microscopic analysis. Each of these methods has advantages and drawbacks. Not all methods are equally suited to all agencies.
5. *There exist several sources of roadway data as well as updates to this information.* For the case study areas, roadway data were available in different levels of detail from the Thomas Jefferson Planning District Commission, VDOT, the City of Charlottesville's Traffic Engineering Department, and the private sector. Some of these GIS coverages may have had a common origin, but each agency may make enhancements or updates to the data set. For persons who desire roadway data that are more accurate than is currently the norm, it will be useful to know where to look for the most current and accurate data set. It is logical that other geographic locations will also have multiple sets of data.
6. *MTRS crash data, or other data that are stored using MTRS's intersection and cross-street format, may be geocoded quickly using addressable street themes.* The match rate will depend heavily on at least two factors: the number of "private property" links coded in the MTRS data and the correspondence between street names employed by the MTRS user and street names as available in the GIS street coverage. Steps may be taken to improve the matching rate, such as identifying landmarks in the MTRS data and associating them with certain attributes in the GIS street coverage. This is of interest to law enforcement officials who do not have access to Virginia's HTRIS database as well as jurisdictions whose local roads are not stored within HTRIS. A caveat is that the precision of the location is only as accurate as the MTRS links; a very long link between two cross streets may be more problematic.
7. *There are activities being conducted by VDOT, DMV, and other organizations that are of mutual interest.* The previously-cited examples of bringing HTRIS crash locations into a GIS or conducting a pilot effort where the locations of crashes would be recorded with a GPS receiver support this finding.

## RECOMMENDATIONS

VDOT's SMS Executive Steering Committee should consider the first two recommendations, as they fall within the purview of safety research. The third recommendation, however, should be directed to VDOT's Data Management Division, which will be overseeing VDOT's GIS operations.

1. *A synthesis of how to systematically identify crash countermeasures should be revisited.* The first three conclusions suggest that researchers have used GIS as an integral tool for crash countermeasure identification, just as one would rely on a word processor, spreadsheet, or statistical package. As discussed earlier in this report, the synthesis would combine statistical and spatial queries. In fact, this synthesis could take the form of a document, primer, or flow chart that conveys how to evaluate a city or region's roadway network from a safety perspective.
2. *Law enforcement agencies should be provided with training for bringing MTRS data into a GIS.* As implied by the fourth, fifth, and sixth conclusions, it is quite possible to begin using GIS now within law enforcement agencies. The forum at which this initiative should be discussed is open to Virginia's SMS Executive Steering Committee. Several options for how training can be accomplished are available: a technical bulletin may be disseminated, an on-site forum where training could be done in person could be convened, or an application that would supplement future MTRS versions could be developed.
3. *There should be a liaison between the SMS Executive Steering Committee and VDOT's GIS Technical Advisory Committee (or an equivalent of this).* As implied by the last conclusion, this liaison need not be complex or time consuming. It could be as simple as having one member of the GIS Committee spend 30 minutes on a semiannual basis briefing the SMS Committee as to GIS developments being undertaken by VDOT. VDOT is currently reorganizing GIS functions that will fall within a newly created Data Management Division. However, the benefits of fostering coordination between VDOT's GIS units and Virginia's SMS persist.

## ACKNOWLEDGMENTS

The author acknowledges the assistance of representatives from several agencies who provided data, insights, or data collection assistance useful for this research: the Albemarle County Police Department (D. Karr and D. Shaw), the City of Charlottesville Police Department (R. Roberts), the City of Charlottesville Traffic Engineering Department (J. Marshall), the Thomas Jefferson Planning District Commission (D. Roberts), Timmons Associates (R. Trott), the University of Virginia Police Department (M. Genay), the Transportation Safety Training Center at Virginia Commonwealth University (R. Breitenbach), VDOT (T. Klinker), and the Virginia Transportation Research Council (G. Allen, S. Brich, R. Combs, G. Fitch, W. Ferguson, M. Hickey, J. Jernigan, A. O'Leary, C. Stoke, and N. Younts). Mention of these agencies and representatives does not, however, imply their agreement with the contents of this report.

