A Case Study in Spatial Misclassification of Work Zone Crashes

Grady Carrick (Corresponding Author)
Florida Highway Patrol
7322 Normandy Blvd.
Jacksonville, FL 32205
904-705-8046
904-695-4110(fax)
gcarrick@ufl.edu

Kevin Heaslip
University of Florida
512 Weil Hall
PO Box 116580
Gainesville, FL 32611
352-392-9537 ext. 1539

Siva Srinivasan
University of Florida
513-A Weil Hall
PO Box 116580
Gainesville, FL 32611
352-392-9537 ext. 1456

Xiaoyu Zhu
University of Florida
512 Weil Hall
PO Box 116580
Gainesville, FL 32611

Word Count: 5,919 + 5(figures and tables) * 250 = 7,169
November 2008

Submitted for Presentation and Publication at the 88th Annual Meeting of the Transportation Research Board, Washington, D.C., 2009

ABSTRACT
Studies associated with work zone crashes are often based on law enforcement traffic crash reports. Work zone crashes are typically segregated from the larger crash data set by special coding on reports that describe the crash as being in or near roadway construction. A spatial representation of the street and nearest intersecting roadway fields of the crash report and the boundaries of the work zone provide an idea where the crash occurred in a GIS environment. When the street location of the crash report is compared to the coded work zone field a validation of the location information is possible. Incongruence between the spatial location and coded report location indicates that more than 1 in 3 work zone crashes in the case study are misclassified.

INTRODUCTION

The traffic crash report filed by police officers typically forms the basis for examination of safety on our roadways. When these reports are aggregated, individual report elements have the potential to illuminate factors associated with vehicular collisions. Unique circumstance crashes, like those occurring in work zones, are typically segregated from the larger traffic crash report data set by means of report codes specific to work areas. Many states employ a simple method (i.e. "yes or no" or "in work area" or "nearby a work area") on their crash form to determine whether a crash is in or near a work zone. This simple method creates ambiguity in reporting that can require more detailed examination in order to have confidence in the data.

Analysis of work zone crashes assumes accuracy in reporting by police. Accuracy in the spatial classification of crashes requires an accurate assessment of the crash as occurring within or outside of a work zone. Misclassification of work zone crashes has the potential to distort crash analysis because it can improperly include or exclude sample cases. In small-sample scenarios, like work zone crashes, such inclusion or exclusion of data can be problematic because of the implications that the results could form policy or budgetary guidance.

Since whole state or national crash data sets can be quite large, examining the issue of reporting accuracy can be complicated. By focusing narrowly on work zone crashes, and additionally the aspect of spatial accuracy, one can adequately narrow the scope of such a proposition. Using a case study format provides additional boundaries that are beneficial for control purposes, with the understanding that such a constriction also bounds conclusions that are produced.

The fundamental issue of this research centers on the accuracy with which crashes are classified as work zone crashes. By comparing the location of crashes and their work zone codes, some determination of accuracy can be made for the case study work zone. This paper presents a literature review that highlights previous work in crash reporting, work zone crash data, and use of Geographic Information Systems (GIS) in crash reporting. Following the literature review the methodology for the research is presented which includes the data collection procedures and the data analysis procedures. To conclude the paper a summary of all research activities are presented as well as recommendations to improve crash data collected within work zones.

LITERATURE REVIEW

Police reports account for the majority of the information regarding motor vehicle accidents. Unfortunately, these reports are often subjective and fail to include important aspects of the crash locations due to form restrictions or human error.

Hall and Lorenz (1) concluded that the "frequency of accidents in construction zones is substantially greater than indicated by the accident record system." Data from the Highway Safety Information System first identified differences in crash report statistics in the mid-1990s. From 1998 to 2000, the Fatality Analysis Reporting System (FARS) indicates a "statistically significant dependence between the way in which work zones are denoted on a state's crash report form and the percentage of fatalities that are coded as occurring in a work zone" (2). Ullman and Scriba (2) analyzed work zone fatalities nationwide and concluded that existing data may underreport the number of fatalities that occur in work zones by as much as 10% due to differences in how information about a work zone crash is captured on standard state crash reporting forms.

Shinar, et. al. (3) found that police data were the most reliable for objective characteristics of crashes such as location, number of passengers, and time. Police data was least reliable for the more subjective aspects of crash reporting such as road character, severity, and road surface composition. Shinar, et. al. (3) also found that police reports often included little information "regarding the presence of driver factors, human conditions and states, and vehicle and environmental road factors and deficiencies." The study also found that, while police officers were competent in assessing human and vehicle factors

contributing to accidents, they were less than competent in assessing the environmental factors which may have contributed to motor vehicle crashes.

