A SPATIAL ANALYSIS OF GEOCODED FARS DATA TO IDENTIFY INTERSECTIONS WITH MULTIPLE OCCURRENCES OF FATAL CRASHES

Kevin M. Majka[1]*, Louis V. Lombardo[2], Barry Eisemann[3], Alan J. Blatt[1], Marie C. Flanigan[1]

1. CUBRC, 4455 Genesee Street, Buffalo, New York 14225, USA, 1-716-632-7500, majka@cubrc.org
2. National Highway Traffic Safety Administration, Vehicle Safety Research, Intelligent Technologies Research Division, USA
3. National Highway Traffic Safety Administration, National Center for Statistics and Analysis, USA

ABSTRACT

Initiatives such as the Cooperative Intersection Collision Avoidance System (CICAS) and the Vehicle Infrastructure Integration (VII) program may provide new opportunities to prevent crashes through advanced vehicle communications. The design and implementation of these programs will benefit from information on the location and frequency of serious injury and fatal intersection crashes. This paper examines intersections in the U.S. using data from the Fatality Analysis Reporting System (FARS) and found that relatively few intersections have had multiple occurrences of fatal crashes from 2001 through 2004. This investigation is the first of its kind to utilize geocoded FARS data for safety research.

KEYWORDS

Fatal crash, Intersection, Spatial Analysis, GIS

INTRODUCTION

In 2002, motor vehicle crashes were the leading cause of death for every age from 3 through 33 [1]. During this year, the total number of fatal crashes in the U.S. reached 38,491 while the number of crash-related deaths reached 43,005. Because of the many young lives lost, motor vehicle crashes ranked 3rd in terms of the years of life lost only behind cancer and heart disease. Intersection crashes in particular, accounted for over 9,600 fatalities in addition to a significant number of serious injuries, estimated in the hundreds of thousands [2]. Although fatal intersection crashes are considered relatively rare events given today's traffic volumes, they still produce a significant negative impact on society. Improvements to infrastructure and the development of safety technologies and systems offer the opportunity to reduce the impact of intersection crashes.

Initiatives such as U.S. Department of Transportation's Cooperative Intersection Collision Avoidance System (CICAS) program and the Vehicle Infrastructure Integration (VII) program may provide new technology and systems to prevent crashes through vehicle-to-vehicle and roadside-to-vehicle communications. The VII Program is working toward deployment of such communications capabilities to prevent vehicles from leaving the roadway and to enhance safe travel through intersections. The CICAS-V program (where 'V' stands for Violations) is aimed at preventing violations of traffic signals and stop signs by

providing drivers with a warning that they are about to commit a violation. The CICAS program also has a stop sign assist component which focuses on countermeasures to assist driver decision-making related to accepting gaps at stop signs at rural high-speed intersections. In addition, CICAS includes a signalized left turn assist component which focuses on countermeasures to assist driver decision-making related to making left turns at signalized intersections. Both components complement the CICAS-V program and are aimed at other potential safety advances at intersections.

The design and implementation of these programs can benefit from information on the location and frequency of *both* serious injury and fatal intersection crashes. The research reported in this paper however, is limited to fatal intersection crashes at signals and stop signs. Analyses of non-fatal crashes such as injury crashes of varying severity and "property damage only" crashes are beyond the scope of this paper, in part, for lack of national geocoded data on such crashes.

It is the purpose of this paper to identify and examine intersections in the United States that have had multiple occurrences of fatal crashes during the years of 2001 to 2004. The data that were analyzed in order to determine the locations of these intersections came from the Fatality Analysis Reporting System (FARS). FARS is a collection of files documenting all qualifying fatal crashes since 1975 that occurred within the 50 states, the District of Columbia, and Puerto Rico. To be included in this census of crashes, a crash had to involve a motor vehicle traveling on a trafficway customarily open to the public, and must result in the death of a person (occupant of a vehicle or a nonmotorist) within 30 days of the crash. In addition, the National Highway Traffic Safety Administration (NHTSA) has an ongoing effort to geocode these data so that it may be more effectively used in spatial analysis.

METHODOLOGY

Using 2001, 2002, 2003, and 2004 FARS crash data (that were previously geocoded by NHTSA's National Center for Statistics and Analysis) for the 50 states and the District of Columbia, we have identified intersections in the United States that have had more than one occurrence of a fatal crash per year in any one of the calendar years 2001, 2002, 2003, or 2004; or within the four year time span of 2001 to 2004. We also further classified these crashes by identifying the traffic control device, the roadway function class, and whether the roadway was part of the US National Highway System.

Intersections where multiple crashes occurred were identified by using spatial and attribute queries of the NHTSA FARS database after the database was imported into a Geographic Information System (i.e., ESRI's ArcGIS 9.1 software). The processing of the data included first, selecting those crashes that were defined as 'non-interchange, intersection' or 'non-interchange, intersection related' crashes, which include all types of intersection crashes such as straight crossing path and left turn across path, and exporting this subset of crashes as the population of crashes used in this study for the 2001 through 2004 time frame. Analyses of the various crash typologies within this population (such as straight crossing path and left turn across path) are possible but beyond the scope of this paper, although these crash typologies may be examined in the future using the geocoded FARS data.