## REFERENCES

1. Goh, P. C. 1993. Traffic Accident Analysis Using Geoprocessing Techniques. *Road and Transport Research*. Vol. 2, No. 2, pp. 76-85.
2. Crespo Del Río, R., Pérez De Madrid, P. A., and Herrero, E. G. 1997. Geographic Information System Applied to Pavement Evaluation by Video-Laser Road Surface Tester Multifunction Device. *Transportation Research Record 1592*. Transportation Research Board, Washington, D.C.
3. Mohle, D., and Long, H. 1996. GIS Helps with Traffic Accident Mitigation. *Better Roads*. Vol. 66, No. 3, pp. 29-31.
4. Miller, S., Johnson, T., Smith, S., and Raymond, L. 1995. Design and Development of a Crash Referencing and Analysis System. *Proceedings of the 1995 Geographic Information Systems for Transportation (GIS-T) Symposium*. American Association of State Highway and Transportation Officials, Washington, D.C.
5. Hovenden, E., Sligoris, J., and Walker, C. 1995. A Roads Spatial Information System: The Victorian Experience. *Road and Transport Research*. Vol. 4, No. 1, pp. 110-121.
6. Meyer, M. D., and Sarasua, W. 1992. Geographic Information System-Based Transportation Program Management System for County Transportation Agency. *Transportation Research Record 1364*. Transportation Research Board, Washington, D.C.
7. American Association of State Highway and Transportation Officials. 1998. *Briefs of Research Problem Statements Considered by the AASHTO Standing Committee on Research*. Research Problem Statement (99-G-31). Washington, D.C.
8. National Association of Governors' Highway Safety Representatives and the U.S. Department of Transportation. 1997. *Revised Draft Guideline for Minimum Uniform Crash Criteria*. Washington, D.C.
9. Slagle, M. R. 1998. Exploring School District Population Changes with GIS. *Geo Info Systems*. Vol. 8, No. 3, pp. 42-44.
10. Bureau of Transportation Statistics. 1998. *Development of a Crash Referencing and Analysis System*. Website maintained by BTS in reference to a visit to North Carolina. Observed again on August 28 1998 at <http://www.bts.gov/gis/outreach/statevisit/nc/appcrash.html>.
11. *Workbook* from the *Traffic Analyst Institute*. (no date). Courtesy of the Transportation Safety Training Center at Virginia Commonwealth University, Richmond, Virginia.

12. Austin, K., Tight, M., and Kirby, H. 1997. The Use of Geographical Information Systems to Enhance Road Safety Analysis. *Transportation Planning and Technology*. Vol. 20, pp. 249-266.
13. Panchanathan, S., and Faghri, A. 1995. Knowledge-Based Geographic Information System for Safety Analysis at Rail-Highway Grade Crossings. *Transportation Research Record 1497*. Transportation Research Board, Washington, D.C.
14. Affum, J. K., and Taylor, M. A. P. 1995. Integrated GIS Database for Road Safety Management. Technology Tools for Transportation Professionals—Moving into the 21st Century: Resource Papers for the 1995 International Conference. Institute of Transportation Engineers, Washington, D.C., pp. 189-193.
15. Raub, R.A. 1997. Occurrence of Secondary Crashes on Urban Arterial Roadways. *Transportation Research Record 1581*. Transportation Research Board, Washington, D.C.
16. Choi, K., and Park, I. 1996. Traffic Accident Analysis with GIS and Statistical Package. *Ninth Symposium on Geographic Information Systems for Transportation*. American Association of State Highway and Transportation Officials, Washington, D.C.
17. Kim, K., Levine, N., and Nitz, L. 1995. Development of a Prototype Traffic Safety Geographic Information System. *Transportation Research Record 1477*. Transportation Research Board, Washington, D.C.
18. Souleyrette, R. R. II, and Sathisan, S. K. 1994. GIS for Radioactive Materials Transportation. *Microcomputers in Civil Engineering*. Vol. 9, No. 4, pp. 295-303.
19. Patel, M. H., and Horowitz, A. J. 1994. Optimal Routing of Hazardous Materials Considering Risk of Spill. *Transportation Research*. Vol. 28A, No. 2, pp. 119-132.
20. Mefford, C. 1995. Using Geographic Information Systems to Plan On-Street Bicycle Routes. *Fifth National Conference on Transportation Planning Methods Application—Volume II*. Transportation Research Board, Washington, D.C.
21. Austin, K. P., Tight, M. R., and Kirby, H. R. 1995. An Advanced System for the Study of Children's Safety on the Journeys to and from School. *Road Safety in Europe and Strategic Highway Research Program*. No. 2A:3 VTI Konferens, Linköping, Sweden.
22. McKnight, A. S., Mosher, C. W., and Bozak, D. J. 1998. *Evaluation of Emerging Technologies for Traffic Crash Reporting*. InfoGroup, Inc. Federal Highway Administration. McLean, Virginia.
23. Aultman-Hall, L., and Hall, F. L. 1998. Ottawa-Carleton Commuter Cyclist On- and Off-Road Incident Rates. *Accident Analysis and Prevention*. Vol. 30, No. 1, pp. 29-43.

