SAFETY-BASED DEPLOYMENT ASSISTANCE FOR LOCATION OF V2I APPLICATIONS PILOT: STOP-SIGN GAP ASSIST APPLICATION

Background

V2I Applications

Vehicle-to-Infrastructure (V2I) is a component of the connected vehicles program. It is a wireless-based communication technology in which the exchange of critical operational and safety data between vehicles and roadway infrastructure is intended to help avoid crashes. Previous research commissioned by the U.S Department of Transportation (DOT) identified eight applications that can provide safety benefits. Among those applications, Red-Light Violation Warning (RLVW), Stop Sign Gap Assist (SSGA) and Curve Speed Warning (CSW) were selected for accelerated evaluation. The Concept of Operations and Systems Requirements for the SSGA application were completed. While a prototype SSGA application was not subsequently developed by the Federal Highway Administration (FHWA), this report provides information that may be used by public agencies to determine if the return on investment is sufficient to warrant application development in the future.

Stop-Sign-Gap Assist (SSGA) Application

The SSGA application is one of potentially high-value V2I applications. The intent of the application is to target crashes that result from poor gap acceptance at two-way stop-controlled intersections. These crashes include stop-controlled motor vehicles that are traveling straight or turning at an intersection. *Gap acceptance*, in this case, is defined as the process by which a driver on the minor road stops at the stop sign, judges the speed of conflicting traffic and the adequacy of "gaps" in traffic, and makes a maneuver (right or left turn) onto or across the major road through a gap. If the gap is too short in length to make the maneuver, a crash results. The intent of a SSGA application is to provide assistance to the driver on the minor road in identifying a suitable gap. One concept uses infrastructure-based instrumentation on the major road to measure gaps in the major road traffic. Such a system can transmit this information to the stopped minor road vehicle to advise the driver of the length of the gap to assist in appropriate maneuver decisions.

Vehicle Deployment

The number of vehicles equipped to receive the SSGA message will affect the system's ability to prevent crashes. As more vehicles are equipped, more crashes may be prevented. The *National Connected Vehicle Field Infrastructure Footprint Analysis* presented estimates for the speed of deployment of equipped vehicles in the nation's vehicle fleet. The deployment scenarios are described as mandates (assuming a requirement is in place) or organic (assuming voluntary installation by automobile manufacturers). Figure 1 presents three scenarios for potential deployment over a 25-year period: a 1-year mandate, a 5-year mandate, and a 15-year organic implementation. The 1-year mandate presents the most aggressive deployment with 60 percent of the vehicles equipped by year 10. The 15-year organic implementation represents the slowest deployment scenario with 20 percent of the vehicles equipped by year 10. For the purposes of this analysis, it is assumed that either a mandate or organic implementation of connected vehicle technologies in vehicles will occur beginning in 2020.

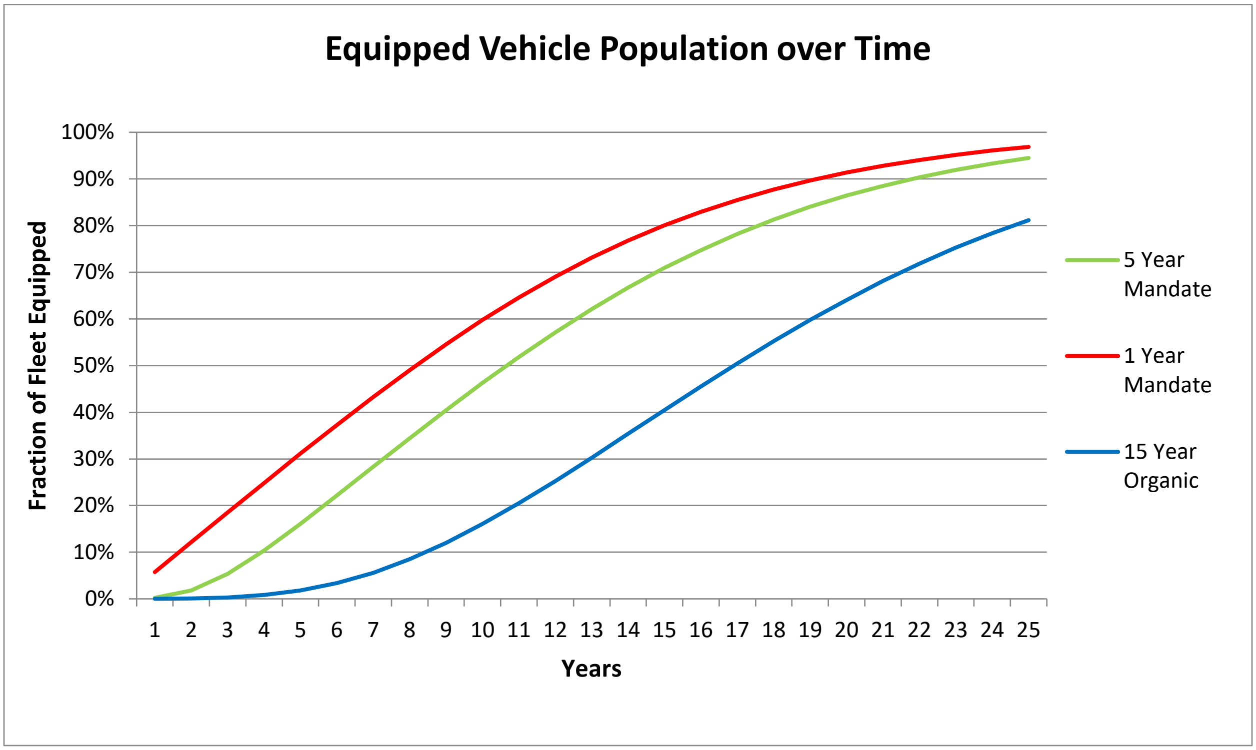

Figure 1. Three potential vehicle deployment scenarios.

(Source: *National Connected Vehicle Field Infrastructure Footprint Analysis*)

The National Highway Traffic Safety Administration (NHTSA) issued the Vehicle-to-Vehicle (V2V) Notice of Proposed Rulemaking on December 20, 2016. This NPRM proposes to mandate V2V communications for new light vehicles over a three-year period, beginning two years after issuance of a final rule.

Infrastructure Deployment

State DOTs and local transportation agencies will be the primary deployers of the infrastructure component of the V2I SSGA systems. The systems will be installed at stop-controlled intersections with the goal of preventing gap-acceptance related crashes. Therefore, their selection of locations for deployment of these systems will be primarily based on the expected occurrence of gap acceptance related crashes. These agencies need guidance in identifying locations that have experienced gap acceptance related crashes and are expected to continue to experience these crashes unless there is an intervention such as the deployment of the SSGA system.

OBJECTIVE

The objective of this effort was to develop guidance for State and local agencies on how to select locations for deployment of the SSGA applications to achieve the greatest benefit to cost ratios. This was accomplished by exploring the occurrence of target crashes, the annual fluctuations in crash occurrence by intersection, and the costs of the target crashes.

The selected identification-location method should have the following characteristics:

- Easy to implement by a State or local agency without rigorous statistical analysis.
- Applied using no more than five years of data.
- Results in the identification of those locations with the most opportunity to reduce target crashes.

This effort concentrated on understanding and characterizing the benefits of the application expressed as the comprehensive cost savings from preventing crashes based on historical crash occurrence.