Next, the points representing the crash sites were converted to a North America equidistant conic projection so that measured distances would remain accurate regardless of latitude.

Buffers with a 100 foot radius were then created for each crash, *for each year independently*, to allow for any error in the accuracy of the original geocoding and accommodate the rather large distance between two crashes that may have occurred on opposites legs of an intersection. After the buffers were created, the interior boundaries between overlapping buffers were dissolved so that crashes occurring within 200 feet of each other are grouped together. The buffers for each year need to be created independently so that the buffer for a crash that occurs in one year does not automatically tie together two adjacent crashes of distances greater than 200 feet which occurred in another year. For example, two crashes in 2002 occur 253 feet apart and are therefore determined to be two separate crashes with separate buffers; in 2003 however, another crash occurs in between the two 2002 crashes such that the buffer of each of the three crashes overlap and could be dissolved into a single crash location (Figure 1). This approach is reasonable if a multi-year time frame is being examined, but inappropriate if individual years are being studied and compared.

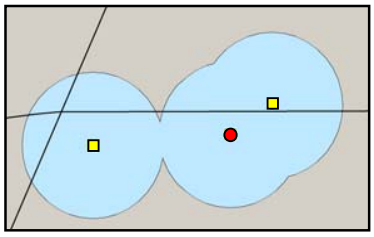

Figure 1 - The red dot represents a FARS intersection crash in 2002, the yellow squares represent crashes in 2003, and the large blue circles are the 100 foot buffers

The crash locations were then 'spatially joined' to the dissolved buffers. This allows the FARS attributes associated with each crash to also be associated with the dissolved buffer. It also enables the analyst to obtain a count of the number of crashes which occurred inside each dissolved buffer so that multiple crash locations can be identified.

The process of spatially joining each crash to its dissolved buffer was subsequently repeated for each year. Additionally, this process was performed for an aggregated data set that contained all of the crashes from the four years, i.e. buffers were created and dissolved simultaneously in order to determine multiple crash locations *over* the four years. Those dissolved buffers with 2 or more crashes (which indicated multiple crashes occurred at the buffered intersection), were then exported into a new file for analysis.

The intersections were then broken down by the type of traffic control device, roadway functional class, and National Highway membership. For each crash that was determined to be part of the 'multiple crash intersection' subset, the appropriate attributes to distinguish these types were queried from the attribute table (which contains the FARS variables). Differences may exist among attributes of crashes at the same intersection and so numbers provided are best estimates. This is due to the fact that intersection summaries are aggregated values (i.e., they are collections of crashes with individual attributes rather than the specific attributes for a single intersection). The intersections therefore may be misrepresented due to changes in attribute values over time, such as the upgrade of a traffic control device, or uncertainty inherent in geocoding and spatial calculations.

FINDINGS

Intersection and intersection related crashes at non-interchange areas were first identified from the geocoded FARS dataset. In this 4-year period, there were 138,430 total geocoded FARS crashes (over 90 percent of all FARS crashes were geocoded by NCSA from 2001 to 2004) and of these, 31,601 crashes were at non-interchange intersections or were intersection related (Figure 2). Multi-crash intersections were then found by identifying intersections that had at least one crash in each year from 2001 to 2004 and also identifying intersections that had multiple crashes over the 4 years.

Crash Type	2001	2002	2003	2004	Total
Geocoded FARS crashes	30,897	35,455	35,710	36,368	138,430
Non-Interchange Intersection Crashes	6,995	8,233	8,214	8,159	31,601
% of total	22.6%	23.2%	23.0%	22.4%	22.8%

Figure 2 – Geocoded Fatal Crashes and Geocoded Fatal Crash Intersection Crashes

Note that the tables present the findings numerically, broken down by the number of geocoded FARS crashes, the number of intersection and intersection related crashes, the number of intersections with fatal crashes by number of crashes, intersections by traffic control device, intersections by functional class, and intersections by membership to the National Highway System (NHS). In Figure 2, the 'Total' column is the row summation of the four preceding years, 2001 through 2004, and represents the total number of FARS geocoded crashes.

In figures three, and five, the final column labeled '2001-2004' is more than the row summation of the four preceding years of data. This is due to the fact that intersections that may have had less than the specified number of crashes in any one year may have had additional crashes in other years. The result is that when examining all four years together, intersections that had the specified number of crashes over the four years are added to the those intersections that had the specified number of crashes in a single year. For example, in Figure 3, the last row for "Intersections with 5 Fatal Crashes" shows that there was only one intersection that experienced 5 fatal crashes over the four year time frame. But in the individual years of 2001, 2002, 2003, and 2004, the entry is zero "Intersections with 5 Fatal Crashes" in each year. This is because in individual years, over the four year time frame, there was one or more fatal crashes at that intersection that totaled to 5 Fatal Crashes in the last column "2001-2004".