Aside from work zone crashes, police reports are often inaccurate in reporting other aspects of motor vehicle crashes. Farmer (4) concluded that police reports are often inaccurate when assessing the severity of injuries resulting from motor vehicle accidents. While the rates of identifying those killed or not injured are usually correct, the varying degrees of non-life threatening injuries are often miscalculated by the reporting officer. Farmer cites Greenberg (5) in asserting that "manner of collision, vehicle identification number (VIN), and seat belt use all have a high discrepancy rate between the police report and reality with the most common error made by police in coding manner of collision was to assign single-vehicle crashes to one of the multiple vehicle categories." This study also found that police reports severely overestimated the use of seat belts. The examination of the literature has shown that there are many areas where subjectivity in crash reporting can breed ambiguity in the crash data. Two technologies being employed to lessen the subjectivity in crash reporting are Global Positioning Systems (GPS) and Geographic Information Systems (GIS).

Geographic Information Systems (GIS) may be defined as defined as "a collection of hardware and software that is used to edit, analyze, and display geographical information stored in a spatial data base (6)." FHWA (6) outlines the advantages of GIS systems over more traditional referencing and analysis methods including the ability to:

- Quickly milepost a crash that was mileposted incorrectly or had no milepost provided.
- Evaluate the problem using the spatial relationships available from the graphical displays compared to the traditional methods of tables and plots.
- Produce presentation graphics and descriptive plots from within the software as opposed to exporting results to another software package for such production.
- Develop an on-line capability to respond to concerns and questions about specific locations.
- More accurately report crashes and roadway features when combined with global positioning system (GPS) technology, thus making the analysis more reliable.
- Conduct corridor analyses by automatically linking adjacent or nearby routes together within defined zones.
- Incorporate non-traditional data bases that are available within GIS, such as land use, zoning ordinances, and population characteristics, into problem identification and evaluation studies.

In addition, the system allows traffic engineers to access various types of supplemental such as scanned versions of an officer's handwritten crash report information without leaving their desks.

Smith (7) states that "GIS has its strength in providing capabilities to model the physical proximity of spatial features particularly the flexibility in modeling spatial objects to suit particular user needs or application requirements." Capabilities in GIS can be used to improve the quality of the linear reference data and, among other data sets, if construction zone data is available, GIS data can be integrated with accident records to provide a reality based snapshot of work zone related crash incidents.

The Centre for Advanced Spatial Analysis (8) argued that "there has been little attempt to merge road traffic incident reduction and road traffic policing within a spatial context partly because of the different organizational structure of the police authorities and because some authorities do not employ analysts for the data because the emphasis of the force and its goals do not concern road traffic policing." This study also asserts the difficulty in showing the importance of data collection and what analysis of the data can do to enable them to improve policing.

Austin (9) argues that current methods for gathering crash data are not accurate enough because police have many duties to perform at a crash scene. In the case of incident analysis, a GIS can locate an incident as an object), as well as a line (8). Austin (9) concludes that a GIS based system will be more valid in reporting crash data and, if adopted by multiple agencies, will save time and money.

METHODOLOGY

The use of case study methodology serves to manage the study of work zone crashes temporally and spatially. The use of a case study is important because it provides the opportunity to examine a sample of data in a detail that would not be possible with the whole population of crashes in the data set. This research methodology included the collection of crash reports in a select work zone location, the depiction of those reports spatially, and subsequently comparison of the *reported* location of the crash with coded "work area" field to the spatial work zone boundaries.

Crash Data Collection

In examining a specific work zone location, as is the case with this study, the statewide crash report data proves difficult to use. Reporting thresholds and issues surrounding timeliness of data pose obstacles. To address these shortcomings, the research team was able to acquire crash data at the local agency level, thereby procuring a data set that could be examined in a timely manner and with proper depth.

Like all U.S. jurisdictions, the State of Florida uses a standardized format for police reporting of traffic collisions. The Florida Traffic Crash Report Form is produced by the Florida Department of Highway Safety and Motor Vehicles (DHSMV) for use by all law enforcement agencies in the state, including municipal, county, and state agencies charged with investigating traffic crashes. Electronic versions of the report format are also approved by the DHSMV, in cases where agencies and officers complete crash reports with in-car computers (10).