DATA EMPLOYED

This analysis used intersection and crash data from the Highway Safety Information System (HSIS). The economic analysis was performed based on crash cost information from the FHWA and the Office of the Secretary of Transportation. These sources of data are described in the following sections.

Intersection and Crash Data

The intersection and crash data for this study came from HSIS. The HSIS is a roadway-based system maintained by the FHWA that provides quality data on a large number of crash, roadway, and traffic variables linked to homogeneous sections of the entire highway system under State control. It is the only multi-State database that allows for the safety analysis of roadway design factors, as it is the only file system with the capability to link roadway inventory and exposure data to crash data for a large sample of primary route mileage, and the only file system to include both roadway sections with and without crashes. It is important to note that HSIS data are only available for State-maintained roadways in each State. As such, in general, HSIS represents more rural areas, because roadways in urban areas are often maintained by a municipality.

Currently, seven States are part of the HSIS: California, Illinois, Maine, Minnesota, North Carolina, Ohio, and Washington. (Historical data from Michigan and Utah are also available, but updated data are no longer captured.) HSIS also includes the City of Charlotte. This study analyzed data for the 10 most recent years of data for California and the three most recent years of data for Minnesota.

There are six types of data files available within HSIS. All States maintain three basic files: a crash file, a roadway inventory file, and a traffic volume file. Additional roadway geometry files are also available within selected States, including a horizontal curve file, a vertical grade file, and an intersection and interchange data file. California and Minnesota were selected for detailed intersection analyses, as these States provide intersection data files, which are critical to this analysis.

The study used ten most recent years of data available for California (2002 to 2011). The study also used three most recent years of Minnesota data (2010 to 2012). More years of California data were necessary for conducting a time series analysis. All 10 years of data were no longer necessary after a recommended time frame had been established and the primary analyses only needed the most recent three years of data. Table 1 presents a summary of California and Minnesota data used for this analysis including the years of data, the number of two-way stop-controlled intersections, and the number of crashes at those intersections. The analysis excluded all-way stop-controlled intersections.

Table 1. Summary of HSIS data from California and Minnesota.

Variable	California	Minnesota
Years Analyzed (Trend Analysis)	10 (2002 – 2011)	N/A
Years Analyzed (Primary)	3 (2009-2011)	3 (2010-2012)
Number of Two-Way Stop-Controlled Intersections	7,777	3,781
Average Annual Two-Way Stop-controlled Intersection Crashes	4,563	1,660
Average Annual Fatal and Incapacitating Injury Two-Way Stop-Controlled Intersection Crashes	206	46
Average Annual Crashes per Intersection	0.587	0.439

Each intersection for California and Minnesota was characterized as urban or rural in the HSIS data. Approximately one third of the stop-controlled intersections were rural in both States.

A list of elements for each dataset is presented in Appendix A.

Crash Costs

The FHWA report *Crash Cost Estimates by Maximum Police-Reported Injury Severity within Selected Crash Geometries* provides mean comprehensive crash costs disaggregated by crash severity, location type, and speed limit.[(1)] The report is a useful reference for determining the cost of crashes and therefore the potential monetized benefits of preventing those crashes. However, the values in the report are based on 2001 dollars which are now out of date. Although not disaggregated by severity, location, and speed limit, the FHWA Office of Policy provides departmental guidance on valuing reduction of fatalities and injuries by regulations or investments. The most recent guidance was provided in the 2015 memorandum, *Treatment of the Economic Value of a Statistical Life (VSL) in U.S. Department of Transportation Analyses – 2015 Adjustment.*[(2)] These values were used to modify the detailed crash costs by applying a proportion (the ratio of the 2015 fatality and the 2001 fatality costs) to the disaggregated 2001 costs to represent the costs in terms of 2015 dollars. Table 2 presents the resulting average cost per crash by the maximum injury severity in the crash in 2015 dollars. This cost represents all speed limits.

Table 2. Average cost per crash based on maximum injury severity (in 2015 dollars).

Maximum Injury Severity in Crash	Cost (in 2015 dollars)
Fatality (K)	$9,901,946
Incapacitating Injury (A)	$533,666
Injury, Non-incapacitating (B)	$197,049
Possible Injury (C)	$110,374
Property Damage Only (O)	$18,374

METHODOLOGY

Identifying Two-Way Stop-Controlled Intersections

The focus of this application is two-way stop-controlled intersections. The California and Minnesota HSIS data files both include an intersection file that identifies the presence of intersections and provides some information on the intersection control present. The intersections included in the file represent various types of intersection control (e.g., signalized, four-way stop-controlled, etc.). In the California file the variable *traffic control type* (TRF_CNTL) was used to identify stop-controlled intersections. Those intersections identified as “Stop Signs on Cross Street Only” were used in this analysis. Additionally the intersections were screened using the junction type variable (JUNCTYPE) to exclude any junctions that were not intersections. As presented in Table 1, there were 7,777 candidate intersections in California.

The Minnesota data file is structured with information at the approach level. That is, for a four-legged intersection there is inventory information for each of the approaches. The variable *approach traffic control* (AP_CNTL) was used to identify intersections with the mainline coded as “Through or One-Way Leaving Intersection) and the minor approach or approaches coded as “Stop Sign.” The overall intersection control was confirmed with the variable *Traffic Control Devices* (TRAFCNTRL) to include intersections description as “Thru Stop Intersection.” The variable *Intersection Description Revised* (TYPEDESC) was used to confirm that the intersection geometry matched the number of approach legs. For example, 3-legged approach intersections were coded as “Tee Intersection” or “Wyee Intersection.” This resulted in 3,781 candidate intersections in Minnesota.

Characterizing Intersections

Stop-controlled intersections in California and Minnesota represent a diverse set of intersections that vary by land use, functional class, and number of approach lanes. The crash experience at the intersection of two-lane rural roads is likely remarkably different from the crash experience at the intersection of four-lane corridors in an urban environment. Poor gap acceptance at rural intersections, which tend to be higher speed, is more likely related to difficulties judging the speed of approaching vehicles.

Given these differences, a State or local agency may want to consider these types of intersections separately. Additionally, an SSGA system may be designed slightly different for a 3-legged intersection than for a 4-legged intersection. Therefore, the intersections were characterized into groups of similar intersections based on number of approaches (3 or 4-legged intersections), area type (urban or rural) and number of approach lanes.

Table 3 presents the intersection groups that the research team developed with the California data and the average number of total intersection crashes for each group. Fourteen groups were developed. Four of the groups are labeled as “others” regarding the number of lanes intersecting which combine those intersections that do not fit in the other groups. In general, those with more lanes on the major on the mainline experienced more crashes.

Table 3. Intersection groups and crashes in California.

