Building a Regional Traffic Crash Data System
- Bridging the Gaps

by

Lina Chen Kulikowski

Broward County Metropolitan Planning Organization

and

Ilir Bejleri

University of Florida, Geoplan Center

Presented at the

26th Annual ESRI International User Conference

August 2006

Abstract

The Metropolitan Planning Organization (MPO) of Broward County, Florida has developed a procedure for mapping and analyzing automobile, bicycle and pedestrian crashes to identify problematic areas in order to prioritize transportation improvement funding and to develop countermeasures.

In the Broward County metropolitan area, traffic crash reports are generated and maintained by various local law enforcement agencies and state safety and transportation departments. This paper discusses challenges and successes of integrating crash databases from multiple agencies as well as the process of crash mapping using a customized ArcGIS geocoding extension, and crash analysis using a combination of ArcGIS geoprocessing, linear referencing, Spatial Analyst, Tracking Analyst and 3D Analyst.

Introduction

Metropolitan Planning Organizations (MPOs) are transportation planning entities, federally funded through the US Department of Transportation, for urbanized areas in the United States. The Broward County Metropolitan Planning Organization (BCMPO) is responsible for the urbanized area in Broward County located in southeast Florida between Miami-Dade and Palm Beach Counties. In response to the federal requirements of MPOs to consider safety explicitly in the transportation planning process – Safety Conscious Planning, BCMPO has various programs to address safety issues of the metropolitan area's transportation system. One of the BCMPO's safety programs focuses on mapping and analyzing traffic crashes using GIS. At present, approximately 54,000 traffic crashes occur in the Broward County metropolitan area every year. Traffic crash reports are generated and maintained by various local law enforcement agencies and state safety and transportation departments. The recent efforts in mapping and analyzing traffic crashes have revealed a number of problems that are related to crash data collection methods, differences in crash data structures and lack of custom crash mapping software tools for the County's specific needs.

Objectives

Mapping and analysis of automobile, bicycle and pedestrian crashes on yearly basis is essential in identifying problematic areas in order to prioritize funding to develop and implement countermeasure. The purpose of this project is to review current crash mapping processes in Broward County and develop solutions and guidelines for future improvements. Several objectives were identified to achieve the long term goal of the County for integrating crash databases from multiple agencies into a countywide Traffic Crash Data System:

- Identify and solve problems associated with Geocoding/mapping of crashes based on addresses and intersection descriptions;
- Identify and solve problems associated with aggregating crashes into nodes and road segments for crash rate analysis;
- Preserve lessons learned and streamline the geocoding process;
- Identify existing GIS tools and develop new GIS tools to provide solutions to existing problems;
- Develop GIS analysis methods to identify high crash locations;
- Develop and propose standards to be followed by all agencies involved.

Problems (the gaps) and solutions (the bridges)

This section will discuss the problems encountered during BCMPO's efforts in building a regional traffic crash system and the solutions developed to address those problems. The solutions have been highlighted in dark red. The problems and solutions are grouped into three categories: traffic crash data collection, mapping traffic crashes, and crash analysis / high crash locations.

Traffic Crash Data Collection

At present, traffic crash reports are generated by numerous local law enforcement agencies including the Broward County Sheriff's Office (BSO) serving the unincorporated areas plus 12 municipalities, 20 municipality Police Departments not served by the BSO, and the Florida Highway Patrol (FHP) serving freeways. Although the traffic crash report forms are standardized throughout the state, agencies use different database systems to maintain the traffic crash records. Some agencies only keep hard copies. Others enter different data fields from the original reports into their databases. Database systems get changed over time and, in some cases, data become inconsistent between the old and the new systems. Although hard copies of traffic crash reports are generally available upon request, the local agencies have different policies for digital data distribution. Some provide data freely, some require an official letter, and some provide only paper printouts of digital data because their systems do not provide the capability to "dump" digital data for distribution.

At the state level, the Florida Department of Highway Safety and Motor Vehicle (DHSMV) maintains a statewide database of traffic crash records. Complete set of crash records are available only for long-form crashes and limited number of data fields are entered for short-form crashes. Typically, DHSMV makes available only the long form crash records. A law enforcement officer must submit a long-form crash report when investigating:

- Motor vehicle crashes resulting in death or personal injury, or
- Motor vehicle crashes in which one or more of the following conditions occur:
  - Leaving the scene involving damage to an attended vehicle or property
  - Driving while under the influence of alcoholic beverages, chemical substances, or controlled substances or with an unlawful blood alcohol level.

An investigating officer may report other traffic crashes on the long-form report.

The yearly crash records are normally released in August of the following year, which creates potential discrepancy between crash locations and the road layer used for geocoding. The short form crash records can be obtained from local agencies, which store the crash data in various digital formats.

Two report numbers are used in the standard traffic crash report forms. The first number, generally referred to as the DHSMV number, is preprinted on the traffic crash report forms and the second number, the local agency number, is filled out by the investigating agency. DHSMV uses the DHSMV number as the unique identifier for its crash records while the local agencies use the local agency report number. The use of different crash tracking numbers makes it difficult to cross-reference data from DHSMV and those obtained from the local agencies.

To bridge the gaps, a spreadsheet matrix was created to compare the different database systems maintained by local agencies. Methods were developed and tailored to import existing data, in different file formats, into a standardized personal geodatabase.

Short term solutions were proposed to standardize crash location fields for geocoding. A set of essential data fields for crash mapping and analysis were identified. These common fields will be proposed to the local agencies for adoption into their databases, e.g. to include the DHSMV report number along with the local report number.

Long term solutions were also proposed, beyond data standardization, for a countywide integrated crash database. This includes the creation of a web-based data entry, GIS-based location validation, and data sharing system for crash data. A web-based system has the advantages of centralized data processing while providing immediate access to up to the minute geocoded crash data.

While the short and long term solutions will take some time to implement, the BCMPO had obtained the 2002 and 2003 long form crash records from DHSMV and had already started to use GIS to develop methods and procedures for crash mapping and analysis. University of Florida provided the technical services in developing ArcGIS tools to automate and enhance the procedures developed by BCMPO. The tools and procedures developed were tested on the 2004 long form crash records from DHSMV.

Mapping Traffic Crashes

Every year, approximately 54,000 crashes occurred in Broward County with an estimated fifty-fifty split of long and short form traffic crashes. Prior to 2002, DHSMV maintained digital records of all long form traffic crash data fields except those fields containing location descriptions. Without the location fields, mapping roughly 27,000 crashes per year requires substantial resources that the County does not have. DHSMV began to record location fields in 2002 which opened the door for geocoding countywide traffic crash locations with the help of GIS.

The Florida Department of Transportation has an ongoing process of mapping traffic crashes occurring on state roads using its Linear Reference System (LRS)

based road layer. The crashes mapped by FDOT make up approximately 60 percent of the long form crashes in Broward County which leaves approximately 10,000 crashes, occurring on county or city roads, to be mapped by BCMPO.

During the process of geocoding crash locations, several crash location problems were revealed and needed to be addressed to increase the success rate of geocoding:

- Alternative street names / alias missing from the streets base layer maintained by the County;
- Misspellings and abbreviations;
- Special locations such as overpasses, railroad crossings, bridges, and shopping center entrances that can not be geocoded using the streets base layer maintained by the County;
- Crash location fields need to be converted to match the style required by geocoding in ArcGIS.

Broward County Planning Services Division maintains a streets base layer with address ranges and an alias table. As additional aliases were discovered through geocoding traffic crashes, they were reported for inclusion. Coordination between the BCMPO and the Broward County Planning Services Division was the key for the success of this process.