Florida State Statute 316.066 establishes a threshold for when a traffic collision must be reported to the police. Any crash in Florida is required to be reported if there is bodily injury, death, or an apparent amount of property damage of $500 or more (11). For crashes that involve death, personal injury, DUI, leaving the scene, or disabling vehicle damage, the Florida Traffic Crash Report: Long Form (HSMV 90003) is required. In all other cases, the use of an abbreviated report format, referred to as the "Short Form" report (HSMV 90006), is allowed. The ratio of "long" and "short" format reports is fairly even for all crash reports statewide. For calendar year 2006, Florida reported 256,200 traffic crashes, involving 370,035 drivers (12).

Obtaining traffic crash data from the statewide repository has the limitation of only including "long form" reports. Therein lies the first obstacle of using the dataset, *reporting thresholds*. The second obstacle of using the statewide crash database involves *timeliness of data*. There is generally a 6 to 18 month lag in data due to input and processing. For some crash analysis, this is not insurmountable, but for the safety analysis aimed at implementing countermeasures it potentially eliminates any possibility for a timely solution. This lag in crash data is particularly challenging when dealing with work zones that are in place for less time then the lag to process data. This lack of data does not provide decision makers the ability to use crash data to change the maintenance of traffic (MOT) plan in a timely manner if there is a safety hazard in the work area.

Because the statewide traffic crash database possesses some limitations for this research effort, an alternative was desirable. Given that the scope of this study is limited geographically to a single work zone in Duval County (Jacksonville), Florida, an opportunity to partner with reporting agencies exists. Another advantage in selecting Jacksonville for this research was that only two law enforcement agencies investigate traffic crashes, the Florida Highway Patrol (FHP) and the Jacksonville Sheriff's Office (JSO).

Obtaining actual crash reports from the agencies charged with investigating the crashes allows for a timely acquisition of data. Reports can be obtained in weeks, rather than months. Additionally, both "long form" and "short form" crash reports can be collected from the reporting agencies. Such a methodology is quite unique, and fortunately for the research team, the scope of the study was limited enough to make the proposition palatable for the agencies involved.

From a legal standpoint, the research team is conscious of obligations to be good stewards of the crash report data, agreeing to abide by applicable laws relating to crash reports and confidentiality. Florida State Statute 316.066(3)(c) establishes that crash reports are exempt from public disclosure for a period of 60-days (11). For this reason, *no report obtained by the research team was dated within the 60-day timeframe* wherein the report would be subject to such limitation. In addition to abiding by this law,

the team further refrained from using individual names, addresses, or other identifying data in any data entry or analysis. While this is not a legal obligation, such a voluntary measure reinforces a commitment to data stewardship. No privacy issues were implicated in the acquisition and use of crash report data for this project.

The crash report data set needed for analysis of work zone crashes in or near the I-95 Trout River Bridge project is bounded by date and location criteria. Filters ensure that only crashes occurring between July 1, 2006 and June 30, 2007 are part of the data set. Additionally, only crashes that occurred on I-95 northbound between State Street and Dunn Avenue, and I-95 southbound between Airport Road and Golfair Blvd. are considered for inclusion. Additional exclusions eliminate collisions occurring on exit ramps, even though they may otherwise be within the geographic bounds of the filter.

While standardized state reporting formats are used by both agencies (FHP and JSO), their approach to report completion and records management differ. These differences required flexible data acquisition approaches. An objective of the research team was to minimize imposition on the contributing agencies.

Florida Highway Patrol data is collected using in-vehicle laptop computer systems. Traffic crash report information is collected on the individual trooper's laptop, and the report is subsequently uploaded electronically via the high-speed download packet access (HSDPA/3G) communication network to the Troop "G" server. The agency completes about 500 crash reports in Jacksonville monthly. To apply the date and location filter criteria, reports for each month were manually examined to mine them for "roadway" and "nearest intersection" information that fit the criteria. The report numbers of all crashes meeting filter criteria were retrieved and printed.

Reports generated by the Jacksonville Sheriff's Office are typically handwritten by the patrol officer. The records management system (RMS) used by the agency allows for search for crash reports through the computer aided dispatch (CAD) records generated when an officer is dispatched to a crash. Location searches using this methodology are not precise, given issues associated with geocoding. A printout of 166 potential candidate report cases was produced by this process. A manual review of the location description eliminated 33 reports that did not meet the filter criteria. A total of 133 paper crash reports were requested from the JSO. Upon review of the crash reports themselves, applying location filters to the "roadway" and "nearest intersection" fields, another 47 reports were excluded.