2001 Summary: We found that in 2001, of the 30,897 geocoded FARS crashes, 6,995 or 22.6% occurred at non-interchange intersections or were intersection related as coded in the FARS database in the 'relation to junction' attribute. Of these 6,995 non-interchange intersection crashes, 155 of them occurred at an intersection where at least one other crash occurred within 100 feet of the first crash. The 155 crashes that occurred in 2001 were found to have been located at 73 intersections where two crashes occurred and at three intersections where three crashes occurred (Figure 3, Figure 4). Of the 76 intersections, 14.5% had no

traffic controls, 50.0% were signalized, and 35.5% had only regulatory signs (Figure 5). Of the 76 intersections, 52.6% were rural and 47.4% were urban as defined by the roadway classification code in FARS (Figure 5). Of the 76 intersections 40.0% were on the National Highway System, while 60.0% were not (Figure 5).

Intersections with Specified Number of Crashes	2001	2002	2003	2004	2001-2004*
Total Intersections with Fatal Crashes	6,916	8,122	8,115	8,066	30,308
Intersections with 1 Fatal Crash	6,840	8,012	8,017	7,974	29,127
% of total	98.9%	98.6%	98.8%	98.9%	96.1%
Intersections with 2 Fatal Crashes	73	109	97	91	1,080
% of total	1.1%	1.3%	1.2%	1.1%	3.5%
Intersections with 3 Fatal Crashes	3	1	1	1	91
% of total	0.0%	0.0%	0.0%	0.0%	0.3%
Intersections with 4 Fatal Crashes	0	0	0	0	9
% of total	0.0%	0.0%	0.0%	0.0%	0.0%
Intersections with 5 Fatal Crashes	0	0	0	0	1
% of total	0.0%	0.0%	0.0%	0.0%	0.0%

*All intersections that had the specified number of crashes in a single year as well as those with the specified number of crashes over the entire time-period

Figure 3 – Number of Intersections with Specified Number of Fatal Crashes

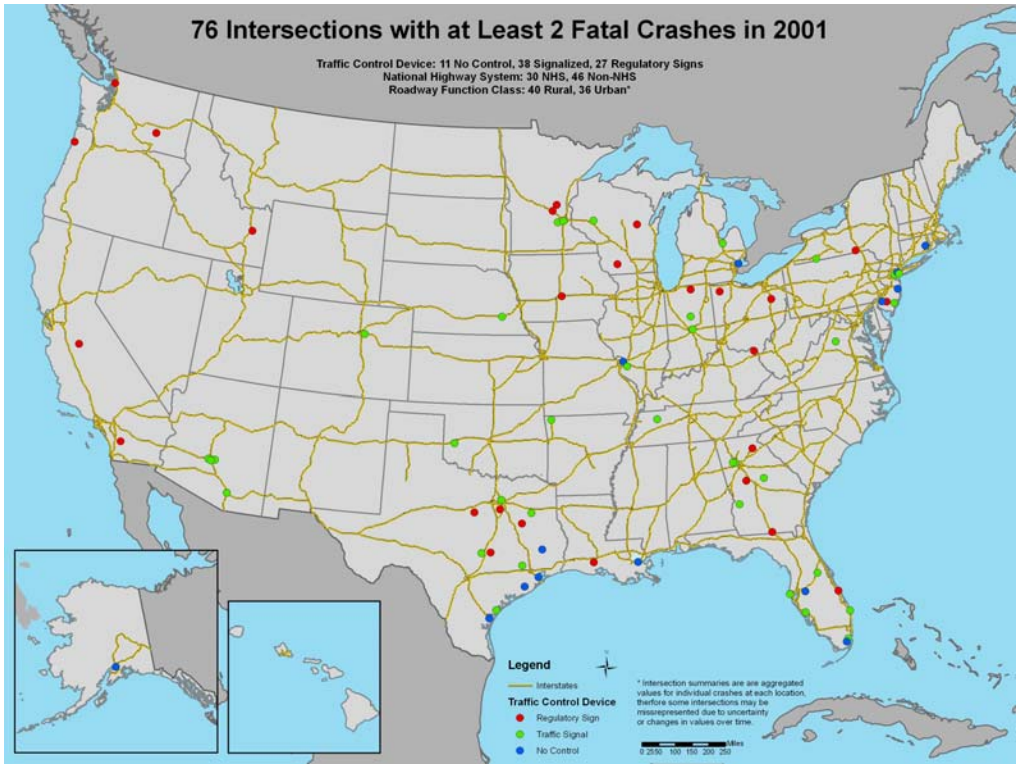

Figure 4 – 2001 Multiple Fatal Crash Intersections by Traffic Control Device

Characteristics of Mutiple Crash Intersections	2001	2002	2003	2004	2001-2004*
Total	76	110	98	92	1,181
No Controls	11	16	14	12	176
% of total	14.5%	14.5%	14.3%	13.0%	14.9%
Traffic Signal	38	52	46	43	550
% of total	50.0%	47.3%	46.9%	46.7%	*46.6%*
Regulatory Sign	27	42	38	37	455
% of total	35.5%	38.2%	38.8%	40.2%	38.5%
Rural	40	53	45	40	520
% of total	52.6%	48.2%	45.9%	43.5%	44.0%
Urban	36	57	53	52	661
% of total	47.4%	51.8%	54.1%	56.5%	56.0%
Non-NHS	46	75	69	59	798
% of total	60.0%	68.5%	71.1%	66.3%	67.9%
NHS	30	35	29	33	383
% of total	40.0%	31.5%	28.9%	33.7%	32.1%