For misspellings and abbreviations, a "data dictionary" tool was developed using VBA with ArcObjects. Commonly misspelled or abbreviated words were stored into the data dictionary by reviewing crash records rejected during initial geocoding. The "data dictionary" tool "cleans up" crash location records using the existing data dictionary and allows users to add new entries into the data dictionary during the "interactive" geocoding process. The data dictionary preserves errors found in the past and saves time for matching new crash records in the future.

For special locations that can not be geocoded using the generic streets layer, a "Points of Interest" point layer was created as an address locator for special locations.

For converting crash location fields into a geocodable style, a data preparation model was built using ArcGIS Model Builder to automatically perform the following steps:

- Combine multiple crash location fields into one "crash_address" field;
- Clean up the "crash_address" field by removing non-location related text such as "unknown";
- Add a city id field required for the address locator style "US streets with zone";
- Standardize "distance from intersection" from a mix of feet and miles into feet.

- Identify crashes occurred away from public roads such as parking lots, and private properties. These crashes are excluded from geocoding.
- Identify crashes already mapped by FDOT. These crashes were already geocoded by FDOT and are excluded from geocoding.
- Tag crashes involving pedestrians or bicyclists.
- Tag the locations with categories of "at intersection", "from intersection", "house address", "state road", and "mile marker". These categories are used later in during the geocoding process.

Several road layers were used in geocoding traffic crashes not mapped by FDOT. The Broward County Planning Services Division maintains a Streets layer that includes all public roads in the county and has address ranges, city codes, and an alias table. This layer is kept up-to-date mainly to support E-911 functions. The Dynamap® streets layer by Tele Atlas (available from FDOT) also has address ranges, city codes and an alias table. However, the city codes in Dynamap® are outdated because they were created using Census 2000 data. The Broward Streets, Tele Atlas Dynamap®, and a "Points of Interest" point layer were used to map traffic crashes occurred on non-state roads.

A "Batch Geocoding" model was built, using ArcGIS Model Builder, to automatically create three address locators from these data layers and automatically geocode traffic crash records. The original idea was to build a model that will create a composite address locator to geocode crashes. Unfortunately, at this time, ArcGIS Model Builder does not support the composite address locator style. As an alternative, the "Batch Geocoding" model creates and runs through the three address locators sequentially which mimics geocoding using a composite address locator. After the "Batch Geocoding" tool has been executed, the resulting address locators can be modified and/or a composite address locator can be built within ArcCatalog. Consequently, the composite address locator, using all three sources of reference data, provided higher success rates in geocoding traffic crashes. An additional reason to create a model to perform batch geocoding was the need to store the user reference segment id for each crash record needed later for crash analysis. The ESRI standard locators do provide the reference ID that is the Object ID of a segment but not the user segment ID. Converting the Object ID to the County segment ID is one of the processes included in the model.

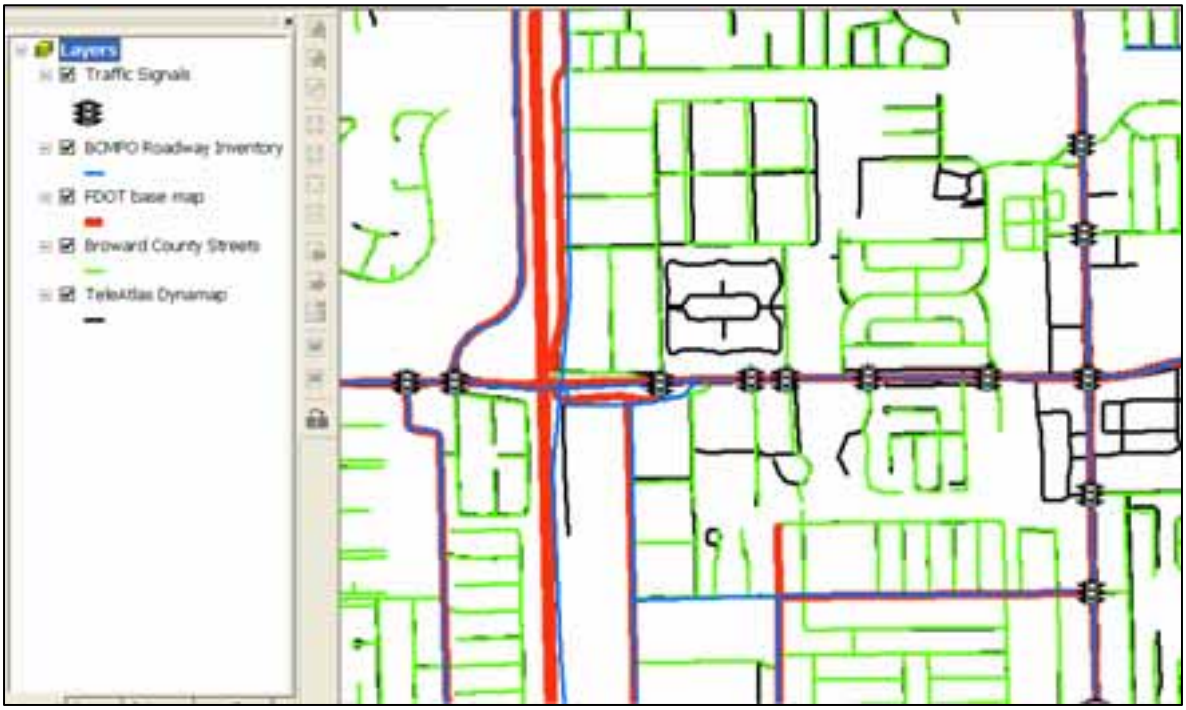
Another problem with mapping traffic crashes is the offset. A very common way of recording a traffic crash location is "on road A, so many feet or miles, from road B". The out-of-box ArcGIS geocoding tool only places the crashes at the intersection of road A and road B. There still is a need to "move" the points to their actual locations if offsets apply.