A total of 302 FHP crash reports and 86 JSO crash reports were gleaned from the report files using the filters for date and location. The report numbers are reasonable since the FHP has primary jurisdiction for patrol on I-95 in the study area. JSO typically assists with crash investigation on the Interstate when no trooper is available. They additionally investigate most crashes on arterials that intersect with highway ramps, the source of most location filter exclusions. Of the 388 reports captured, 185 reports were "short form" and 203 "long form".

The crash reports were coded into an excel spreadsheet to allow for analysis using commercially available statistical software. This effort focused on coding that described the date, time, and location of the crash, as well as the specific code for "work area" and "roadway conditions".

It is worth noting that a separate "Engineer's Maintenance of Traffic (MOT) Evaluation at Accident Site" form is used when a traffic crash is reported in a Florida work area, or as the Manual on Uniform Traffic Control Devices (MUTCD) terms them, temporary traffic control (TTC) zone (13). Under-utilization of this reporting mechanism is very common (14), and such an engineering-type review is actually only completed in serious injury crashes (15). Because their use is sporadic, these engineering reports were not considered as a source of crash data for this study.

Crash Report Spatial Data Collection

Rendering traffic crashes spatially finds roots in "pin maps" used by police and engineers to determine where crashes occur. Other methods using links/nodes and route mileposts have also been used to create a spatial visualization of crash locations. Since the advent of geographic information systems (GIS), safety analysts have sought ways to improve upon the concept. A significant hindrance to spatially depicting crash locations has always been problems associated with location geocoding. Obtaining

accurate location data is complicated, given the fact that crashes rarely occur at conventional "address" locations that are typical in geocoding.

The research effort examines traffic crashes in a very specific way, limited by a relatively narrow geographic boundary. The size of the data set, 388 cases, is sufficiently limited to allow for manual plotting of crash locations using GIS tools with minimal investment of resources. Manual plotting of crash locations, based on traffic crash report information, is viewed as the best solution for this project.

Google Earth provides an innovative way to locate crashes using aerial images of roadways and other features. Images have the advantage over traditional two-dimensional maps in that image details allow the user to view specific to lanes, markings, and other features that may be present. When these details are compared with those reflected on the report diagram or in the narrative, greater precision is possible. In most cases, the ruler functionality in Google Earth allows for measuring distances in feet or miles to the nearest intersection.

Each crash location was plotted using the placemark functionality in the Google Earth application. Placemark names were keyed to match the investigating agency report number for later combining with the crash report data set. Google Earth was only used to take advantage of imagery to accurately define crash locations in the spatial environment. Google Earth captures the latitude/longitude of a point when the placemark feature is used. While providing a good platform for identifying crash locations as points, analysis of spatial data is better accomplished using tools like ESRI's ArcGIS. To convert the Google Earth file to a usable shapefile, the favorite place file is saved as a keyhole markup language (KML) file. Using an application named KML2SHP, the KML file can be transformed into a shapefile for use in ArcGIS. In ArcMAP, the shapefile projection is defined in a known coordinate system since the conversion process does not provide a projection.

With the crash locations plotted in a shapefile, additional layers for the Duval County boundary, roadways and waterways were added to the project. The roadway file imported lines from the Florida State Basemap that met the geographic bounds of the study area. Unnecessary roadways were screened in such a way that all roadways excluding I-95 were deleted. The northbound and southbound directions of I-95 were segregated to allow for manipulation of data points that may fall on either side of the roadway. Since each direction of the roadway was made-up of multiple line segments, those segments were dissolved and subsequently reparsed into 10 meter segments. This step allows delineation of roadway segments to be accomplished easily.

Since the work zone study area encompasses the "Trout River", creating spatial information about the waterway is visually helpful. The Florida Geographic Data Library (FGDL) possesses layers for waterways in Duval County, and appropriate layers were added to the project so that the "Trout River" is depicted. The line file was clipped to match the boundaries of the study area.

The Manual on Uniform Traffic Control Devices (MUTCD) is a national guideline for agencies that establish traffic control in temporary traffic control (TTC) zones, or work areas. Actual maintenance of traffic (MOT) indexes were obtained from the construction project contractor, wherein the placement of traffic control devices associated with the work zone were established. Point features were created to depict the location of advanced warning signs, tapper, and begin/end work area signs. From these features, specific locations were defined by selecting roadway segments to match areas found in the MUTCD; 1) before advanced warning area 2) advanced warning area 3) transition area 4) activity area 5) termination area (13). Each of these "zones" was assigned a unique symbology. When coupled with the direction of the roadway (northbound or southbound), additional characterization of the work zone locations results. Each crash record contains precise location information, as well as representation with the five "zones" described above.