Intersection Group	Number of Intersections in Group	Average Annual Number of Total Intersection Crashes for Group
3-legged, urban-2 mainline lanes x 2 cross street lanes	769	492.7
3-legged, urban-4 mainline lanes x 2 cross street lanes	515	476.0
3-legged, urban-6 mainline lanes x 2 cross street lanes	149	183.0
3-legged, urban others	99	92.0
4-legged, urban-2 mainline lanes x 2 cross street lanes	286	325.7
4-legged, urban-4 mainline lanes x 2 cross street lanes	316	412.7
4-legged, urban-6 mainline lanes x 2 cross street lanes	62	106.7
4-legged, urban others	40	80.3
3-legged, rural-2 mainline lanes x 2 cross street lanes	3,536	1152.0
3-legged, rural-4 mainline lanes x 2 cross street lanes	299	186.0
3-legged, rural others	196	59.3
4-legged, rural-2 mainline lanes x 2 cross street lanes	1,205	781.0
4-legged, rural-4 mainline lanes x 2 cross street lanes	218	176.7
4-legged, rural others	87	39.3

Table 4 presents similar information for the Minnesota data. Generally, groups of fewer than 30 intersections are too small to be representative. The intersections in Minnesota included a few intersections categorized as "other" where the number of lanes were not provided. These few intersections are not included in the groupings but are considered in other parts of the analysis. This resulted in a total of eight groups in Minnesota. The average annual number of crashes for each group varies more than the groups in California.

Table 4. Intersection groups and crashes in Minnesota.

Intersection Group	Number of Intersections in Group	Average Annual Number of Total Intersection Crashes for Group
3-legged, urban-2 mainline lanes x 2 cross street lanes	627	264.7
3-legged, urban-4 mainline lanes x 2 cross street lanes	132	158.7
4-legged, urban-2 mainline lanes x 2 cross street lanes	410	190.7
4-legged, urban-4 mainline lanes x 2 cross street lanes	101	139.3
3-legged, rural-2 mainline lanes x 2 cross street lanes	1,087	336.3
3-legged, rural-4 mainline lanes x 2 cross street lanes	136	127.7
4-legged, rural-2 mainline lanes x 2 cross street lanes	1,173	325.3
4-legged, rural-4 mainline lanes x 2 cross street lanes	96	93.0

Identifying Target Crashes

Target crashes for this application are those resulting from a vehicle on the stop-controlled minor approach colliding with a vehicle on the perpendicular uncontrolled major approach. These are crashes where the driver of the vehicle on the minor approach entered the intersection without an adequate gap in approaching traffic. These crashes include several types of crossing path crashes because the minor road vehicle may have been going straight, turning left, or turning right.

In the California data files, several variables were used to identify target crashes at the candidate intersections. The variable *INTS/Ramp ACC Location* (INT_RAMP) was used to identify those crashes coded as occurred at or outside an intersection. These crashes were limited to those coded as occurring within 250 feet of the intersection. Although the target crashes occur at the intersection, a slightly larger distance was used to allow for variances in how the reporting officer recorded the crash. Single vehicle crashes were excluded from the analysis using the number of vehicles involved. The variable *direction of travel* (DIR_TRVL), which provides the cardinal direction of travel of each involved vehicle, was used to identify that the crash involved a vehicle from the minor approach and a vehicle from the major approach. This was compared with the characterization of the vehicles movement using the variable *movement preceding accident* ($MISCACT_{1,2}$) which provides the movement of each vehicle preceding the crash (e.g., proceeding straight, making right turn, making left turn). The resulting set is all intersection crashes that involve one vehicle from the minor approach and one vehicle from the major approach.

Similarly, in the Minnesota data files, several variables were used to identify target crashes at the candidate intersections. The variables *diagram of accident code* (ACCDIGM) and *type of accident* (ACCTYPE) were used to identify multi-vehicle collisions that were characterized as right angle, left turn, or right turn. Similar to California, the involved vehicles' approach directions were compared using the vehicle level variable *direction vehicle was traveling* (VEH_DIR) to confirm that the crash involved vehicles were from opposing directions. The location of the crash within the intersection was determined using the variables *location of first harmful event (*LOC_HARM) and *relation to intersection* (LOC_TYPE). The variable *action prior to accident* ($MISCACT_{1,2}$) was used to eliminate unusual maneuvers such as a vehicle making a u-turn or going the wrong way into opposing traffic.

The resulting number of average annual target crashes in California and Minnesota are presented in Table 5. Using the number of intersections identified in Table 1, the number of average annual target crashes per intersection is also calculated. The table also includes the number of candidate intersections with one or more target crashes in the analysis period and the average annual target crashes for those intersections. Notably, the number of annual target crashes is very small given the number of candidate intersections.

Table 5. Average annual target crashes by dataset

Variable	California	Minnesota
Average Annual Target Crashes	822.3	176.0
Average Annual Target Crashes Per Intersection	0.104	0.047
Number of Intersections with One or More Target Crashes in Three Years	1,301	369
Average Annual Target Crashes per Intersection for Intersections Experiencing One or More Target Crashes over Three Years	0.624	0.477

Crash Severity

Minnesota and California both use the KABCO scale to identify the maximum reported injury in a crash. Table 6 presents the distribution of maximum injury severity for target crashes and other intersection crashes for California and Minnesota. The totals presented here are for three years of data. Notably, the target crashes are more severe than other crashes in both California and Minnesota with 24 percent of the target crashes in California resulting in K or A or B compared to just 16 percent of the other intersection crashes and 21 percent of the target crashes in Minnesota resulting in K or A or B compared to just 13 percent of other crashes.

Table 6. Summary of crash severity distribution for California and Minnesota data.

Maximum Reported Crash Severity	California		Minnesota	
	Target Crashes	Other Intersection Crashes	Target Crashes	Other Intersection Crashes
K (fatal)	45 (2%)	142 (1%)	14 (2.6%)	23 (0.5%)
A (incapacitating injury)	111 (5%)	321 (3%)	19 (3.6%)	82 (1.8%)
B (non-incapacitating injury)	417 (17%)	1,308 (12%)	89 (16.9%)	459 (10.3%)
C (possible injury)	734 (30%)	2,614 (23%)	145 (27.5%)	961 (21.6%)
O (property damage only)	1,127 (46%)	6,871 (61%)	261 (49.4%)	2,928 (65.8%)

The research team explored limiting the target crashes used in the selection of candidate intersections to the more severe crashes (e.g., fatalities and incapacitating injuries) since some agencies limit the severities used in their network screening analysis. However, the process presented in this report uses all target crashes, a narrowed focus compared to total crashes that are generally used in network screening. Narrowing the focus further to include only those target crashes that resulted in fatalities, type A, or type B injuries would base the selection on those intersections that had demonstrated the most severe target crashes, but would also greatly limit the sample of intersections. For example, in Minnesota just over 100 intersections would remain for a three year period. This subset is too narrow to account for annual fluctuations in crashes, and could lead to an agency installing the system at a location that was prioritized high on the list based on one or two crashes that are due to annual fluctuations, and do not necessarily represent a pattern of target crashes at the intersection. Therefore, this analysis considered target crashes of all severities.

Note that crash severity is considered in the calculation of benefit to cost. The recommended use of all severities is to identify those intersections where the target crash is occurring consistently across several years of data.

Timeframe

The overall objective of this effort was to develop a method to identify intersections that were good candidates for the applications based on crash data. To accomplish this objective, a method was needed to identify the timeframe that State and local agencies should use in their analysis of candidate intersections. In general, intersection crash counts fluctuate at any given intersection from year to year. One can reduce variation with more years of data, but operational or design changes may have been implemented over time. This is particularly likely at intersections that experience a high frequency of crashes, as improvements may be implemented in response to crash occurrence.