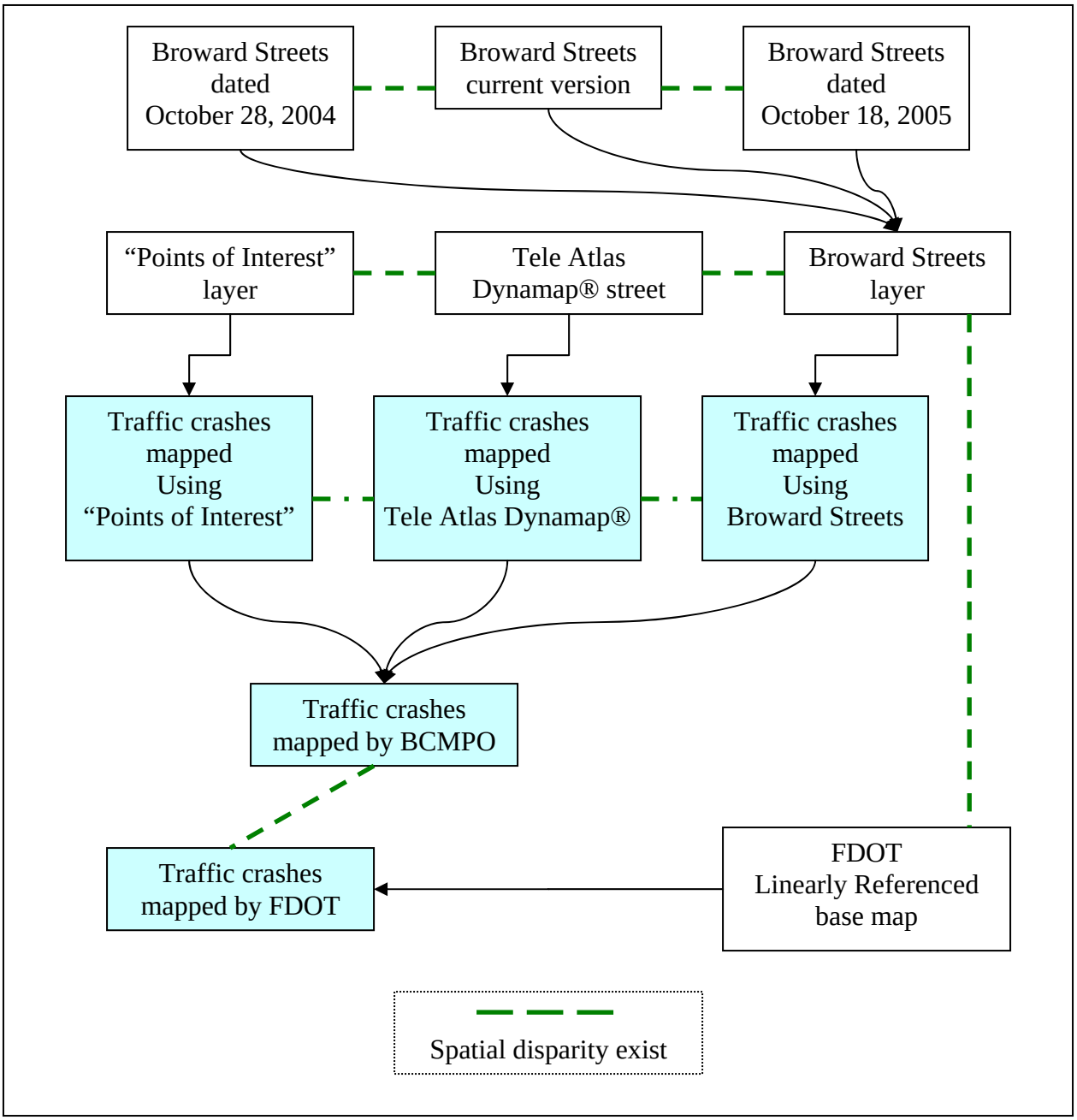
An "apply offset" tool was developed using ArcGIS VBA to move the geocoded traffic crashes to their true locations using the offset values recorded in the DHSMV crash records and the Broward Streets base layer.

Crash Analysis and High Crash Locations

The purpose of mapping traffic crashes is to identify problematic areas in order to prioritize funding to develop and implement countermeasures. Two commonly used measures for identifying high crash locations are *crash rate* and *crash severity index*. In order to develop these measures, additional steps need to be taken after the traffic crashes are mapped. First the traffic crashes need to be aggregated by intersections and/or road segments, number of crashes and traffic volumes at intersections and/or along road segments are required to calculate the crash rates, and the crash severity index is calculated based on the number of crashes as well as crash severity types of "fatality", "injury", and "property damage only".

The first problem encountered when aggregating traffic crashes to intersections or road segments is the geographical inconsistency among traffic crashes mapped by different sources. The Florida Department of Transportation (FDOT) maintains a Linear Reference System based road layer that includes major roads classified as collectors and above. This layer was used to map traffic crashes that occurred along state roads. Traffic crashes geocoded by BCMPO are from a composite address locator of three sources – Broward Streets layer, Tele Atlas Dynamap® street layer, and a "Points of Interest" point layer. Geographical inconsistency from the base layers used in geocode and mapping contributed to the geographical inconsistency between crashes geocoded by the BCMPO and those mapped by FDOT. Below is a snapshot depicting spatial disparity among the different data layers used in this project.

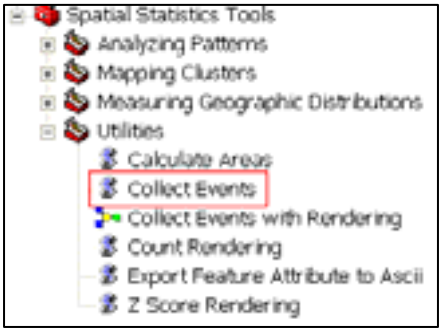




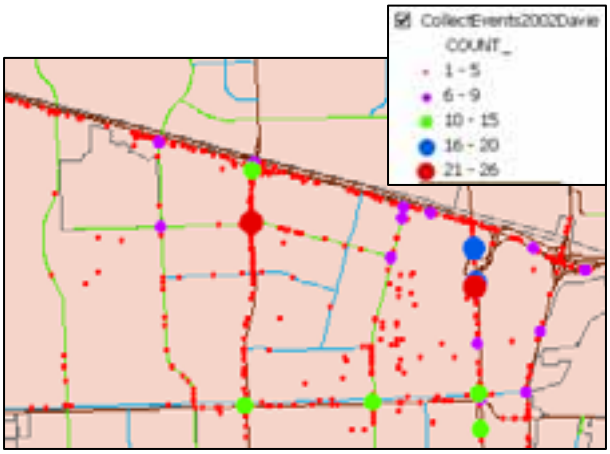
To address the above problem, all traffic crashes were transformed into one consistent reference system – Broward Streets. Traffic crashes geocoded using the TeleAtalas Dynamap® street layer and the "Points of Intersect" point layer were "snapped" to Broward Streets to either the nearest node or the nearest segment. Traffic crashes mapped by FDOT had been assigned to a node system maintained by FDOT. Therefore, these crashes were brought into the Broward Streets reference system through node matching between the FDOT nodes and the Broward Streets nodes.

As Broward County continues to realign road segments to improve the position accuracy of the Streets base layer, a simplified linear referencing technique was implemented to enable mapping crash locations, as events, along the future versions of the Broward Streets base layer. M values were added into the Broward Streets base layer using the "Create Routes" tool from the Linear Referencing Toolbox. As mentioned previously, one of the steps in the "Batch Geocoding" model assigns the Broward Streets segment IDs and measures along the segments to traffic crashes. This task can also be accomplished using the "Locate Features Along Routes" tool from the Linear Referencing Toolbox.

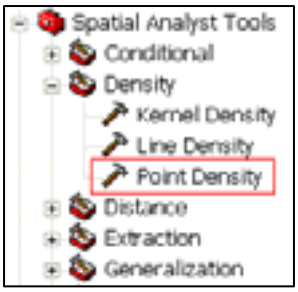
The second problem that needed to be addressed was how traffic crashes can be aggregated into intersections and/or road segments once they are brought into a consistent reference system. Three methods were explored to aggregate traffic crashes.

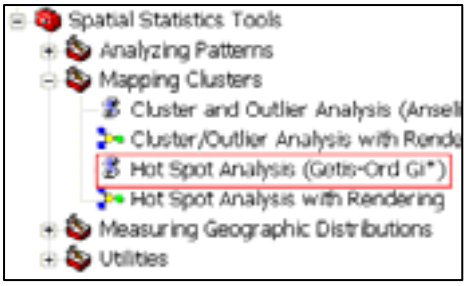


The first method used the "Collect Events" tool from the Spatial Statistics Toolbox that came with ArcGIS 9.0. The "Collect Events" tool converts incidents into weighted point data with each point representing a unique location with an attribute value of total number of incidents occurred at that location. The image below at left displays traffic crashes as they were mapped. A common problem with this kind of display is that multiple crashes at the same location are not visually detectable because they are "stacked". The image below at right tells a very different story once the crashes were aggregated by "Collect Events". Although the "Collect Events" tool provides a very simply and quick solution of detecting locations with high number of incidents, it has a shortcoming of only aggregating incidents at unique locations. It does not provide aggregation of traffic crashes that are, for example, "250 feet around an intersections".