The GIS data and crash report data are readily merged on the key field of "Investigating Agency Report Number". This was the "name" field in the GIS data set and the "ID" field in the crash report data set. Once merged, all of the attributes of each data set are available for the spatial and statistical analysis described in subsequent sections of this report.

ANALYSIS

The focus of this study is to determine the accuracy with which work zone traffic crash reports are coded. By examining the reporting requirements for Florida officers, the expectation for classification of work zone crashes is established. By plotting the location of the crash and the location of the work area, a determination can be made regarding how the work zone section should have been coded. Comparing the expected coding and actual coding can illuminate the degree of accuracy involved with work zone classification.

To set the stage for analysis, an understanding of the Florida reporting requirements is in order. Like every U.S. state, Florida has a unique traffic crash report form. While there is a national guideline for traffic crash report design in the Model Minimum Uniform Crash Criteria (MMUCC), these criteria are not mandatory (16). Divergent approaches to collecting work zone data are quite evident in the various state forms, and Florida deviates from the MMUCC with respect to this variable. For simplicity, the terms "work area", "work zone", and "temporary traffic control" (TTC) zone will be used interchangeably in this report. This is necessary because the guiding documents discussed herein use different terminology.

The Florida Traffic Crash Report Form contains a coded section for "work area" that is accompanied by three attribute options; 01-None, 02-Nearby, and 03-Entered. The MMUCC views work zone data with greater detail capturing elements related to location, type of work, and the presence of workers. Table 1 is a representation of the MMUCC "work zone" data elements. While the question of whether the crash occurred in or near a work zone is similar to that in the Florida report, the MMUCC views location with greater resolution further defining occurrence as before the first warning sign, in the advance warning area, transition area, activity area, or termination area.

The *Instructions for Completing the Florida Uniform Traffic Crash Report Forms* is a manual produced by the Florida Department of Highway Safety and Motor Vehicles (DHSMV) to guide officers in the completion of the crash report form. The previous version (2002) and the most recent version (2008) are identical with respect to "work area" reporting instructions. The section of the manual that directs completion of the "work area" coding for the report describes a work area and subsequently expounds upon the three unique codes that are used. "A work area is defined as that area designated by the presence of a flag person, cones, barricades, drums, arrow boards, pavement markings, signage or other traffic control used to separate workers and their equipment from other functions."(17) The description includes utility work and also stipulates that workers need not be present for the definition. While some might argue that a crash occurring before a work area, but in an upstream queue created by the work area, should be defined as a "work area" crash, this is not consistent with Florida reporting instructions. Since we are examining the input of officers, their directions must prevail, with the understanding this may be a shortcoming in the Florida reporting methodology.

Having coded the case study work zone in the GIS using descriptors 1) before advanced warning area 2) advanced warning area 3) transition area 4) activity area 5) termination area, applying the resolution of the MMUCC is possible. Applying the Florida coding to this greater level of detail requires some extrapolation. Since the conditions for "Nearby" include, "…designated work area in the vicinity of the traffic crash;"(17) applying the code to crashes that occurred in the advanced warning and taper areas appears reasonable. The requirement for coding a crash as "Entered" is, "…during the sequence of events related to the traffic crash, one or more of the involved vehicles or pedestrians were within the boundaries of a designated work area."(17) This would be analogous to the work area or activity area.

Using SPSS, cross tabulations on the variables of crash report codes and the GIS location codes produces a contingency table. When these variables are compared, misclassification of work zone crashes are apparent in each of the GIS locations as well as each of the report codes for work area. Table 2 is a reproduction of the cross tabulation using crash record counts.

Examining row and column percentages for the variables illustrates the magnitude of misclassification of work zone crashes in the case study. Tables 3 and 4 are representations of the row and column percentages for the cross tabulation of the variables representing crash report codes and GIS location codes.

Examining the column numbers (Table 2) and percentages (Table 4), we can evaluate the degree to which the work area variable on the Florida report was accurately applied by officers.

There were 99 crashes that were coded by officers as "none" for the work area variable. In 62.4% of cases where the crash report was work area coded as "none", the crash actually occurred in the advanced warning, taper, or work area. Advanced warning area crashes accounted for the majority of these instances, while 13.5% were clearly within the work area.

Eighty-seven crashes were coded as "nearby" by officers. In 16.9% of cases where the report was coded as "nearby", it should have been coded in the work area or "entered". Since the term "nearby" and its definition in the Florida instruction booklet can be interpreted very broadly, one might assume that all other locations (before advance warning, advanced warning, taper, exiting area) are technically accurate. Even given this broad classification, 16.9% remains a concerning number for misclassification.