Most agencies use historical crash data of some form in their network screening to identify locations that are expected to experience future crashes, and therefore require some form of remediation. In a sophisticated analysis, safety performance functions (SPFs) can be developed to predict future crashes based on past crashes and other factors such as volume. An SPF is a statistical model developed to estimate the “typical” crash frequency for a specific type of roadway entity, based on the traffic volumes and key characteristics. However, one of the goals of this effort was to develop a method that is easy to implement by a State or local agency without rigorous statistical analysis.

An analysis was conducted to select an approach to best identify those intersections that were expected to continue to experience crashes and are potential candidates for this system. Ten years of intersection data in California were used for this part of the analysis. The analysis used the intersection groups and target crashes described in the previous sections. Additionally, the research team screened out intersections with missing traffic volume information (traffic volumes for either the mainline or cross-street) and intersections with information that appeared to be incorrect (negative, abnormally small, or abnormally large numbers for daily traffic volumes). The team used volume thresholds for the traffic volume signal warrant in the Manual on Uniform Traffic Control Devices (MUTCD) to guide this filtering process. While traffic volume is not the only warrant for traffic signals, removing those intersections with traffic volumes that exceed these criteria reduces the likelihood of including signalized intersections that are mislabeled as stop-controlled in the dataset. The research team used the intersection milepost to calculate the distance between consecutive intersections on the same route. The team made the decision to remove all intersections that are less than 200 feet from another stop-controlled intersection or less than 350 feet from a signalized intersection. This decision was made based on the nature of crash reporting and the precision of crash location identification. When intersections are closely located it is infeasible to distinguish between crashes occurring at one intersection or the other.

The final dataset includes a total of 7,777 stop-controlled intersections with stop control on the minor road. All 7,777 intersections had AADTs for both major and minor approaches, as well as crash counts. These are necessary variables for the analyses conducted in this effort.

As previously stated, a method is needed to identify those intersections that consistently experience the target crashes. In order to achieve this, the research team considered several methods based on the three desired characteristics outlined in the objectives section (i.e., easy to implement, based on five years of data or less, and results in the identification of those locations with the most opportunity to reduce target crashes).

The following measures were evaluated, using all crash severities:

- Annual crash frequency.
- Two-year crash frequency.
- Three-year crash frequency.
- Four-year crash frequency.
- Five-year crash frequency.
- Ten-year crash frequency.
- Proportion of target crashes to total crashes.

Five out of six measures tested met the first two characteristics (i.e., easy to implement and based on less than five years of data). The 10-year crash frequency was included in the analysis to see if many years of data would help improve the results. This could help an agency to make the decision whether to use as much data and they have available.

To assess the ability of each method to meet the last characteristic (i.e., results in the identification of those locations with the most opportunity to reduce target crashes), a baseline measure or ground truth measure was needed for comparison. Instead of looking at the raw crash counts or crash rate, the team used the potential safety improvement (PSI) based on an Empirical Bayes (EB) approach as the baseline. This is the method recommended by the Highway Safety Manual and recent research.

A PSI is the difference between the expected number of crashes (long term average) for a roadway entity (in this case, an intersection) and the "typical" number of crashes for that entity, predicted by a safety performance function (SPF). An SPF is a statistical model developed to estimate the "typical" crash frequency for a specific type of roadway entity, based on the traffic volumes and key characteristics. The EB-adjusted expected number of crashes is the long term average for a specific entity after adjusting for regression to the mean and random fluctuation over time.

EB-adjusted PSI

$$PSI_{EB} = N_{expected} - N_{predicted} = (1 - w) \times N_{observed} - (1 - w) \times N_{predicted}$$

Where:

- PSI_{EB} is the potential safety improvement based on the Empirical Bayes method.
- $N_{expected}$ is the expected number of crashes (long term average) for this intersection, corrected by the EB method
- $N_{predicted}$ is the average number of crashes predicted by the SPF based on similar intersections
- $N_{observed}$ is the number of observed crashes for this intersection
- w is the weight for EB-based correction

The above descriptions of PSI are illustrated in Figure 2. More detailed descriptions of the SPFs and the EB method are provided in Appendix B.

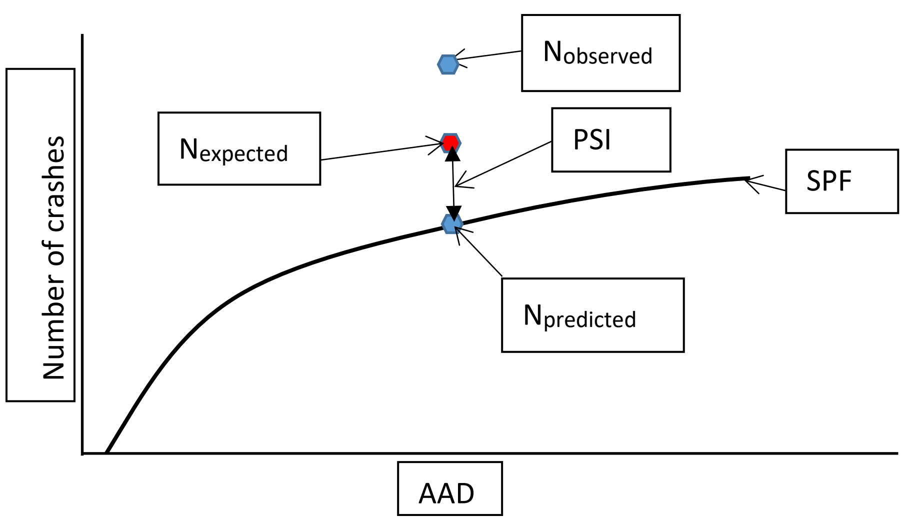

Figure 2. Concept of Potential Safety Improvement (PSI).

Two different SPFs were developed, one for 3-legged and one for 4-legged intersections. The research team estimated model parameters using several functional forms for each group of intersections presented in Table 3. Two separate SPFs for 3-legged and 4-legged intersections were found to be the best among the options examined. Other intersection characteristics were coded and included in the SPFs as indicator variables (e.g., rural vs. urban, mountainous or rolling vs. flat terrain, number of lanes on mainline). Using the SPFs, the team estimated the predicted numbers of target crashes ($N_{predicted}$), and then calculated the EB-adjusted expected numbers of target crashes ($N_{expected}$) and the respective PSIs for all 7,777 intersections. These intersections were then ranked based on PSI as well as the other seven measures being evaluated. The team followed a fractional ranking approach for breaking ties. In this approach, observations with equal values receive the same ranking number. This ranking number is the mean of what they would have been under an ordinal ranking approach.

Each of the seven rankings was compared against the PSI-based ranking. For each pairing between an alternative ranking and the PSI-based ranking, a Spearman's rank-order correlation analysis was used to evaluate how close the alternative rankings are to the PSI method.

Table 7 illustrates how three different alternative ranking measures may be compared with the PSI method and evaluated using the Spearman rank-order correlation. The example uses fictional data for 10 intersections to illustrate the method. The first column is the PSI-based rankings. The second, third and fourth columns show how these same 10 intersections are ranked based on one-year, two-year and three-year crash frequencies, respectively. The bottom row of this table is the Spearman's coefficients which indicate the strength of the statistical association between PSI-based and the other crash frequency based rankings. [Note a higher value of the Spearman's coefficient indicates a better alternative.] With a Spearman's coefficient of 0.903, in this example, the three-year crash frequency-based rankings are much closer to the PSI-based rankings than are the two-year rankings and are thus considered the better alternative to the ground truth (i.e., the PSI method). Again, the data in this table and the results are provided as an example to illustrate the method.