24. Siegel, J., and Yang, S. 1998. TMS: Indianapolis's Data Traffic Cop. *Geo Info Systems*. Vol. 8, No. 5, pp. 38-40.
25. Saccomanno, F. F., Chong, K. C., and Nassar, S. A. 1997. Geographic Information System Platform for Road Accident Risk Modeling. *Transportation Research Record 1581*. Transportation Research Board, Washington, D.C.
26. Affum, J. K., and Taylor, M. A. P. 1996. GIS-Based System for the Systematic Determination of Road Accident Problems and Causal Factors. *Roads '96 Conference Proceeding-Part 5*, ARRB Transport Research Ltd., Victoria, Australia.
27. Opiela, K. S. 1998. National Cooperative Highway Research Program Research Results Digest, Number 221, *Development of System and Application Architectures for Geographic Information Systems in Transportation*. Transportation Research Board, Washington, D.C.
28. Lamm, R., Guenther, A. K., Choueiri, E. M. 1995. Safety Module for Highway Geometric Design. *Transportation Research Record 1512*. Transportation Research Board, Washington, D.C.
29. McPherson, L., Shuller, E., Rosendahl, F., and Tilley, C. 1996. A GIS Application Used for Highway Safety Studies. *Ninth Symposium on Geographic Information Systems for Transportation*. American Association of State Highway and Transportation Officials, Washington, D.C.
30. Stokes, R. W., and Marucci, G. 1994. *Development of a Prototype GIS-Based Accident Mapping Program for Kansas*. Kansas State University, Manhattan, Kansas.
31. *Final Document*. November 12, 1997. (Unpublished data, provided to VDOT by a consultant).
32. O'Neil, W. A., and Harper, E. 1997. Location Translation within a Geographic Information System. *Transportation Research Record 1593*. Transportation Research Board, Washington, D.C.
33. Pawlovich, M. D., and Souleyrette, R. R. 1996. A GIS-based Accident Location and Analysis System (GIS-ALAS). *1996 Semisesquicentennial Transportation Conference Proceedings*. Iowa Sate University. Ames.

## APPENDIX: HOW TO USE MTRS DATA WITHIN A GIS COVERAGE

The following steps illustrate one method for converting MTRS data, version 5.0, for use within a GIS coverage. There are many ways to accomplish this task, including the creation of a subroutine that would be more automated than what is shown here. Selection of the best method will be dependent on the user's familiarity with the various software packages. For ease of illustration, conversion of the Charlottesville crash data from MTRS version 5.0 to a PC ArcView format are presented for use with ArcView GIS version 3.0a.

Once the data have been converted to an ASCII format, these steps may be summarized as building the lookup table between link numbers and street names, geocoding the intersection crash locations, geocoding the midblock crash locations, identifying special landmarks used within MTRS, and troubleshooting in order to increase the match rate.