*All intersections that had the specified number of crashes in a single year as well as those with the specified number of crashes over the entire time-period

Figure 5 – Characteristics of Intersections with Multiple Fatal Crashes

2002 Summary: In 2002 35,455 FARS crashes were geocoded while 8,233 or 23.2% occurred at non-interchange intersections or were intersection related as coded in the FARS database in the 'relation to junction' attribute. Of these 8,233 non-interchange intersection crashes, 221 of them occurred at an intersection where at least one other crash occurred within 100 feet of the first crash. The 221 crashes that occurred in 2002 were found to have been located at 109 intersections where two crashes occurred and at one intersection where three crashes occurred (Figure 3, Figure 6). Of the 110 intersections, 14.5% had no traffic controls, 47.3% were signalized, and 38.2% had only regulatory signs (Figure 5). Of the 110 intersections 48.2% were rural and 51.8% were urban as defined by the roadway classification code in FARS (Figure 5). Of the 108 intersections 31.5% were on the National Highway System, while 68.5% were not (Figure 5).

2003 Summary: For 2003 there were 35,710 geocoded FARS crashes, 8,214 or 23.0% occurred at non-interchange intersections or were intersection related as coded in the FARS database in the relation to junction attribute. Of these 8,214 non-interchange intersection crashes, 197 of them occurred at an intersection where at least one other crash occurred within 100 feet of the other crash. The 197 crashes that occurred in 2001 were found to have been located at 97 intersections where two crashes occurred and at one intersection where three crashes occurred (Figure 3, Figure 7). Of the 98 intersections, 14.3% had no traffic controls, 46.9% were signalized, and 38.8% had only regulatory signs (Figure 5). Of the 98 intersections 45.9% were rural and 54.1% were urban as defined by the roadway classification code in FARS (Figure 5). Of the 98 intersections 28.9% were on the National Highway System, while 71.1% were not (Figure 5).