In 196 of the sample cases for the Trout River work zone, the officer applied the code "entered", denoting that the definition of a work zone was met. The actual reported location of 52 of those crashes was before the advanced warning or in the exiting area after the work zone, representing 16.5% of the total.

Examining the row numbers (Table 2) and percentages (Table 3) can explain the actual location, with respect to the work area code applied by the officer. Twenty crashes that actually occurred in the work area were coded as "none" and an additional 25 coded as "nearby", indicates that work zone crashes appear to be under-reported by 30.4%.

40.3% of crashes occurring before the advanced warning area and 65.7% of those occurring in the exiting area were improperly coded as having occurred in the work area or "Entered". Crashes clearly before or after the work area makeup 25.5% of those coded by officers as "work area" crashes in the study area.

The most common work area reporting mistake for officers is not recognizing that the crash occurred in a work zone, 62 of 388 crashes or 15.9%. This tends to reinforce under reporting concerns of other research (1,2). The juxtaposition however is the potential "over reporting" observed in 52 of 388 crashes or 13.4%. In these cases reports were coded as "entered" however they clearly occurred outside of the work area. It appears that the rate of misclassification is both exclusive and inclusive, albeit the rate of exclusion slightly higher. The problem for crash analysis lies in properly recognizing variables that impact work zone crashes. While the raw number of crashes may not change significantly through applying these stricter rules for spatial accuracy, with nearly 1/3 of the crashes misclassified as "in" or "out" of the work zone, any resulting analysis is potentially affected. This duality of erroneous reporting may actually be more troublesome than simple under reporting.

The issue of misclassification is further increased with the consideration of "nearby" data, despite using a very liberal interpretation that includes advanced warning and taper crashes. The combined effort of misclassification involves 139 of the 388 case study crashes, or 35.8%. This figure should cause caution for anyone doing analysis of work zone crashes, where the report code is used to segregate reports from the larger set of crash data.

Analysis of misclassification between the two reporting agencies (FHP and JSO), showed that variable was not significant. Environmental variables of "wet roads" and "dark" lighting conditions were also not found to be significant.

CONCLUSIONS

Analysis in this case study of 388 work zone crashes indicates that misclassification of crashes occurring in the work zone is fairly common. Though this is only a case study, the implications for the larger, statewide dataset, cannot be determined. One might expect similar misclassification problems exist, given this sample did involve multiple reporting agencies. Three factors need to be addressed to improve this situation, report format, officer training, and spatial data collection precision. These conclusions are consistent with previous research conducted by Hall and Lorenz and Ullman and Scriba.

The Florida Traffic Crash Report Form is similar to the MMUCC guidance with respect to work zone data collection; however the use of the "nearby" attribute appears to complicate the issue. A binary

code that captures "in" or "outside" a work zone may be a better option. The greater detail achieved by addition of all MMUCC work zone data elements would likely help clarify the issue as well.

It appears that better written instructions and perhaps better officer training are also needed to improve the capture of work zone data in Florida crash reporting. The definition of a work area in the Florida instructions appears to be somewhat vague. Again, the inclusion of greater report detail and perhaps a diagram of the work zone layout may be useful.

Greater precision in the capture of spatial information is needed when comparing the location of crashes with the location of work zones. This project relied on the *reported* crash location, described by the street and nearest intersection location provided by the officer. Experience would indicate that these data elements are likely subject to the same officer reporting errors that are present in other fields. By using global positioning systems (GPS) to capture spatial information, greater precision can be achieved in locating crashes. This will enable more precise determinations concerning the relationship to work zones in the case of major projects, but not necessarily in the instance of utility or other short term activities. With the additional use of GPS to capture crash locations, the conclusions herein can be bolstered. This research has established, in a case study approach, that more than 1 in 3 work area crashes are misclassified. The methodology is replicable and may be an advisable precursor to more detailed analysis of work zone crashes.

REFERENCES

1. Hall, J. W., Lorenz, V. M., "Characteristics of Construction Zone Accidents," Transportation Research Record 1230, Transportation Research Board, Washington, D.C., 1989.

2. Ullman, Gerald L. and Tracy A. Scriba. "Revisiting the Influence of Crash Report Forms on Work Zone Crash Data." Transportation Research Record 1847. Transportation Research Board, Washington, D.C., 2004.

3. Shinar, David, John R. Treat, and Stephen T. MacDonald. "The Validity of Police Reported Accident Data." Accident Analysis and Prevention. Volume 15, Number 3. Pgs. 175-191. Great Britain. 1983.