Table 7. Example of ranking evaluation method based on fictional data

PSI-based ranking	1 year crash frequency ranking	2 year crash frequency ranking	3 year crash frequency ranking
1	5	3	1
2	6	4	2
3	4	2	5
4	8	8	4
5	2	1	3
6	1	2	7
7	9	10	6
8	10	7	9
9	3	9	10
10	7	5	8
Spearman's Rank Correlation Coefficient	**0.176**	**0.485**	**0.903**

Identifying Potential Benefits

The primary anticipated benefit of the SSGA application is the reduction in target crashes and the fatalities and injuries resulting from those crashes at intersections where the systems are used. This anticipated crash benefit is the focus of this analysis. As discussed in the data section, the cost of a target crash can be monetized by the severity of a crash. This monetary cost of a crash is considered an economic benefit if a crash is avoided by the SSGA application.

Differences in Candidate Intersections

The intersections identified as candidates for the SSGA application were reviewed to identify differences in these intersections compared to other intersections in the State. Specifically, the following characteristics were explored:

- Volume (mainline, cross-street, estimated total entering).
- Width (Lane widths or crossing widths).
- Functional class of major roadway.
- Speed.
- Location type (interchanges, terrain).

- Intersection attributes (varies by State but may include lighting, channelization, etc.).

This information may be useful to FHWA, States, and other agencies for future efforts in analyses of safety issues and design of countermeasures. The findings of the comparison are presented in the Results section below.

RESULTS

The following sections present the results of California and Minnesota data analyses, including the identification of critical intersection types, the selection of a timeframe for use in the identification of candidates, the identification of the top ranked intersections in each State, potential deployment scenarios for each State, and a comparison of candidate intersections to other intersections.

Critical Intersection Types

As previously discussed, agencies may want to explore different types of environments (e.g., rural or urban, different intersection types) separately. For example, in some States, funds are allocated separately for rural roads. Table 8 presents the average annual target crashes for each of the 14 intersection groups in California that were introduced in Table 3. This table is limited to those intersections that experienced one or more target crashes during a three-year period. The number of intersections in each group is also displayed in the table. Each of the groups averaged less than one target crash per year per intersection. The table also provides an average cost per target crash based on the average severity distribution of the target crashes in the group and the average cost per crash severity presented in Table 2. The average annual number of target crashes is multiplied by the average cost of the target crash to get the average annual costs of target crashes per intersection for each group in the final column.

Table 8: Average annual target crashes and average cost by intersection group in California for intersections with at least one target crash.

Intersection Group (number of intersections in group with at least one target crash)	Average Annual Number of Target Crashes	Average Cost of Target Crash	Average Annual Intersection Target Crash Costs
3-legged, urban-2 mainline lanes x 2 cross street lanes (119)	0.479	$150,473	$72,075
3-legged, urban-4 mainline lanes x 2 cross street lanes (126)	0.590	$188,164	$111,007
3-legged, urban-6 mainline lanes x 2 cross street lanes (43)	0.705	$97,610	$68,857
3-legged, urban others (20)	0.517	$431,897	$223,147
4-legged, urban-2 mainline lanes x 2 cross street lanes (112)	0.652	$185,380	$120,828
4-legged, urban-4 mainline lanes x 2 cross street lanes (134)	0.721	$194,516	$140,322
4-legged, urban-6 mainline lanes x 2 cross street lanes (27)	0.938	$102,364	$96,046
4-legged, urban others (18)	0.963	$70,925	$68,298
3-legged, rural-2 mainline lanes x 2 cross street lanes (258)	0.521	$200,595	$104,444
3-legged, rural-4 mainline lanes x 2 cross street lanes (39)	0.513	$611,586	$313,634
3-legged, rural others (14)	0.476	$575,608	$274,099
4-legged, rural-2 mainline lanes x 2 cross street lanes (315)	0.690	$435,557	$300,512
4-legged, rural-4 mainline lanes x 2 cross street lanes (64)	0.635	$615,632	$391,183
4-legged, rural others (12)	0.667	$87,909	$58,606

While intersection groups could be prioritized by the Average Annual Number of Target Crashes, the addition of target crash costs to the intersection groupings provides more insight to the prioritization. Crash costs can capture both frequency and severity of crashes. However, it should also be noted that crash costs are heavily weighted by a fatal crash. The rural groups experienced more severe crashes and as such, the average costs for those groups were higher. This is likely due to the higher approach speed at these rural intersections.

The group of rural 4-legged, 4 x 2 intersections (that is, rural intersections with a four-lane mainline and a two-lane cross street) averages 0.635 target crashes per year on average. Based solely on average target crashes, it appears this is not a priority type of intersection. However, once the average cost of target crashes is considered, the importance of this type of intersection becomes greater. The average cost of a target crash at rural 4 x 2 intersections is over $600,000 because the crashes at these intersections are more severe, likely a reflection of the higher speeds. If averaged by cost, this intersection group would be the highest priority of the 14 groups.

Table 9 presents similar information for Minnesota. As with the California data, only those intersections that experienced one or more target crashes in a three-year period are included in the table. As evidenced by the average cost of target crashes at each intersection, the rural r-legged intersections (both four-lane mainlines and two-lane mainlines) are a priority.

Table 9. Average annual target crashes and average cost by intersection group in Minnesota for intersections with at least one target crash.

Intersection Group (number of intersections in group with at least one target crash)	Average Annual Number of Target Crashes	Average cost of Target Crash	Average Annual Intersection Target Crash Cost
3-legged, urban-2 mainline lanes x 2 cross street lanes (56)	0.45	$ 220,684	$ 98,520
3-legged, urban-4 mainline lanes x 2 cross street lanes (34)	0.56	$ 76,322	$ 42,650
4-legged, urban-2 mainline lanes x 2 cross street lanes (53)	0.40	$ 57,363	$ 22,729
4-legged, urban-4 mainline lanes x 2 cross street lanes (31)	0.59	$ 255,358	$ 151,018
3-legged, rural-2 mainline lanes x 2 cross street lanes (37)	0.42	$ 90,943	$ 38,508
3-legged, rural-4 mainline lanes x 2 cross street lanes (22)	0.52	$ 115,837	$ 59,673
4-legged, rural-2 mainline lanes x 2 cross street lanes (113)	0.42	$ 735,178	$ 312,288
4-legged, rural-4 mainline lanes x 2 cross street lanes (20)	0.75	$ 756,348	$ 567,261

Selection of Timeframe and Candidate Intersections

As discussed in the methodology section, seven measures were explored to determine the best method for State and local agencies to use crash data to identify candidate locations for SSGA systems. The six different measures tested using California data included:

- Annual crash frequency.
- Two-year crash frequency.
- Three-year crash frequency.

- Four-year crash frequency.
- Five-year crash frequency.
- Ten-year crash frequency.
- Proportion of target crashes to total crashes.