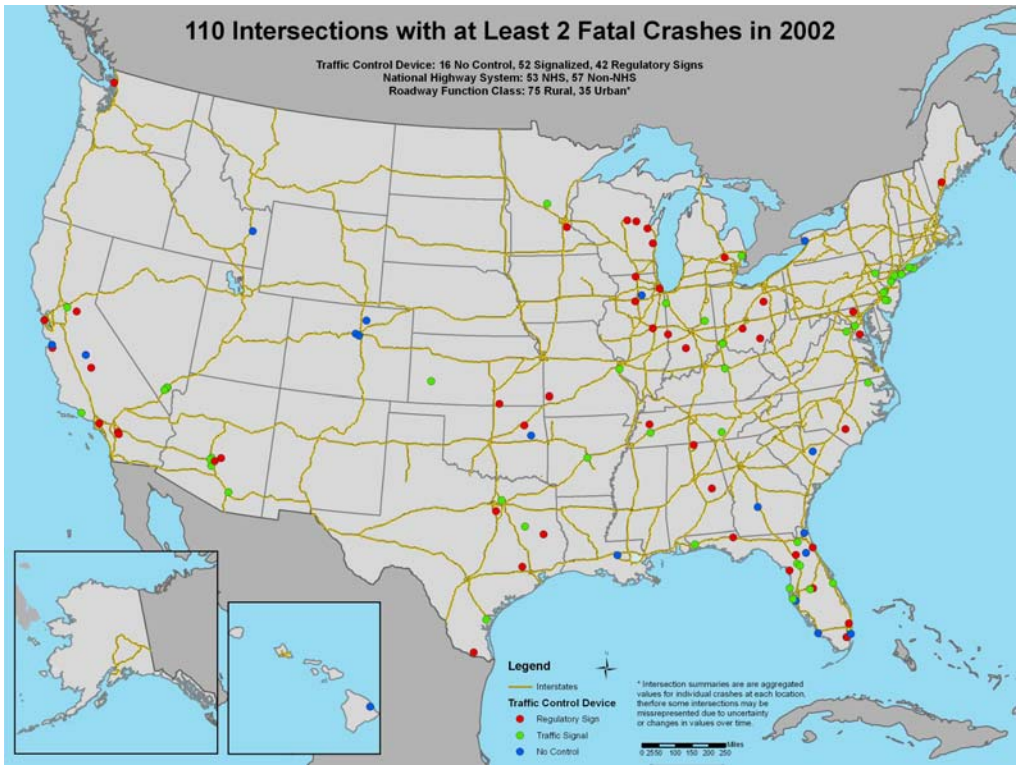

Figure 6 – 2002 Multiple Fatal Crash Intersections by Traffic Control Device

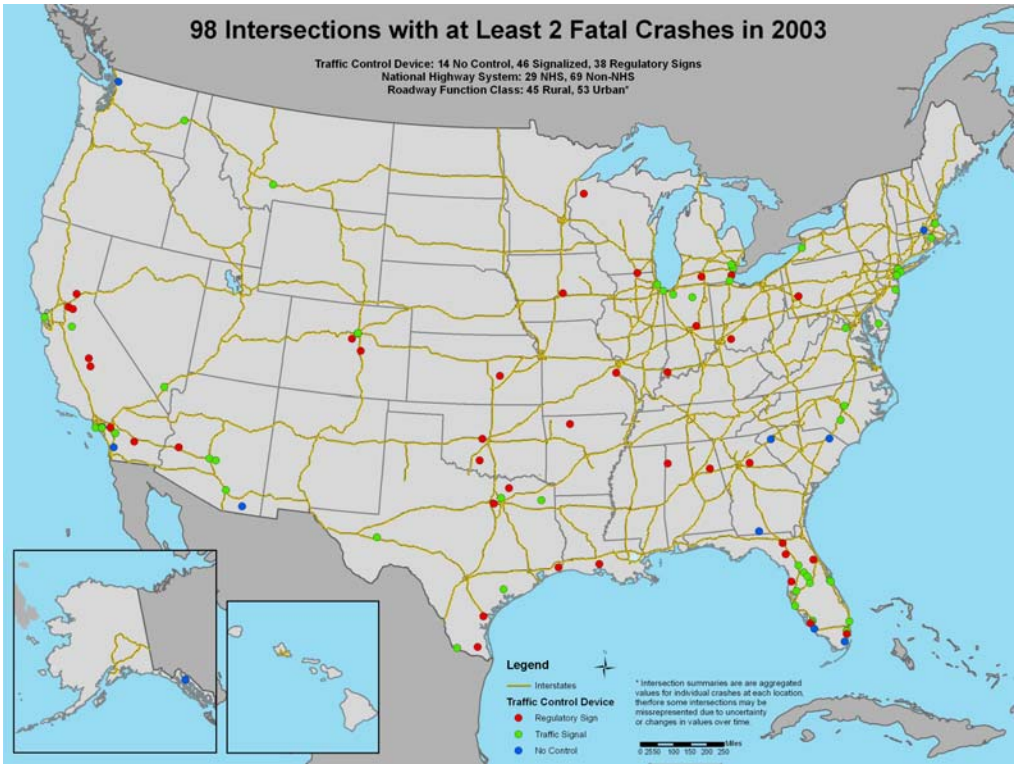

Figure 7 – 2003 Multiple Fatal Crash Intersections by Traffic Control Device

2004 Summary: 36,368 FARS crashes were geocoded in 2004. Of those 8,159 or 22.4% occurred at non-interchange intersections or were intersection related as coded in the FARS database in the relation to junction attribute. Of these 8,159 non-interchange intersection crashes, 185 of them occurred at an intersection where at least one other crash occurred within 100 feet of the other crash. The 185 crashes that occurred in 2004 were found to have been located at 91 intersections where two crashes occurred and at one intersection where three crashes occurred (Figure 3, Figure 8). Of the 92 intersections, 13.0% had no traffic controls, 46.7% were signalized, and 40.2% had only regulatory signs (Figure 5). Of the 92 intersections 43.5% were rural and 56.5% were urban as defined by the roadway classification code in FARS (Figure 5). Of the 92 intersections 33.7% were on the National Highway System, while 66.3% were not (Figure 5).