4. Farmer, Charles M. (2003) 'Reliability of Police-Reported Information for Determining Crash and Injury Severity', Traffic Injury Prevention, 4:1, 38 — 44.

5. Greenberg, L. (1996) *Police Accident Report (PAR) Quality Assessment Project*, Report No. DOT HS 808 487, National Highway Traffic Safety Administration, U.S. Department of Transportation, Washington, DC.

6. Federal Highway Administration "GIS- based Crash Referencing and Analysis Systems." US Department of Transportation.. Washington, D.C. 1999.

7. Smith, Richard C. "Implementing GIS-based Highway Safety Analyses: Bridging the Gap." http://gis.esri.com/library/userconf/proc00/professional/papers/PAP888/p888.htm

8. Centre for Advanced Spatial Analysis. "Review of Current Practices in Recording Road Traffic Incident Data: With Specific Reference to Spatial Analysis and Road Policing Policy." University College London. 2003. http//www.casa.ucl.ac.uk/working_papers/paper67.pdf

9. Austin, KP. "The Identification of Mistakes in Road Accident Records- Part One: The Use of Geographic Information Systems." Institute of Transport Studies. University of Leeds. 1993.

10. State of Florida. Florida Statutes; Chapter 316.068 Crash report forms. State of Florida, 2007

.

11. State of Florida. Florida Statutes; 316.066 Written report of crashes. State of Florida, 2007

12. .State of Florida, Department of Highway Safety and Motor Vehicles (DHSMV), *Traffic Crash Statistics Report 2006*. 2007, DHSMV, 8 January 2008. http://www.hsmv.state.fl.us/hsmvdocs/CS2006.pdf .

13. United States. Manual on Uniform Traffic Control Devices: For Streets and Highways. [Washington, D.C.]: ATSSA/ITE/AASHTO, 2003.

14. Spanhour, Lisa and Mtenga, Primus. Analysis of Work Zone MOT Data Collection and Usage Procedures. FM No. 4066331B201. FDOT. Tallahassee, Florida. May 2002.

15. Washburn, Scott S. and Carrick, Grady T. *Data Collection Needs for Work Zone Incidents*. Final Report. Southeast Transportation Center, Univ. of Tennessee, Knoxville, TN. August 2006.

16. NHTSA (National Highway Traffic Safety Administration). Guideline for Minimum Uniform Crash Criteria (MMUCC), 2nd Edition. Washington, D.C.: National Highway Traffic Safety Administration. 2003.

17. State of Florida, Department of Highway Safety and Motor Vehicles (DHSMV), *Instructions for Completing the Florida Uniform Traffic Crash Report Form.* 2002, DHSMV, 24 July 2008. http://www.flhsmv.gov/hsmvdocs/CM.pdf

LIST OF TABLES AND FIGURES

Table 1 – MMUCC Work Zone Data Elements – Section C19

Data Element	Attribute
Was the crash in or near a construction, maintenance or utility work zone?	Yes
	No
	Unknown
Location of the crash	Before the First Work Zone Warning Sign
	Advance Warning Area
	Transition Area
	Activity Area
	Termination Area
Type of Work	Lane Closure
	Lane Shift/Crossover
	Work on Shoulder or Median
	Intermittent or Moving Work
	Other
Workers present?	Yes
	No
	Unknown

Table 2 – Cross Tabulation – GIS Location and Work Area Report Code

		work area				Total
		data lost	none	nearby	entered	
GIS the field that describes the work zone location of the crash	Before advanced warning	1	30	12	29	72
	Advance warning area	2	38	40	30	110
	Transition area	1	4	6	12	23
	Activity area	1	20	25	102	148
	Termination area	1	7	4	23	35
Total		6	99	87	196	388

Table 3 – Cross Tabulation – GIS Location and Work Area Report Code – Row Percentages

		work area				Total
		data lost	none	nearby	entered	
the field that describes the work zone location of the crash	Before advanced warning	1.40%	41.70%	16.70%	40.30%	100.00%
	Advance warning area	1.80%	34.50%	36.40%	27.30%	100.00%
	Transition area	4.30%	17.40%	26.10%	52.20%	100.00%
	Activity area	0.70%	13.50%	16.90%	68.90%	100.00%
	Termination area	2.90%	20.00%	11.40%	65.70%	100.00%
Total		1.50%	25.50%	22.40%	50.50%	100.00%