### Convert Binary Data to ASCII.

With MTRS data, one can use a utility called *bin2asc*, which stands for binary-to-ASCII. This file is stored as *bin2asc.exe*. The accident data files, in this case *accident.89*, *accident.90*, and so forth through *accident.98* are placed in this same subdirectory. First, double click on *bin2asc*, press "1" for "binary-to-ASCII" conversion. Second, press "2" for accident file. Third, enter *accident.89*, which is the binary file to be processed. Fourth, enter *crashc89*, which is the name we will give to the resulting ASCII file. Fifth, repeat this process for each of the annual binary files, such that you end up with ten different ASCII files: *crashc89*, *crashc90*...*crashc98*. Sixth, these crashes may then be combined within a single file. Seventh, a row of raw ASCII data may be printed from this file and it may be compared to an interactive file from the MTRS application to determine the meaning of the different variables. Eighth, the data may be loaded into a spreadsheet such as Microsoft Excel. Finally, the MTRS accident update module may be run, using the F1 key to obtain help for learning the meaning of each variable.

### Use the MTRS Lookup Table to Replace Street Numbers with Street Names

An easy way to construct the lookup table is to use a spreadsheet. The MTRS file *street.dat* contains the links between the street names and the road numbers. One may place these data in the spreadsheet and give the block of cells a name. In this case, one may use the name *range*. Then, for each crash, associate the number and link. For example, if cell K2 contains one of the intersecting link numbers, the statement `VLOOKUP(K2, RANGE, 2, FALSE)` will return the name associated with that link. (The reason for the "2" is that the first column of *range* contains the link numbers and the second column contains the link names; based on the link number in the first column one wants the *vlookup* statement to return the street name in the second column. The "false" statement is preferred because one only wants a match if the link enumerated by the officer matches a link from our lookup table; this makes the criteria for matching very stringent.) In most cases, each crash will be associated with two or three intersecting links. Hence, one would have two or three VLOOKUP statements for each crash—two for an intersection location and three for a midblock location. For example, if a crash simply

occurred near the intersection of Altamont Street and High Street, then there would only be two numbered links. On the other hand, if a crash occurred on Grady Avenue between 14<sup>th</sup> Street and 15<sup>th</sup> Street, then there would be three numbered links. The difference in how these crashes are represented is significant. For exporting to the GIS format, it simplifies matters to then concatenate every two intersecting streets into a single variable for each crash. For an intersection of two streets where the names (not the link numbers) are stored in the columns BC2 and BD2, the statement CONCATENATE (BC2, " ", "&", " ", BD2) will yield a single intersection for the crash that can then be geocoded with the PC ArcView software. For crashes that occurred at an intersection, there would only be a single variable, which will be called *intersect1*. For crashes that occur at a midblock location, there will be two variables—*intersect1* and *intersect2*. The rows may then be sorted by midblock versus intersection locations and stored as a dbase IV file, which will be imported into ArcView as two separate files—one for the intersection locations and one for the midblock locations.

### Geocode Intersection Crashes

Within ArcView, initially it may be necessary to build a geocoding index. This is done by making the streets theme active, going to theme properties, pressing the geocoding icon, and then pressing O.K. Next, the crash database file containing only intersection-crashes is imported. In ArcView, open the file that contains the street coverage, go to Add Table, then add the dbase table that was just created in the spreadsheet application. Next, go back to Views, highlight the street coverage theme, and go to View\Geocode addresses. In the popup menu that follows, change the address field to the variable name that contains the intersecting streets joined with the “&” symbol. In this case, this was the variable *intersect*.

In practice, a time-consuming task was redoing the geocoding process several times in order to catch crashes that were not located (or not located properly). The following points need to be considered:

- *The street names in MTRS are not necessarily the same in the roadway coverage.* For example, a street known as BYPASS 250 in MTRS is identified as U.S. HWY 250 BYPASS in the GIS data. Order is important—the city data might have a street called “Main Street East,” whereas the GIS data might refer to the street segment as “East Main Street” and misspellings may occur either in the MTRS data or the GIS data. For example, a portion of Hydraulic Road was characterized as “Woodlands” Road.
- *Some streets in the roadway coverage were simply not identified.* For example, many ramps were simply given the name “ramp.”
- *Some of the crashes had been simply coded as intersecting with “private property”* where the location of this property was unknown.