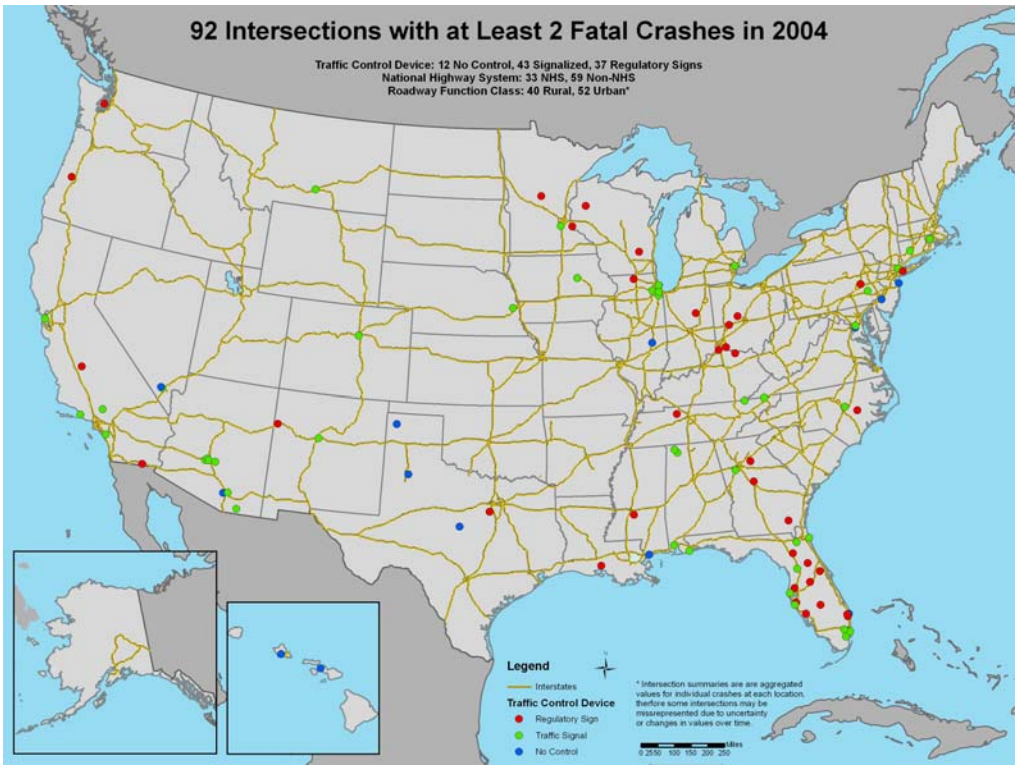

Figure 8 – 2004 Multiple Fatal Crash Intersections by Traffic Control Device

2001 to 2004 Summary: In addition to identifying intersections that had multiple crashes in any one of the four years, it is also beneficial to identify intersections that had two or more crashes at a single intersection at any time during the four-year timeline. This means that an intersection can be flagged for study if more than one crash occurred in any single year *or* if the total number of crashes occurring at that intersection during the timeframe 2001 to 2004 is greater than one.

2001 to 2004 Summary - Intersections with at Least 2 Fatal Crashes: Over the four-year time period 138,430 FARS crashes were geocoded and 31,601 crashes occurred at non-interchange intersections or were intersection related as indicated in the previous section. The identification of multiple crash intersections this time however was found by identifying intersections that had at least two crashes total, during any given year or combination of years. Of these 138,430 FARS crashes, 2,474 occurred at a non-interchange intersection where at least one other crash occurred within 100 feet of the first crash.

The 2,474 crashes that occurred from 2001 to 2004 were found to have been located at 1,080 intersections where two crashes occurred, at 91 intersections where three crashes occurred, at nine intersections where four crashes occurred, and one location where five crashes occurred (Figure 9). The addition of 1080, 91, 9, and 1 intersections sums to a total of 1,181 intersections that had multiple fatal crashes over the time period of 2001 to 2004. Of the 1,181 intersections where multiple crashes occurred 13.0% of them had no traffic controls, 46.7% had traffic signals, and 38.8% had regulatory signs. Of the 1,181 intersections 44.0% were rural and 56.0% were urban as defined by the roadway classification code in FARS. Of the 1,166 intersections 32.1% were on the National Highway System, while 67.9% were not.

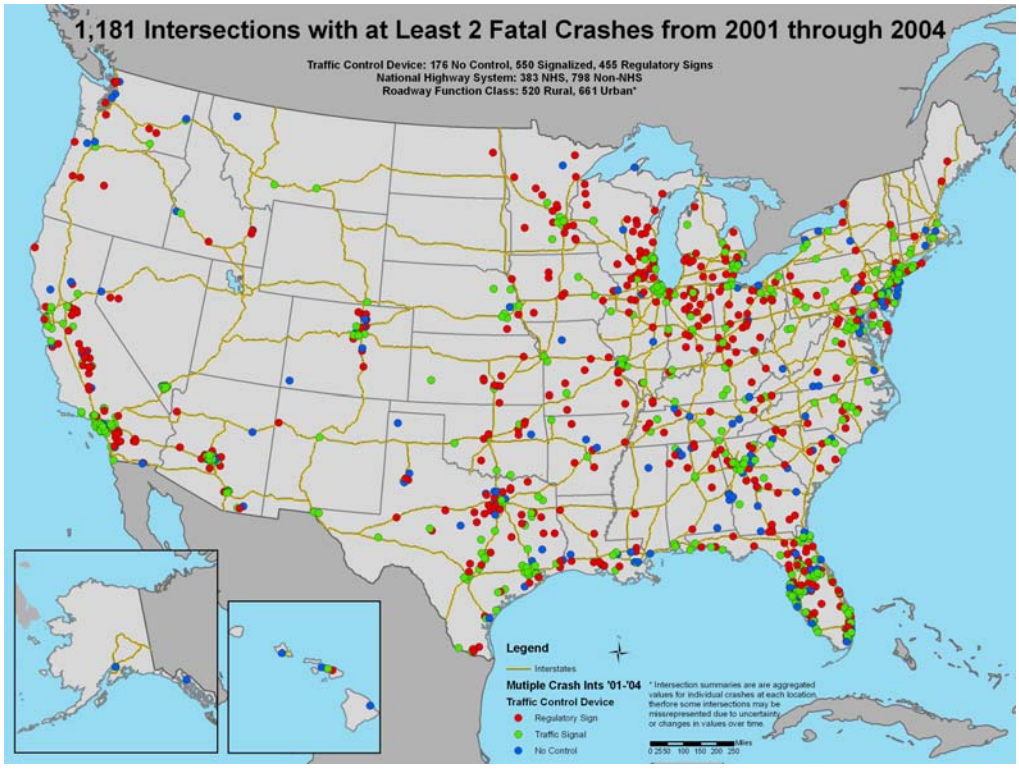

Figure 9 – 2001-2004 Multiple Fatal Crash Intersections by Traffic Control Device

2001 to 2004 Summary - Intersections with at Least 3 Fatal Crashes: By looking at the entire data set from 2001 to 2004 it is also possible to start identifying specific and manageable number of intersections for further study. In this study we have identified 101

intersections that have 3 or more crashes, 91 having three crashes, nine having four crashes, and one having five crashes (Figure 10). Of the 101 intersections where three or more crashes occurred in 2001, 2002, 2003, and 2004, 5.9% of them had no traffic controls, 53.5% had traffic signals, and 40.6% had regulatory signs. Of the 101 intersections 52.5% were rural and 47.5% were urban as defined by the roadway classification code in FARS. Finally, the 101 intersections were classified as 45.5% National Highway System members and 54.5% non-National Highway members.