As noted in the Methodology section, in this analysis, all 7,777 two-way stop-controlled intersections were ranked using the PSI-based method and each of the alternative methods. The research team also performed an additional analysis using only the top 10 percent of intersections based on PSI-ranking. The following table shows the Spearman's rank correlation coefficients between PSI-based rankings and the rankings based on the seven options for the entire dataset and a subset of PSI-based top 10 percent.

Table 10. Comparison of six methods to PSI-based method for identifying priority intersections.

Ranking Method	Spearman's coefficient based on all intersections	Spearman's Coefficient based on top 10 percent of PSI-ranked intersections
1-year crash frequency rankings	0.362	0.48
2-year crash frequency rankings	0.46	0.702
3-year crash frequency rankings	**0.533**	**0.931**
4-year crash frequency rankings	0.42	0.85
5-year crash frequency rankings	0.345	0.812
10-year crash frequency rankings	0.138	0.729
Proportion of 3 year target crashes to 3 year total crashes	0.53	0.208

The results show the rankings based on three-year crash frequency are the closest to the ones based on PSI (i.e., the largest Spearman's rank correlation coefficient). The results suggest the three-year crash frequency is a better representation of the long term average than other alternatives, including the four-year and five-year averages. This could be a result of changes in operational and design characteristics.

The EB-based PSI approach is considered a more reliable estimate of the long-term safety performance of an entity. If an agency has the resources and capability, this approach can provide more reliable results than average crash frequency alone. It is more sophisticated and reliable. However, it is not suggested for use by the agencies unless they have the resource and capability to perform this type of EB-based analysis. The EB-based approach violates the first among three desired characteristics: ease of implementation. Based on the analysis results, the **three-year crash frequency method holds the most promise** for providing a reliable method that achieves the desired characteristics and the results indicate it is sufficient to use the three-year average.

Based on this three-year frequency of crashes, the intersections in each dataset were ranked in priority order for implementation. The highest crash frequency of these intersections in each agency are listed in priority order in Appendix C for California. The lists provide the three-year average number of total and target crashes for each of the high priority intersections.

Demonstration of Benefits

As previously discussed, the anticipated benefit of the SSGA systems is the monetized benefit of a reduction in target crashes. The potential economic benefit of a system at a specific candidate intersection will be influenced by the number and severity of expected target crashes, the effectiveness of the system in preventing target crashes, and the deployment of equipped vehicles. This is best demonstrated by selecting example intersections from each dataset and calculating the expected benefit.

The following section presents four examples, two for each dataset for California and Minnesota:

- Table 11 presents an intersection in California with **10 Target Crashes** per year on average. The example assumes that vehicle deployment follows the **5-Year Mandate** presented in Figure 1.
- Table 12 presents an intersection in Minnesota with **5 Target Crashes** per year on average. The example assumes that vehicle deployment follows the **5-Year Mandate** presented in Figure 1.
- Table 13 presents the same California intersection with **10 Target Crashes** per year on average that was presented in the first example. However, this example assumes that vehicle deployment follows the **15-Year Organic** penetration presented in Figure 1.
- Table 14 presents the same Minnesota intersection with **5 Target Crashes** per year on average that was presented in the second example. However, this example assumes that vehicle deployment follows the **15-Year Organic** penetration presented in Figure 1.

The start of system installation and vehicle penetration in all four examples is assumed to be 2020. For each scenario presented, the system is assumed to be 95 percent effective in reducing crashes when communicating with an equipped vehicle. Complete effectiveness (i.e., 100 percent) was not used because there may be some drivers who receive the gap assistance message that may not heed the message.

The examples assume that total and target crashes would continue at current levels at the intersections without the installation of the system. Therefore, every year the same number of target crashes would be expected without the intervention. This assumption is a simplification intended for illustrative purposes. In reality, many other factors may affect the occurrence of crashes at the intersection, such as changes in traffic volume.

The tables all provide the total anticipated crashes prevented. These are represented to the nearest tenth. In reality, a tenth of a crash prevented is not possible (i.e., either a crash is prevented or it occurs). However, the table is intended to demonstrate the benefit that can be achieved over the twenty year period. The tables also provide the crash cost savings (based on the distribution of severity at the example intersection) and the percent of target crashes reduced each year.

The number of crashes anticipated to be prevented increases each year the system is in place because there is an increase in the penetration of connected vehicle technologies in the vehicle fleet in subsequent years. As would be expected, the 5-year mandate results in more crashes being prevented sooner as a result of more aggressive penetration.

Note that all of the costs presented in these examples are presented in 2014 dollars. Inflation is not considered, again for simplification of the examples.

Table 11. Estimated number of crashes prevented and crash cost saved over time with 5-year mandate deployment scenario at an intersection and with 10 target crashes expected without intervention (California).

Year	Deployment (percent)	Total crashes prevented	Crash cost saved	Percentage
2020	0.22	--	--	--
2021	1.79	0.2	$19,617	1.7%
2022	5.34	0.5	$58,521	5.1%
2023	10.33	1.0	$113,206	9.8%
2024	16.08	1.5	$176,220	15.3%
2025	22.14	2.1	$242,631	21.0%
2026	28.29	2.7	$310,029	26.9%
2027	34.42	3.3	$377,207	32.7%
2028	40.43	3.8	$443,070	38.4%
2029	46.25	4.4	$506,852	43.9%
2030	51.84	4.9	$568,112	49.2%
2031	57.14	5.4	$626,195	54.3%
2032	62.10	5.9	$680,551	59.0%
2033	66.70	6.3	$730,962	63.4%
2034	70.92	6.7	$777,209	67.4%
2035	74.76	7.1	$819,291	71.0%
2036	78.21	7.4	$857,100	74.3%
2037	81.30	7.7	$890,963	77.2%
2038	84.03	8.0	$920,881	79.8%
2039	86.43	8.2	$947,182	82.1%
2040	88.03	8.4	$964,717	83.6%
20 Year TOTAL		**87.3 Crashes**	**$10,065,798**	**48%**

Table 12. Estimated number of crashes prevented and crash cost saved over time with 5-year mandate deployment scenario and an intersection and with 5 target crashes expected without intervention (Minnesota).

Year	Deployment (percent)	Total crash prevented	Crash cost saved	Percentage
2020	0.22	--	--	--
2021	1.79	0.1	$14,927	1.7%
2022	5.34	0.3	$44,530	5.1%
2023	10.33	0.5	$86,141	9.8%
2024	16.08	0.8	$134,090	15.3%
2025	22.14	1.1	$184,624	21.0%
2026	28.29	1.3	$235,909	26.9%
2027	34.42	1.6	$287,026	32.7%
2028	40.43	1.9	$337,144	38.4%
2029	46.25	2.2	$385,676	43.9%
2030	51.84	2.5	$432,291	49.2%
2031	57.14	2.7	$476,487	54.3%
2032	62.10	2.9	$517,848	59.0%
2033	66.70	3.2	$556,208	63.4%
2034	70.92	3.4	$591,398	67.4%
2035	74.76	3.6	$623,420	71.0%
2036	78.21	3.7	$652,189	74.3%
2037	81.30	3.9	$677,956	77.2%
2038	84.03	4.0	$700,722	79.8%
2039	86.43	4.1	$720,735	82.1%
2040	88.03	4.2	$734,077	83.6%
20 Year TOTAL		43.6	$8,393,398	47.8%

Table 13. Estimated number of crashes prevented and crash cost saved over time with 15-year organic deployment scenario and an intersection and with 10 target crashes expected without intervention (California).