- *Street name changes may occur at different points.* In one location, a physical road known as “Main Street” changes to “University Avenue”; the MTRS system and the roadway coverage data denote the change at different locations.
- *Ramps are a special case.* Names such as “250 E/B OFF RAMP EMMET N” hindered the matching process; therefore, shortened ramp names needed to be assigned.
- *Landmarks not shown in the roadway network file may be used in MTRS,* such as “Barracks Road Shopping Center.” A lookup table may be created and implemented, either in a spreadsheet environment or in a GIS; an example is shown in Table A-1. It may also be necessary to modify the street coverage if landmarks are not shown.

**Table A-1: Example Lookup Table**

ALIAS	ADDRESS
HYDRAULIC ROAD & K-MART LOT	350 HYDRAULIC ROAD
K-MART LOT & HYDRAULIC ROAD	350 HYDRAULIC ROAD
KROGER PK LOT & EMMET ST	2050 EMMET ST
EMMET ST & KROGER PK LOT	2050 EMMET ST
KROGER PK LOT & HYDRAULIC ROAD	500 HYDRAULIC ROAD
HYDRAULIC ROAD & KROGER PK LOT	500 HYDRAULIC ROAD

It is recommended that one experiment with the location conversion process at this stage, since being able to accurately locate crash intersections is fundamental to MTRS and will save time when it is necessary to pinpoint midblock crashes.

### **Geocode Midblock Crashes**

A midblock crash location is one that occurs between two street intersections. Various methods were attempted for geocoding these locations, such as obtaining the street address in closest proximity of the intersections. However, additional customization of this geocoding method to account for the case when a street segment crossed zip code boundaries would have been necessary. Therefore, the ArcView script *View.AddXYCoordToFTab* was used to associate each intersection with an XY coordinate and the resulting XY coordinates were averaged. In practice, one may geocode the two intersections for each crash as two separate coverages; then obtain the XY coordinates. Next, one may join the resultant tables, and then geocode the average of the two intersection coordinates. An alternative would be to create a lookup table that associates the specific addresses or midblock coordinates with the desired midblock locations.

## Summary of MTRS Conversions

This method is only as accurate as the coding scheme that was originally used in MTRS. Furthermore, one possible disadvantage with the method used for geocoding midblock crashes is that for a road segment that has a sharp curve, one could end up with a crash location not on the roadway. If this is a concern, then the address-based method to locate the crashes can be used. Note also that MTRS 6.0 allows GPS coordinates to be included directly as part of the crash record.

### What if Greater Precision for Locating the Crash is Required?

It may be the case that one wishes to record certain crash locations with greater precision than what is available in MTRS. This may be addressed in the same manner as was done with the Albemarle County crashes shown in figure 9. The County identified two sections of road where they knew many crashes regularly occurred, but where precise location information was not available. Precise location information was lacking because the locations are stored within the MTRS system as simply being on a particular segment, where in these cases the segment is quite long.

Segments corresponding to the two county roads (Route 22 and Route 231) were defined as a route theme in Arc/Info. This means that one can pinpoint any location as a function of its distance from the start of the route, such as 322 m (0.2 mi), from the beginning of Route 22. Then, the FR-300P was used to place the crash as it had been measured by the officer. The remainder of the crash data shown on the FR-300P report system were converted from the MTRS system and associated with the crash location.

Finally, the individual crash sites were visited. Where the crash diagram allowed one to identify the crash location more precisely and traffic conditions allowed a measurement to be taken safely, the distances cited on the FR-300P were recomputed. For example, a crash that occurred on Route 231 was cited in the report as being 322 m (0.2 mi) south of Route 615. A measurement taken with a wheel suggested that the first point of impact was approximately 81 m (264 ft) further than that location. It appeared that one could obtain a more precise distance figure for 26 out of the 142 crashes shown on Route 22 and Route 231.