Table 4 – Cross Tabulation – GIS Location and Work Area Report Code – Column Percentages

		work area				Total
		data lost	none	nearby	entered	
the field that describes the work zone location of the crash	Before advanced warning	16.70%	30.30%	13.80%	14.80%	18.60%
	Advance warning area	33.30%	38.40%	46.00%	15.30%	28.40%
	Transition Area	16.70%	4.00%	6.90%	6.10%	5.90%
	Work area	16.70%	20.20%	28.70%	52.00%	38.10%
	Termination area	16.70%	7.10%	4.60%	11.70%	9.00%
Total		100.00%	100.00%	100.00%	100.00%	100.00%

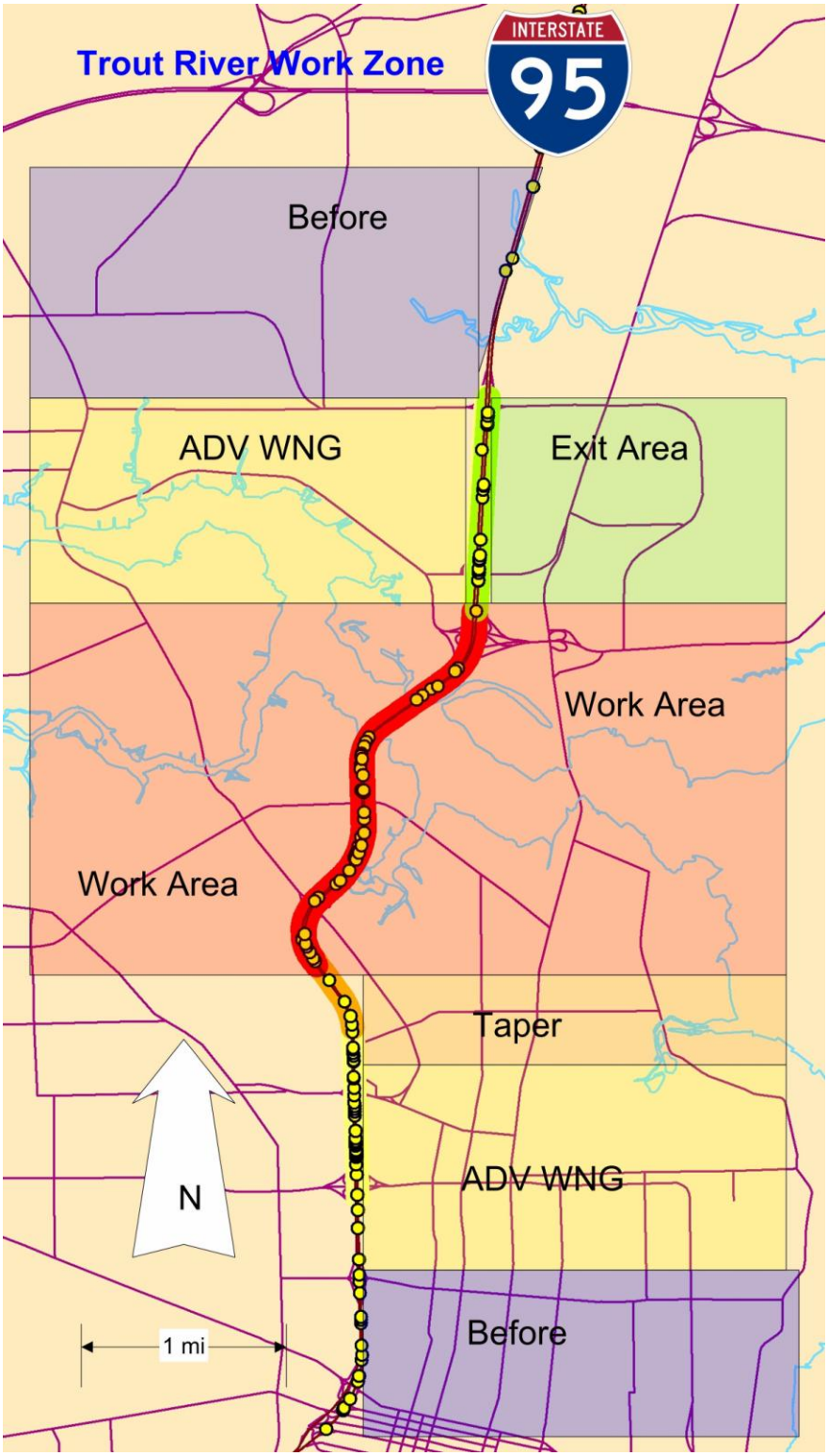

Figure 1 – Spatial Representation of Misclassified Crashes