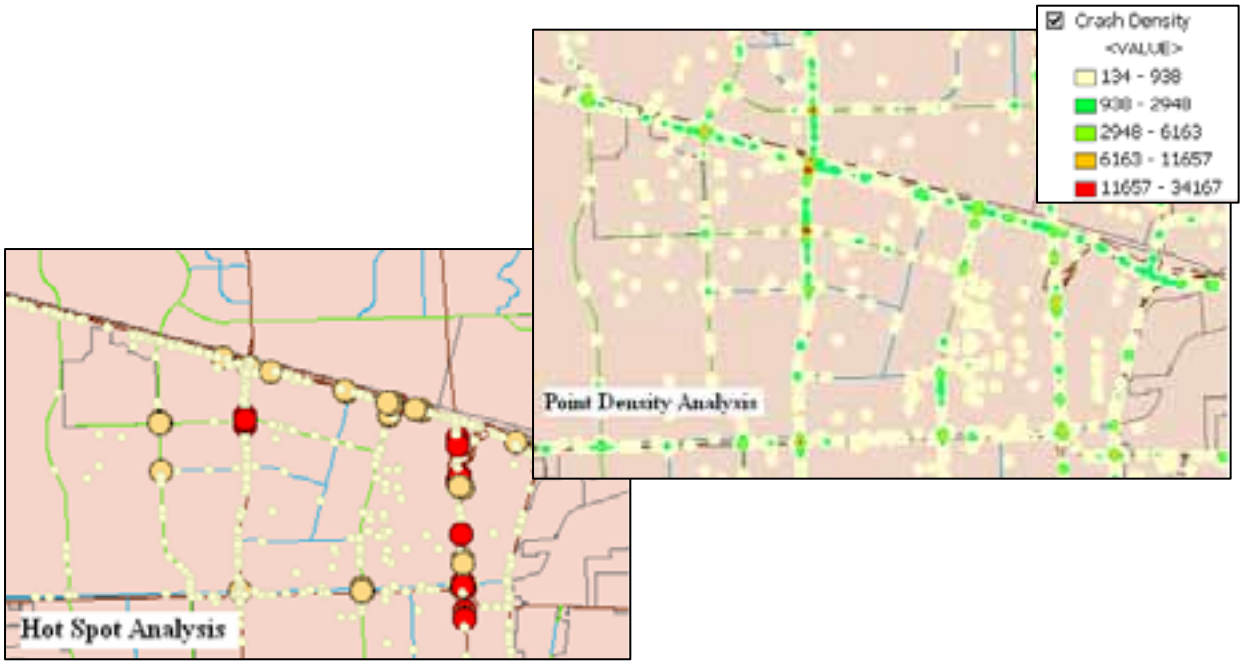


The second method used the "Point Density" tool from the ArcGIS Spatial Analyst Extension and the "Hot Spot Analysis" tool from the Spatial Statistics Toolbox that came with ArcGIS 9.0. The "Point Density" tool can be applied directly on the traffic crash point layer. However, the "Hot Spot Analysis" tool requires a weighted point layer such as the one created by the "Collect Events" tool from the first method.





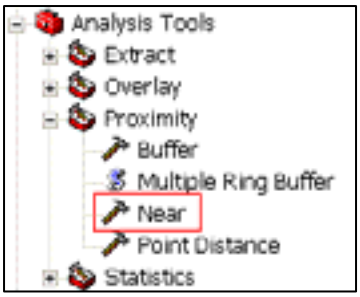
The image at left below displays the results from the "Hot Spot" analysis, which does not provide an accurate picture of high crash locations. This is because the "Hot Spot Analysis" tool does not aggregate neighboring events as expected. It only "checks" if a high number is next to another high number to find clusters of extremely high incidents. The image at right below displays the results from the "Point Density" analysis where areas highlighted in red have the highest density of crash incidents. Notice the density displays are not as prominent as those symbolized in "Collect Events" and "Hot Spot". This is because the density values are stored in a grid data layer. The grid cell size and cell values are confined by the parameters used for the point density analysis - in this case, 50 feet for grid size and 250 feet for search radius. Although the displays can be exaggerated by using a larger search radius and grid size, the results can be skewed because crashes outside of the influence area of a "hot spot", such as the typically used 250 feet around an intersection, will be included in the density calculation.

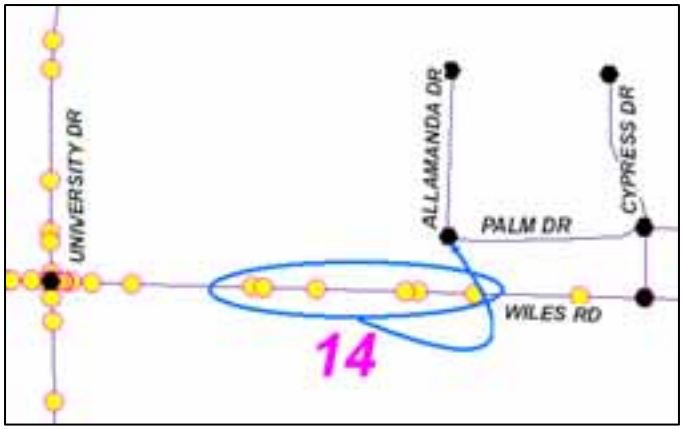


There is, however, yet another shortcoming from the "Point Density" analysis. For example, as illustrated in the graphic at right, one minor intersection, Almond Ave @ Las Olas Blvd, is less than 250 feet away from two major intersections, State Road A1A W (Seabreeze Blvd) @ Las Olas Blvd and State Road A1A E @ Las Olas Blvd. By using a 250 feet search radius, crashes occurred at the two nearby intersections are included in the calculation of total number of crashes as well as crash density for Almond Ave @ Las Olas Blvd. Furthermore, since this intersection carries significantly less traffic than the two major intersections, it will receive a higher crash rate than the other two and rank higher in the problematic location list.

The third method eliminates the above shortcoming by assigning the "nearest" node ids to traffic crashes before aggregation. With each traffic crash assigned to its nearest node, number of crashes at nodes can be accurately derived through database summary. Again, three methods were explored in this project to assign the "nearest" node ids to traffic crashes.