Year	Deployment (Percent)	Total crash prevented	Total crash cost saved	Percentage
2020	0.02	--	--	--
2021	0.09	0.0	$986	0.1%
2022	0.31	0.0	$3,397	0.3%
2023	0.83	0.1	$9,096	0.8%
2024	1.81	0.2	$19,836	1.7%
2025	3.38	0.3	$37,041	3.2%
2026	5.60	0.5	$61,370	5.3%
2027	8.49	0.8	$93,042	8.1%
2028	11.99	1.1	$131,398	11.4%
2029	16.03	1.5	$175,672	15.2%
2030	20.49	1.9	$224,549	19.5%
2031	25.27	2.4	$276,933	24.0%
2032	30.25	2.9	$331,508	28.7%
2033	35.35	3.4	$387,399	33.6%
2034	40.47	3.8	$443,509	38.4%
2035	45.53	4.3	$498,961	43.3%
2036	50.47	4.8	$553,098	47.9%
2037	55.23	5.2	$605,263	52.5%
2038	59.77	5.7	$655,017	56.8%
2039	64.04	6.1	$701,811	60.8%
2040	68.15	6.5	$746,853	64.7%
20 Year TOTAL		**45.2 crashes**	**$5,209,886**	**26%**

Table 14. Estimated number of crashes prevented and crash cost saved over time with 15-year organic deployment scenario and an intersection and with 5 target crashes expected without intervention (Minnesota).

Year	Deployment (Percent)	Total crash prevented	Total crash cost saved	Percentage
2020	0.02	--	--	--
2021	0.09	0.0	$751	0.1%
2022	0.31	0.0	$2,585	0.3%
2023	0.83	0.0	$6,921	0.8%
2024	1.81	0.1	$15,093	1.7%
2025	3.38	0.2	$28,186	3.2%
2026	5.60	0.3	$46,698	5.3%
2027	8.49	0.4	$70,798	8.1%
2028	11.99	0.6	$99,984	11.4%
2029	16.03	0.8	$133,673	15.2%
2030	20.49	1.0	$170,865	19.5%
2031	25.27	1.2	$210,725	24.0%
2032	30.25	1.4	$252,253	28.7%
2033	35.35	1.7	$294,782	33.6%
2034	40.47	1.9	$337,477	38.4%
2035	45.53	2.2	$379,672	43.3%
2036	50.47	2.4	$420,867	47.9%
2037	55.23	2.6	$460,560	52.5%
2038	59.77	2.8	$498,419	56.8%
2039	64.04	3.0	$534,026	60.8%
2040	68.15	3.2	$568,299	64.7%
20 Year TOTAL		**22.6 Crashes**	$4,532,633	**26%**

Large-Scale Consideration for Agency-Wide Deployment Levels

As previously discussed, this analysis illustrates how interested agencies could focus on implementing SSGA systems at those intersections that have the most target crashes based on a three-year average of target crash occurrence. For each individual intersection, the agency can conduct a cost benefit analysis. An agency may also want to set a goal for a systemic deployment (e.g., top five percent of all intersections) or a goal for reducing the number of target crashes agency-wide (e.g., reduce target crashes by 50 percent agency-wide over twenty years). For this broader scale consideration, the cumulative distribution of target crashes should be considered. The research team conducted an analysis for all candidate intersections in California and Minnesota to demonstrate the benefit of these simple graphs. To develop these graphs, an agency would need a listing of candidate intersections and the average annual target crashes at each.

Figure 3 and Figure 4 present the cumulative distribution of average annual total target crashes (total target crashes over three years divided by three) compared to the number of candidate intersections that experienced **at least one target crash** in California and Minnesota, respectively. As shown on the graphs, 10 percent of these intersections are responsible for nearly 30 percent of the total target crashes in California and 26 percent in Minnesota. The percent is consistent for the two datasets.

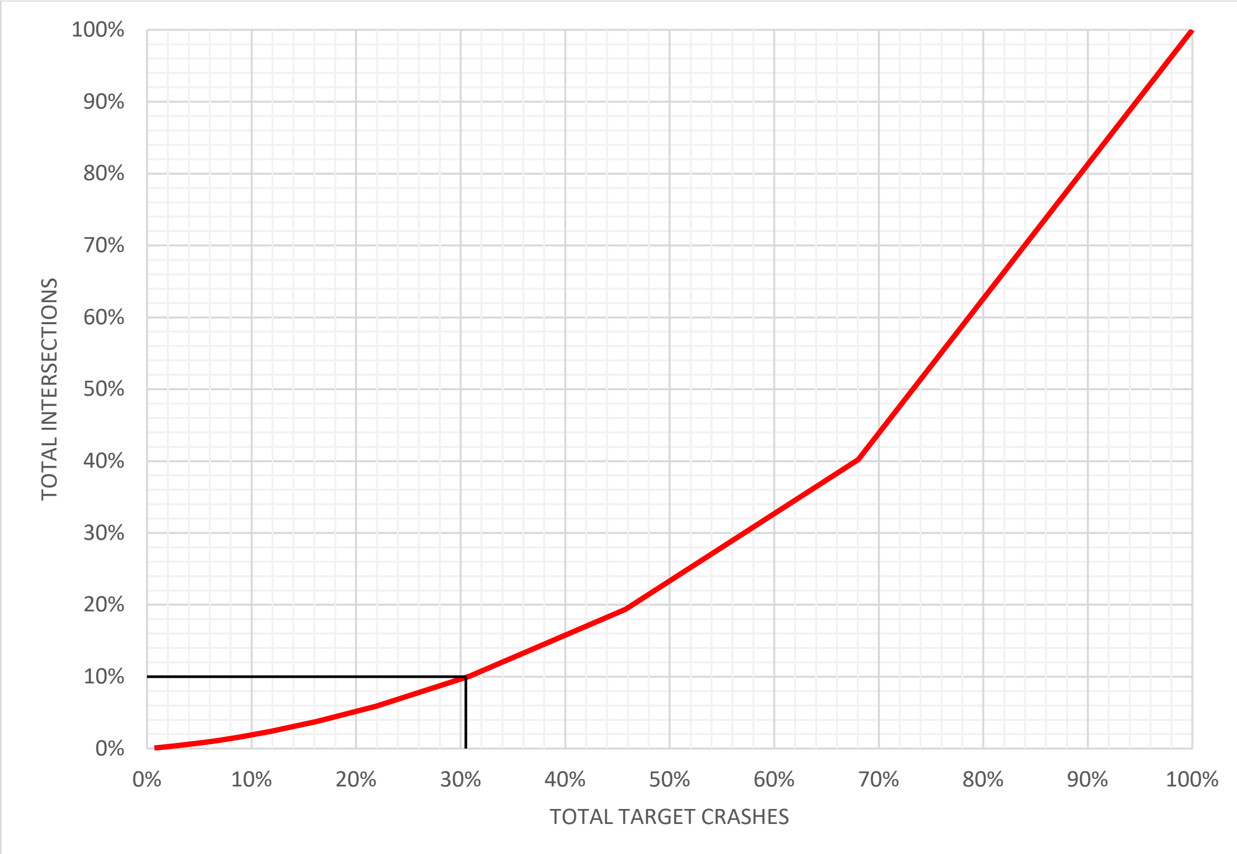

Figure 3. Relationship between cumulative number of stop-controlled intersections and cumulative target crashes in California.