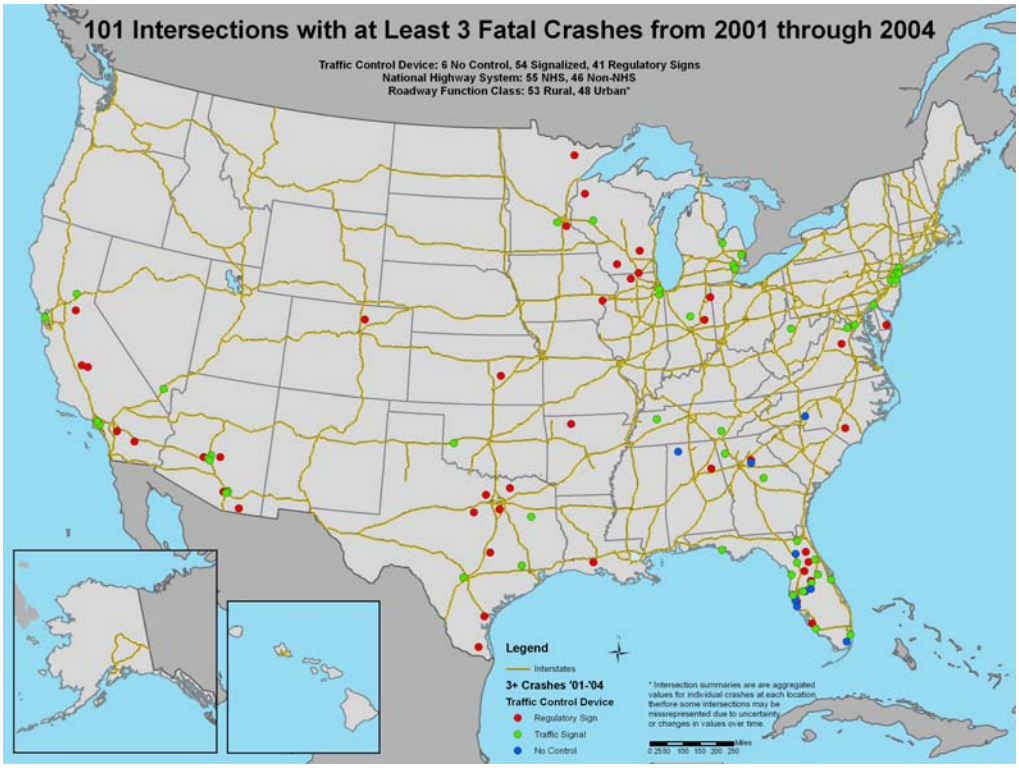

Figure 10 – 2001-2004 3+ Fatal Crash Intersections by Traffic Control Device

2001 to 2004 Summary - Intersections with at Least 4 Fatal Crashes (With at Least 1 Crash in Each Year): Of the total 138,430 FARS crashes for 2001 to 2004, 12 crashes occurred at a non-interchange intersection where at least one other crash occurred within 100 feet in *each* of the four years. These twelve crashes occurred at three different intersections, where each intersection experienced four crashes each (Figure 11); it was not possible to have an intersection with less than four crashes because we were looking for four successive years of crashes; more than four was possible but was not found in this study. Of the three intersections where multiple crashes recurred in 2001, 2002, 2003, and 2004, all had some form of traffic controls: one was signalized, and two had regulatory signs. Of the three intersections, all three were rural (as opposed to urban) as defined by the roadway classification code in FARS and none were on the National Highway System.

Figure 11 – 2001-2004 Successive Fatal Crash Intersections by Traffic Control Device

CONCLUSION

It is evident that the intersections where multiple crashes have occurred are unevenly distributed across the states, in both overall number and also when normalized by population. In terms of overall intersections, Florida ranks first with 24 intersections with more than 3 fatal crashes from 2001 to 2004, followed by Arizona with 11, Texas with 10, California with 9, Wisconsin with 5, and 21 other states with between 1 and 4 intersections with more than 3 fatal crashes (Figure 12, Figure 13).

Considering that those states with the highest number of intersections with more than 3 fatal crashes are also among the most populous US states, the overall number of intersections was normalized by the year 2000 population of that state to calculate an adjusted value. Based on these calculations, Delaware ranks first with 2.6 intersections with 3 or more fatal crashes per one million people, followed by Arizona at 2.1, the District of Columbia at 1.7, Florida at 1.5, Wisconsin at 0.9, and the remaining 21 states with values between 0.6 and 0.2 (Figure 12).