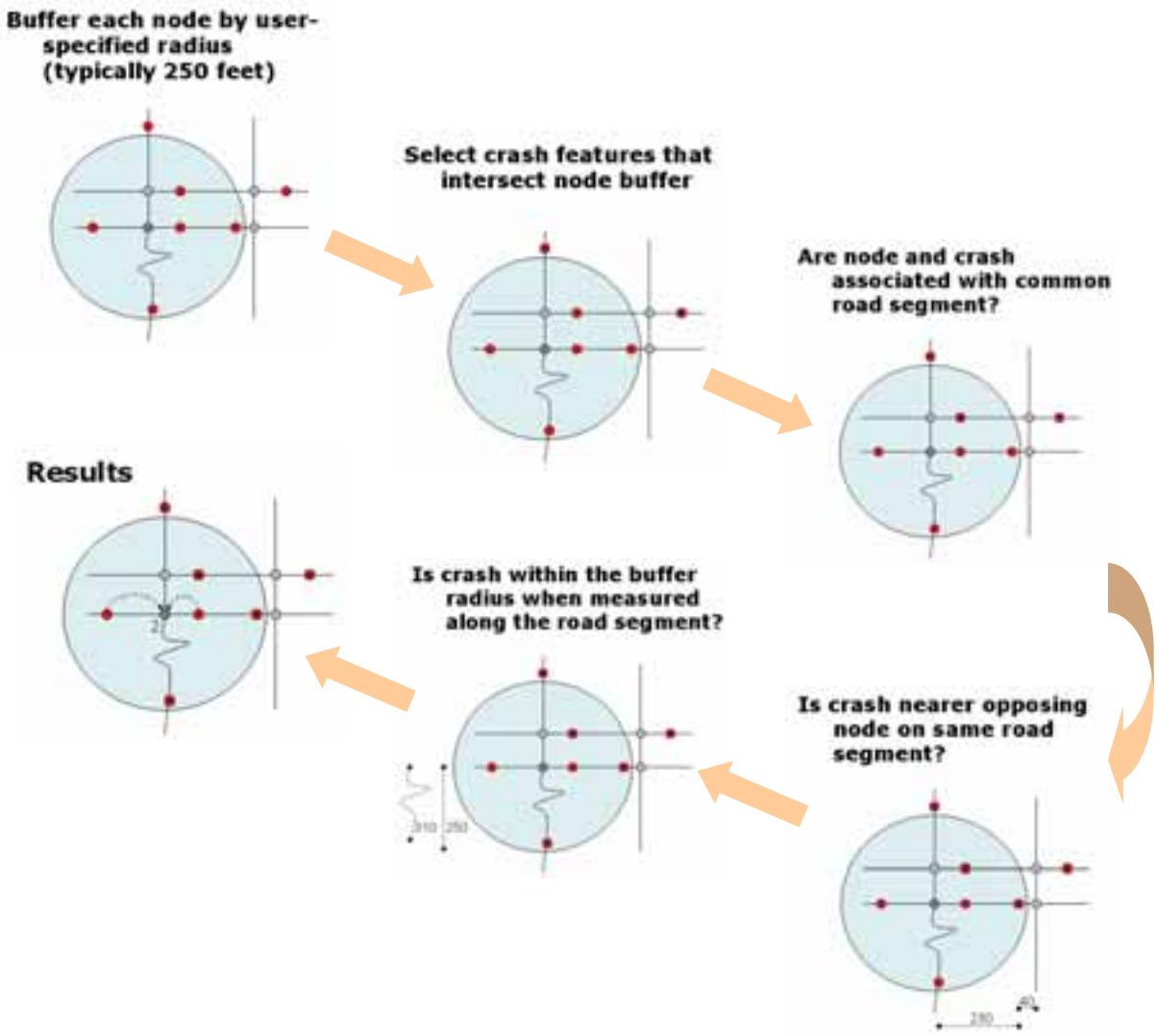
The first method used the out-of-box ArcGIS "Near" tool from the Analysis Toolbox. The "Near" tool determines the distance from each point in the input features to the nearest point or polyline in the near features, within the search radius. Although the "Near" tool provides a very simply and quick solution of finding the nearest node for a crash, it has a shortcoming of using "air distance" search instead of tracing along the road network. As illustrated below, fourteen (14) traffic crashes on Wiles Road were incorrectly assigned to the node at the intersection of Allamanda Drive @ Palm Drive because their are closer to that node by air distance than from the intersection of Wiles Road @ University Drive and the intersection of Wiles Road @ Cypress Drive.





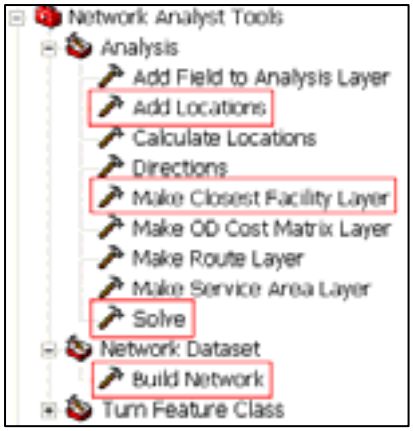
The 'Near' tool requires ArcInfo license. In absence of such license e.g. for users that have ArcView or ArcEditor licenses the same results can be achieved by performing a spatial join of intersection nodes to crash points. This method has the same "air distance" problem shown above.

The second method involved the development of a "Crash Aggregation" tool using ArcGIS VBA. This process accomplishes the Node ID assignment and aggregation of crashes to intersection nodes at the same time. This tool produces similar Node ID assignment results as a third method that uses the ArcGIS Network Analyst. However, it provides a stand-alone solution and does not require additional cost to purchase the ArcGIS Network Analyst Extension. Below is a flow chart of the logic used in developing the "Crash Aggregation" tool. At one point, a traffic crash will not receive a node ID assignment if it's outside of the search radius of any node. If a different search radius is desired, the entire process will have to be repeated. The tool was further generalized to assign node ids without the distance limitation and it provides an optional buffer distance used to aggregate crashes within a desired distance.



Further, a personal geodatabase schema was designed to build a transportation network of multiple segment and node layers that supports the "Crash Aggregation" tool. The relationship class of geodatabase provides a perfect "bridge" to tie all road networks, road features, and traffic crashes together into a geo-relational database framework. A master node table was created and relationships were established between the master nodes and several network data layers including the Broward Streets layer, the FDOT node layer, the traffic signal inventory, the priority intersection inventory, the roadway inventory, an a "Points of Interest" point layer. The BCMPO maintains priority intersection and roadway inventories with traffic volumes and other roadway characteristics. By relating the traffic crashes and the intersection and/or roadway inventory through the master node ids, further crash analyses such as crash rate calculation can be streamlined.

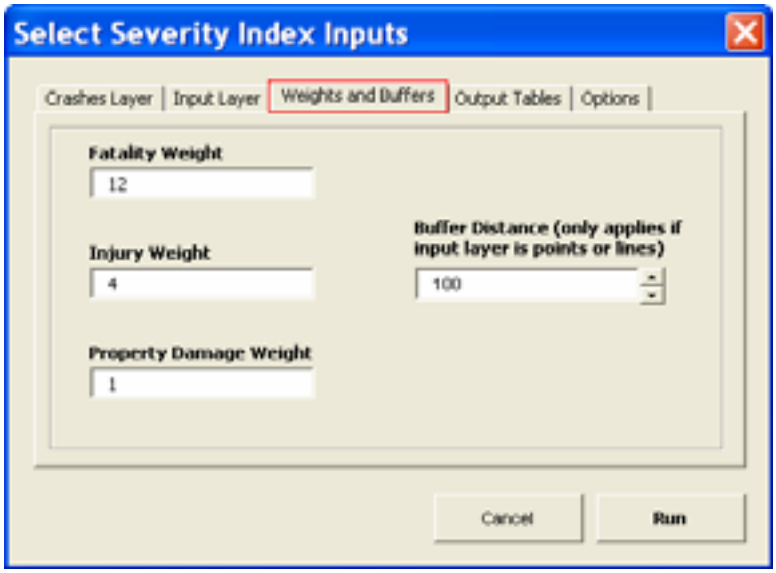
During the course of developing the "Crash Aggregation" tool, Network Analyst Extension became available when ArcGIS 9.1 was released. A third method was then developed to assign node ids to traffic crashes using the "Closest Facility" function from the ArcGIS Network Analyst Extension. The "Closest Facility" function finds the best route from incidents (traffic crashes) to their closest facilities (intersection nodes) and returns the travel cost (distance) for each route. The same technique was also used in matching the FDOT nodes and the Broward Streets nodes.



The steps taken to assign street node ids to traffic crashes using the "Closest Facility" function of ArcGIS Network Analyst are as follows:

- Copy the Broward Streets line layer into a new or existing feature dataset in a geodatabase;
- Create a new Network Dataset from the Broward Streets layer; this will add a "junction" layer into the feature dataset; Build the network;
- Create a node layer from the "junction" layer; use "spatial join" to identify number of street segments at each node and delete nodes with less than 3 segments; this will create an "intersection" node layer to be used as "facilities";
- Create a new "closest facility" session in ArcMap and load the point layer of mapped traffic crashes as "incidents" and the "intersection" node layer as "facilities";
- Solve the "closest facility" problem and the resulting "routes" layer will provide a table of "incidents", their closest "facilities", and the "impedance" (distance in this case).

Finally, this project added an interactive GIS tool to calculate the crash severity index and crash rate for street segments or nodes representing BCMPO unfunded projects. This tool can assist in assigning priorities to projects based on traffic crash patterns. The user can rapidly create a crash severity index for segments or nodes of a street network as well as a corridor or an area. The tool selects traffic crashes within the buffer distance of the input layer, aggregates them by crash types, calculates the crash severity index and crash rate, and generates output tables. By joining the output tables and the input layer, crash severity index and crash rate results can be classified and rendered in a map format. Point, line and polygon types are all acceptable as project inputs and the user has the flexibility of choosing the variables that define the formula for crash severity index calculation as illustrated by the graphic at left.



The formula for calculating the crash severity index for each project multiplies the aggregated count of each crash type within a selected buffer distance of the input feature by a weighting factor determined by the user in the GUI. Three crash types are considered in the formula: fatal ($count_{FA}$), injury ($count_{PI}$), and property damage only ($count_{PDO}$). The sum of all three crash types is then divided by the total number of crashes.

Crash Severity Index:

$$SI = \frac{A \times count_{FA} + B \times count_{PI} + C \times count_{PDO}}{total\#ofcrashes}$$

The formula for calculating the crash rate is the total number of crashes (C) within a certain user defined distance of an input feature (segment or node) divided by the product of the total years encompassed by the crash data (Y) and traffic volume (AADT) at each feature (M). Traffic volumes must be provided to obtain node or segment crash rates but is optional for polygon type projects. For nodes and polygons, M is the AADT multiplied by 365 days divided by 1,000,000. For segments, M is the product of segment length, AADT figures, and 365 divided by 1,000,000. If the project type is polygon and the traffic volume is not applicable, then the area of the polygon is used as M in the crash rate calculation.

Crash Rate:

$$CR = \frac{C}{Y \times M}$$

Implementation

ArcGIS Model Builder was used in this project to build models to automate procedures, preserve knowledge gained in the past, and provide an easy environment for future modifications. For example, one of the steps in the "DHSMV data preparation" model selects crash records occurred in parking lots or private properties by applying a pre-defined query on the "crash_address" field. The pre-defined query now contains 22 different variations of coding parking lots, driveways, etc. A user can easily open that element in the model and modify the query with different or new entries. By using the "DHSMV data preparation" model, one can perform by a click of the button a task that typically takes days.

Visual Basic for Applications (VBA) is object-oriented programming language that can be used to customize an existing computer program using an existing set of objects, each associated with properties and methods. The advantage of VBA for customization is that many programs, including ESRI's ArcMap, contain a VBA interface and are configured to work with VBA scripts. ESRI's ArcObjects is a set of objects used in the operations of ArcGIS desktop applications.

Several tools were developed in this project using VBA for ArcGIS. One example is the "Crash Severity Index" tool. The Crash Severity Index Tool was created using specific ArcObjects that control procedures such as the selection of features and manipulation of attribute tables. The "Crash Severity Index" Tool VBA program is organized around a form that is the graphical user interface (GUI). Tabs on the form contain user defined inputs and commands, such as the "run" button, that are linked to procedures in the VBA code. Program modules group subroutines and specific functions arranged thematically according to task. For example, one module contains code related to all geodatabase operations, another deals with functions that join tables to feature datasets.

A class module, which provides a way to encapsulate common procedures and to create new class objects, was included to handle output feature summary tables. Ultimately, class modules can help create cleaner code that can be debugged more efficiently.

Finally, all the tools and models developed for this project were consolidated into a toolbox, as illustrated at left, for easy access and execution.

Results

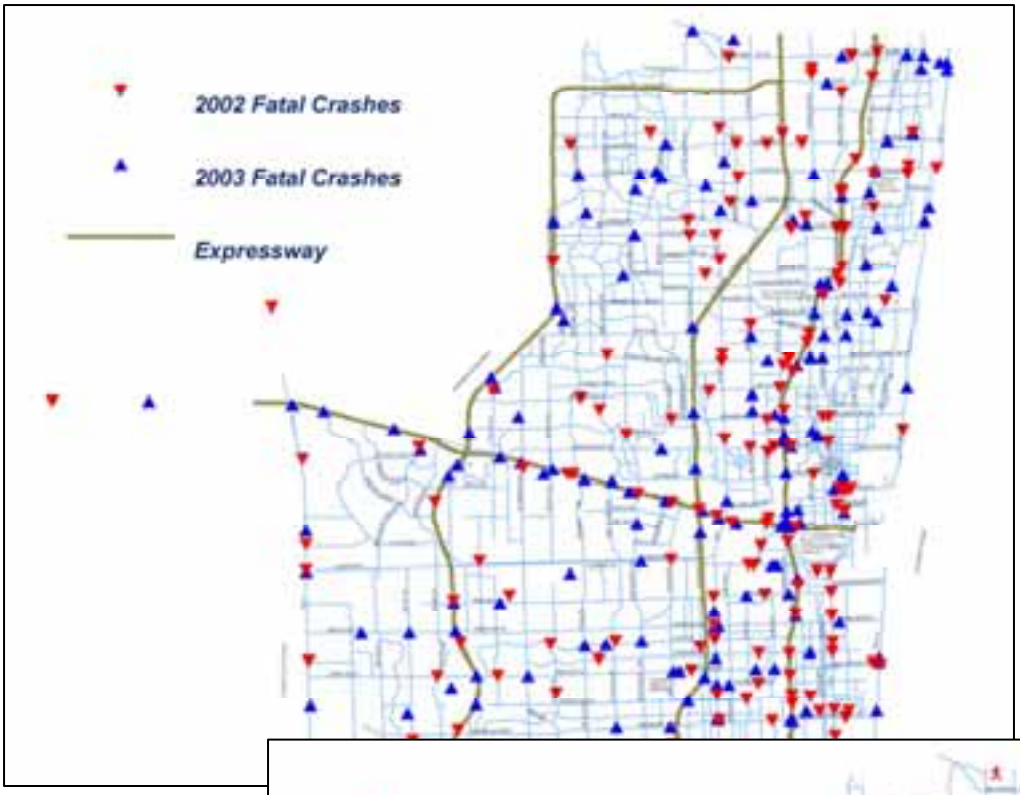
Prior to this project, the BCMPO was looking into the creation of a regional traffic crash data center. The regional traffic crash data center will house staff and resources to enter and maintain all of the traffic crash reports in a centralized database. Such a system will cost an estimated $250,000 per year to operate. Efforts to secure funding for such a system were fruitless. By using existing data available from DHSMV and automating the traffic crash mapping and analysis, the BCMPO is now able to conduct crash analyses needed for safety programs.