This investigation illustrates the application of geocoded FARS data to support traffic safety analysis, in general, and the VII and CICAS program in particular. Although the results of this investigation show that relatively few intersections have had multiple fatal crashes during the time period examined, the number of motor vehicle crash fatalities that occur at intersections remains very high. The continued collection and geocoding of FARS data will no doubt enhance this study and provide additional insight into the patterns and factors of intersection and intersection related fatal crashes, and hopefully provide insight and support to traffic safety managers at all levels.

Table 6. Intersections with 3 or more Fatal Crashes by State and Type

State	No Control	Signalized	Regulatory	Total	Pop 2000	Per 100K Pop.
Alabama		1	1	2	4,447,100	0.0450
Arizona		7	4	11	5,130,632	0.2144
Arkansas			1	1	2,673,400	0.0374
California		4	5	9	33,871,648	0.0266
Colorado			1	1	4,301,261	0.0232
Delaware		1	1	2	783,600	0.2552
District of Columbia		1		1	572,059	0.1748
Florida	5	13	6	24	15,982,378	0.1502
Georgia		2	2	4	8,186,453	0.0489
Illinois		2	1	3	12,419,293	0.0242
Indiana		1	2	3	6,080,485	0.0493
Iowa			1	1	2,926,324	0.0342
Kansas			1	1	2,688,418	0.0372
Louisiana			1	1	4,468,976	0.0224
Michigan		4		4	9,938,444	0.0402
Minnesota		1	2	3	4,919,479	0.0610
Nevada		1		1	1,998,257	0.0500
New Jersey		3		3	8,414,350	0.0357
New York		3		3	18,976,457	0.0158
North Carolina		1	1	2	8,049,313	0.0248
Oklahoma		1		1	3,450,654	0.0290
Tennessee		2		2	5,689,283	0.0352
Texas		3	7	10	20,851,820	0.0480
Virginia	1	1		2	7,078,515	0.0283
West Virginia		1		1	1,808,344	0.0553
Wisconsin		1	4	5	5,363,675	0.0932
Total	**6**	**54**	**41**	**101**		

Figure 12 – Intersections with 3 or more Fatal Crashes by State and Type

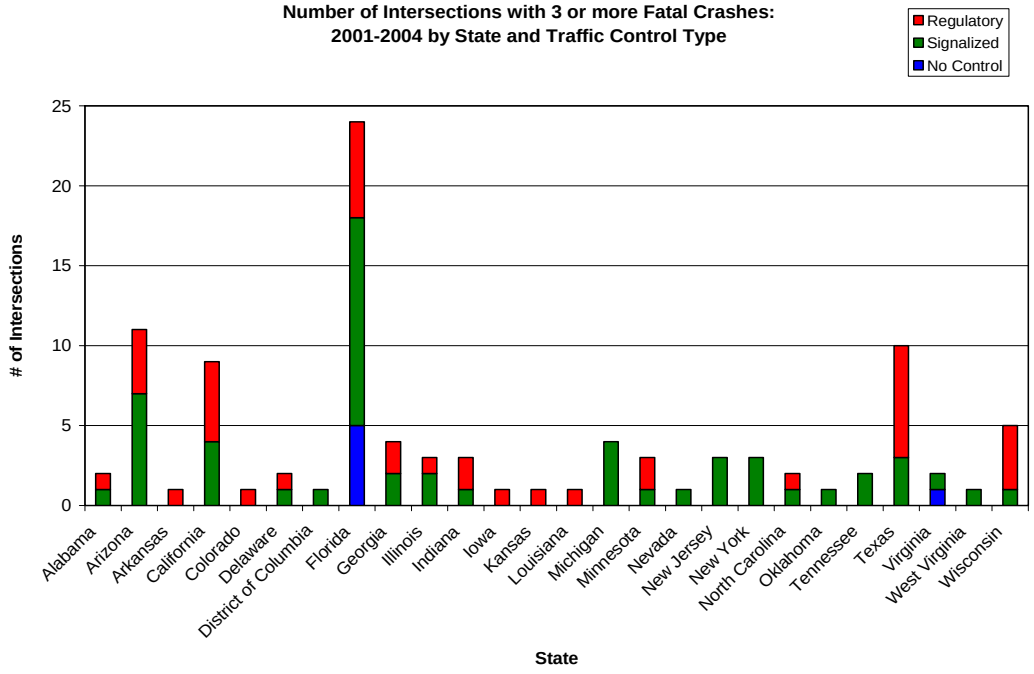

Figure 13 – Intersections with 3 or more Fatal Crashes by State and Type

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Joseph Carra at NHTSA for his continued support and interest in this project. In addition, we would also like to acknowledge Dr. James Chang and Mr. Daniel Cohen at Mitretek Systems for sharing their work and providing us with valuable insights into the analyses of fatal intersection crashes.

REFERENCES

[1] National Highway Traffic Safety Administration, NCSA Traffic Safety Facts Research Note DOT HS 809 831, January 2005

[2] The National Intersection Safety Problem, Issue Brief, Institute of Transportation Engineers, FHWA, April 2004; http://www.ite.org/library/IntersectionSafety/problem.pdf