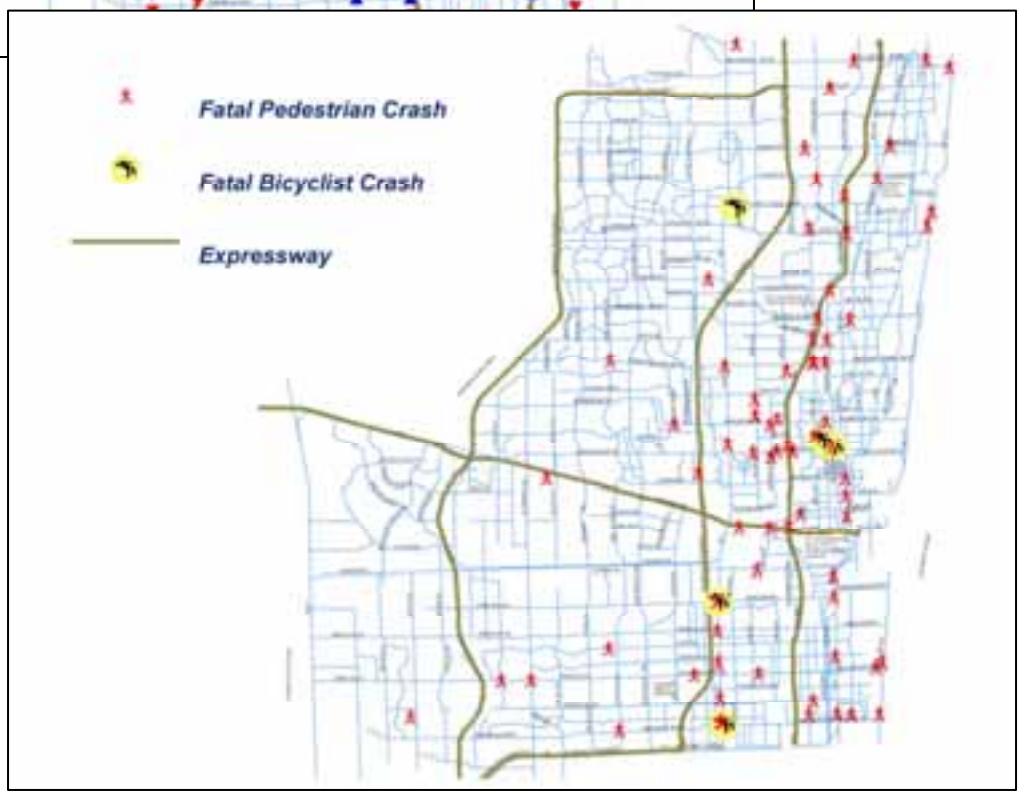
Below are some examples of how the mapped traffic crashes have been used by the BCMPO in crash pattern analyses.

The intersection of NW 19 Street @ NW 31 Avenue experienced the highest crash rate in 2002 and 2003 when compared to other intersections that are not part of the state road system. This location is currently under further safety review for improvement needs.

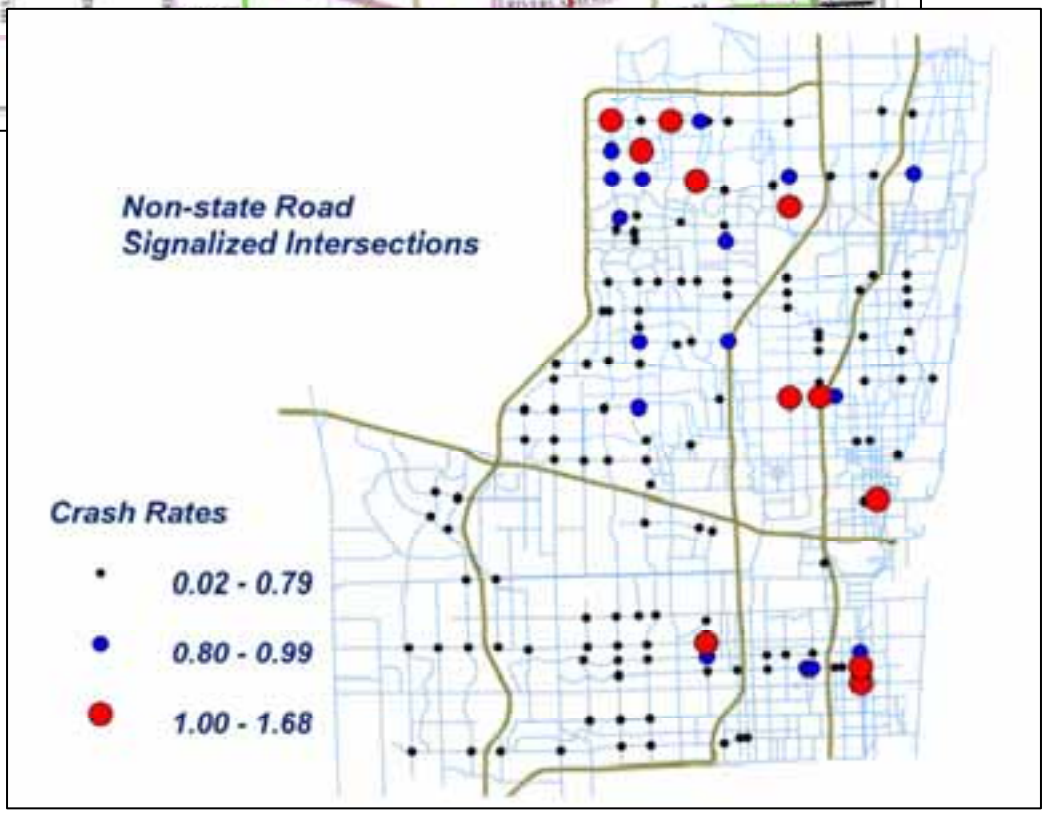
Year 2002 and 2003 Traffic Crashes Along NW 19 Street between SR 7 and NW 31 Avenue

Fatal traffic crashes occurred in 2002 and 2003:

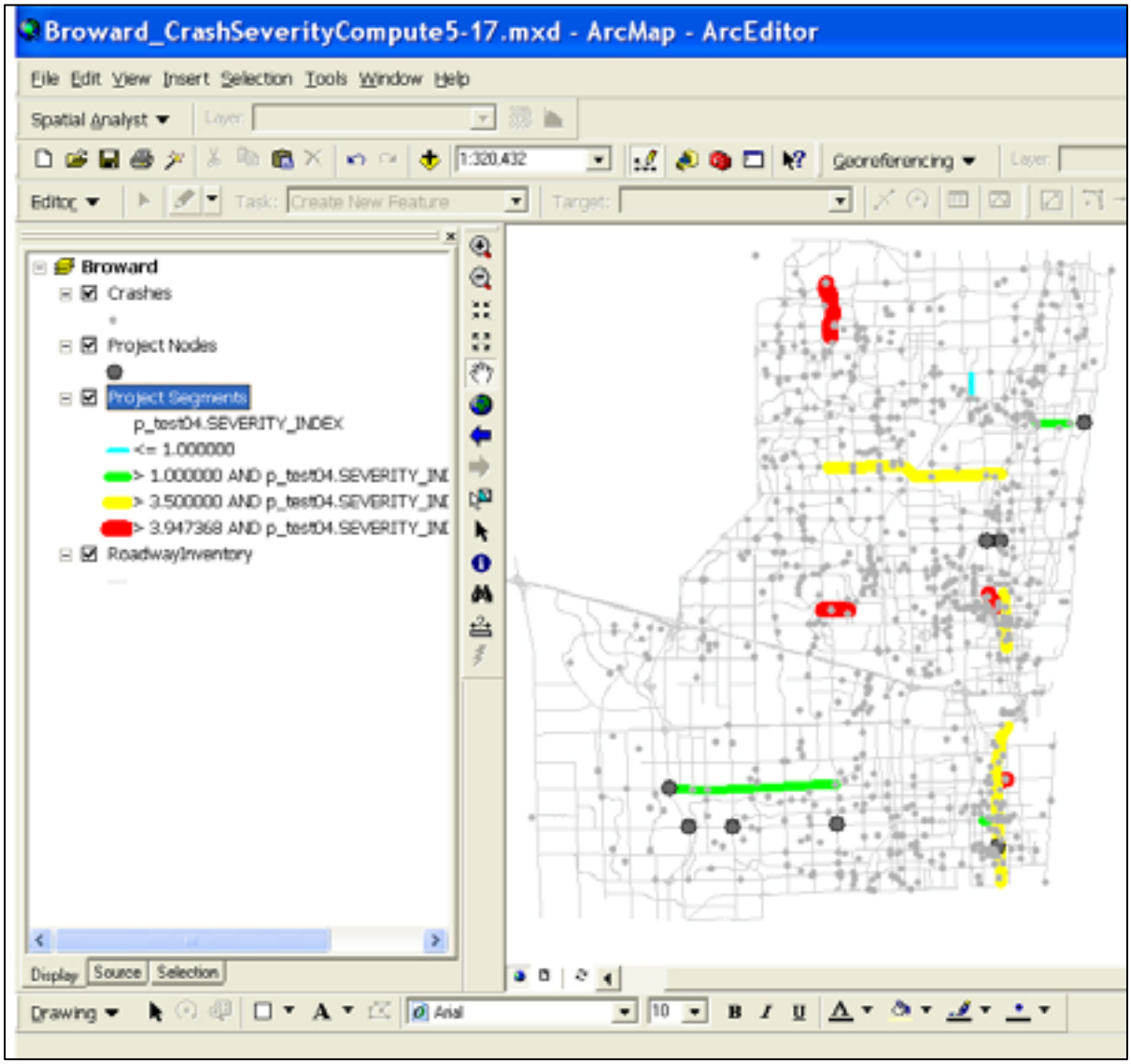




Crash rate comparisons for roadway segments and intersections:



Preliminary results of the crash severity index scores for the BCMPO unfunded roadway projects, classified into four categories:



Lessons Learned

ArcGIS Model Builder

One important lesson learned while building *large and complex* models was the need to develop a method on paper before its implementation in model builder. By clarifying the major steps of the model on paper, the time to develop and especially to modify the model can be reduced. This also provides a reference system to validate the results of the model by comparing them with the results generated by the manual method.

A second important lesson was that the use of hard coded paths during the model development may lead to problems when the model variables are converted to parameters. One typical case is the problem with hard coded query expressions. In Model Builder the query expressions don't change when the input variable names are changed. The ideal solution to this problem is to export the model into a Python or VB script and modify it accordingly to build queries that change based on the values of the input parameters. An alternative solution implemented in this project was the duplication of the input data into a new temporary geodatabase to control the values of the input variable. For example, the data layers that are queried are copied to a temporary geodatabase with a predefined name which is used during the remainder of the model.

Incremental development and testing is essential in building large models. This method proved very efficient in developing the batch geocoding model to find and debug problems in isolated part of the models.

Element/variable naming and the visual layout of the model should be considered in the early stages of the model development and not at the end of the process. This helps in making the model self explanatory and easy to read and debug. Visual grouping and labeling of logical parts of the model contribute to efficient development, testing, distribution and maintenance.

Perhaps the most problematic issue encountered during the development and especially during the iterative modifications of a large model was its problematic size management by the ESRI software. For example, for the crash data preparation model, it was noticed that a model of about 2 MB original size was increased to 11 MB and later to over 30 MB when relatively small changes were made to the model. This impaired significantly the development work by increasing dramatically time required for the opening and the saving of the model. For example when the model size grew to over 30 MB it took about 15-20 minutes just to open the model for modification.

One method to work around this serious problem is to copy all the elements of the model into a new blank model. This reduces the size of the model back to its original size, however the custom layout and the original labels are lost. One

needs to do additional work to bring the visual organization of the model to its desired layout and add the lost labels. This typically doubles the model size, however the new size is still manageable unless further changes are made to the model. A tip received through users' forum is to delete the tbx history file before opening the model. This solution didn't prove very useful. Another solution is to break the model into several small models and link them together. In conclusion, at this time, ArcGIS Model Builder does not provide a clean solution to this problem.

VBA and ArcObjects

Development of the VBA tools with ArcObjects called for creative problem solving, especially when bugs in ESRI's code caused objects to not work as expected. One of the unresolved issues in the "Crash Severity Index" tool concerns the failure of the ArcMap's selection tool on feature layers after the conclusion of the "Crash Severity Index" tool procedure. This problem may be associated with the recent 9.1 version of ArcMap since it could not be reproduced in version 9.0.

As with any application development, time required to complete the coding is usually a tough one to estimate. In some cases new ideas were discovered during the process and they may require change of methodology or additional coding. As a result, this project was extended for four months beyond its original plan of six to eight months.

Future Studies

Throughout this study, it was identified that most of the problems were a derivation of poor quality of crash data locations and multiple, non-standardized street centerlines.

To improve the quality of the data input, a likely long-term solution is the use of GPS and data entry with mobile computers in the field. Such efforts are currently under experimentation in Florida; however, there is no single standard statewide system. In addition, the implementation of such solution requires investments and training and subsequently is unlikely to be fully operational in the next few years – considering the number of agencies involved in crash reporting throughout Broward County.

A major step beyond data standardization for a countywide integrated crash database would be the creation of a web-based data entry and data sharing system for crash data. Such system has the advantages of centralized data processing while providing municipalities immediate access to up to the minute crash data.

Broward County would host an internet server with a server-based application that would allow all agencies responsible for crash data (sheriff's office and municipalities) to enter crash data on an enterprise level central database using a web client. The server would contain a RDBMS system (e.g. Oracle or MS SQL Server or any other system that the County uses) in which will reside a fully normalized database suitable to hold both attribute and spatial location of crash data. The application will contain functions for data entry, data query, analysis, mapping and reporting. Each agency will be able to access the system through appropriate privileges. The geocoding of crashes will be done implicitly in the data entry phase using validation rules to match the County street network. For special cases that cannot be validated, the interface will provide an ArcIMS viewer to allow the user to consult the County street map enhanced by aerial photography to accurately determine the location of crashes. The user interface would also provide a standard format for data entry, minimizing errors and eliminating the need to compile and process municipality-level crash data for the County. Software on the server could also generate weekly or monthly crash reports to pinpoint hotspots of crash activities at the municipal and county level. Due to its client-server architecture, the system can be accessed by a much larger audience that needs crash data thus eliminating the need to print and mail paper reports, cutting the cost of printing and mailing as well as the overall time involved.

Such system has several benefits. The agencies will not need to own proprietary standalone software for crash data, the data will be available to all interested parties as soon as it is entered in the database which will make the statistics and the analysis available instantaneously, the standardization of the information is implicitly ensured by design through the data entry, the data will be synchronized with the county GIS street centerlines which will be used to validate crash locations and ensure nearly 100% match rate and subsequently will eliminate the need for traditional geocoding while ensuring data accuracy and database integrity. In addition, the system will be compatible with field data collections which with time will replace the data entry component of the system.

Acknowledgements

This project would have been impossible without the cooperation from various agencies involved in traffic crash reporting and safety. The following agencies have provided data and assistance that contributed to the success of this project:

- Florida Department of Highway Safety and Motor Vehicles;
- Florida Department of Transportation;
- Broward County Urban Planning and Redevelopment Department, Planning Services Division;
- Broward County Sheriff's Office;
- City of Coral Springs,

- Town of Davie;
- City of Fort Lauderdale;
- City of Margate;
- City of Miramar;
- City of Pembroke Pines;
- City of Plantation;
- City of Pompano Beach (served by BSO); and
- City of Sunrise.

This project was completed with assistance from the following researchers and students of University of Florida:

- Kevin Johns
- Eric Kramer
- DeAndrae Spradley
- Nathaniel Wingfield

Finally, encouragements from others played a key role in the publishing of this paper:

"*You have done a lot of valuable groundwork in cooperation with others. When you do report on some particularly noteworthy uses of the new database, others may be very interested in how you were able to assemble the information. Keep up the good work.*" – from someone with a lot of experience and wisdom.

Contact Information

Lina Chen Kulikowski
Information Systems Analyst
Broward County Metropolitan Planning Organization
Broward County Transportation Planning Division
115 S. Andrews Ave. #329H
Fort Lauderdale, FL 33301-1818
Phone: 954.357.6610
lkulikowski@broward.org

Dr. Ilir Bejleri, Ph.d.
Assistant Professor
University of Florida, Geoplan Center
431 Architecture
Gainesville, FL 32611
Phone: 352.392.0997 ext 432
ilir@ufl.edu