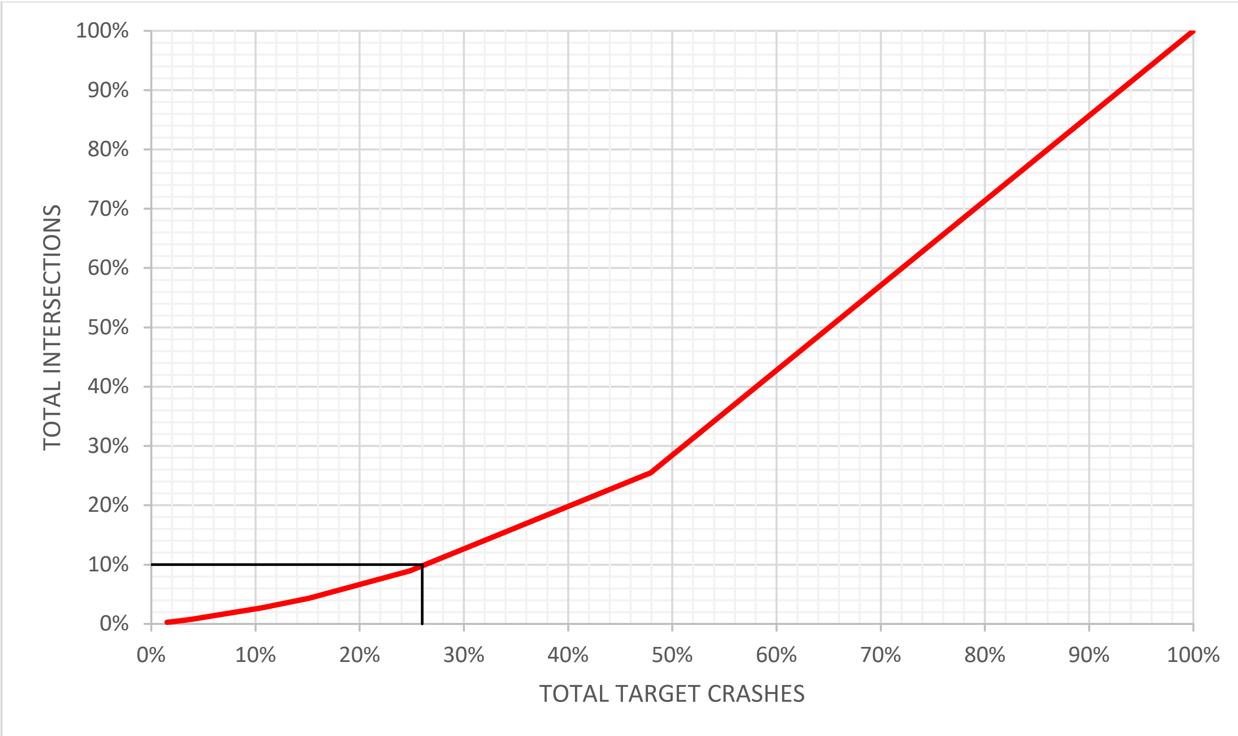

Figure 4. Relationship between cumulative number of stop-controlled intersections and cumulative target crashes in Minnesota.

An agency may also want to consider the severity of the target crashes. Translating the maximum injury severity of the crashes to crash costs is a useful way to account for severity in these graphs. Figure 5 presents the cumulative distribution of annualized target crash costs compared to the number of candidate intersections in California that experienced at least one target crash in the last three years. The impact of deploying at the top 10 percent of intersections is more poignantly expressed once severity is included. The top 10 percent of intersections represents approximately 75 percent of total target crash cost. A similar graph for Minnesota is presented in Figure 6. The top 10 percent of intersections represent more than 80 percent of total target crash cost in Minnesota

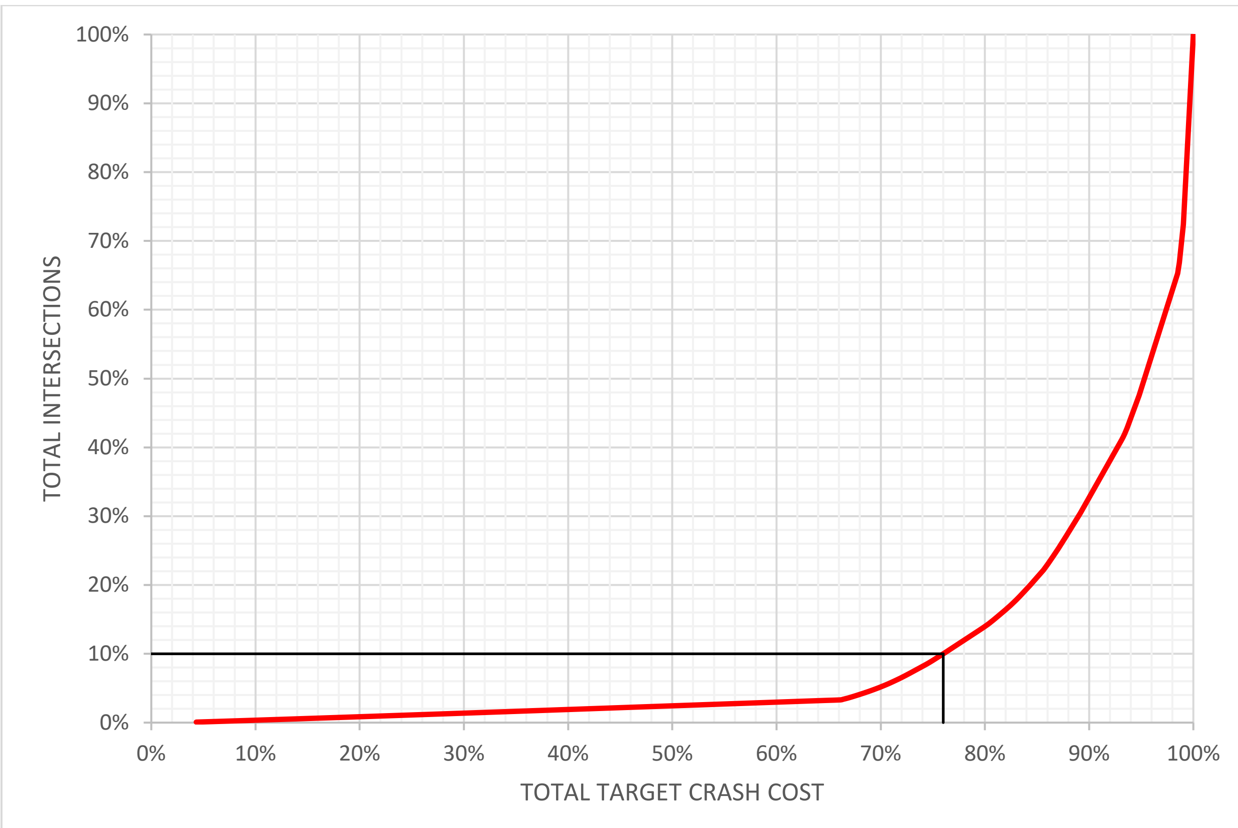

Figure 5. Relationship between cumulative number of intersections and cumulative target crash cost in California.

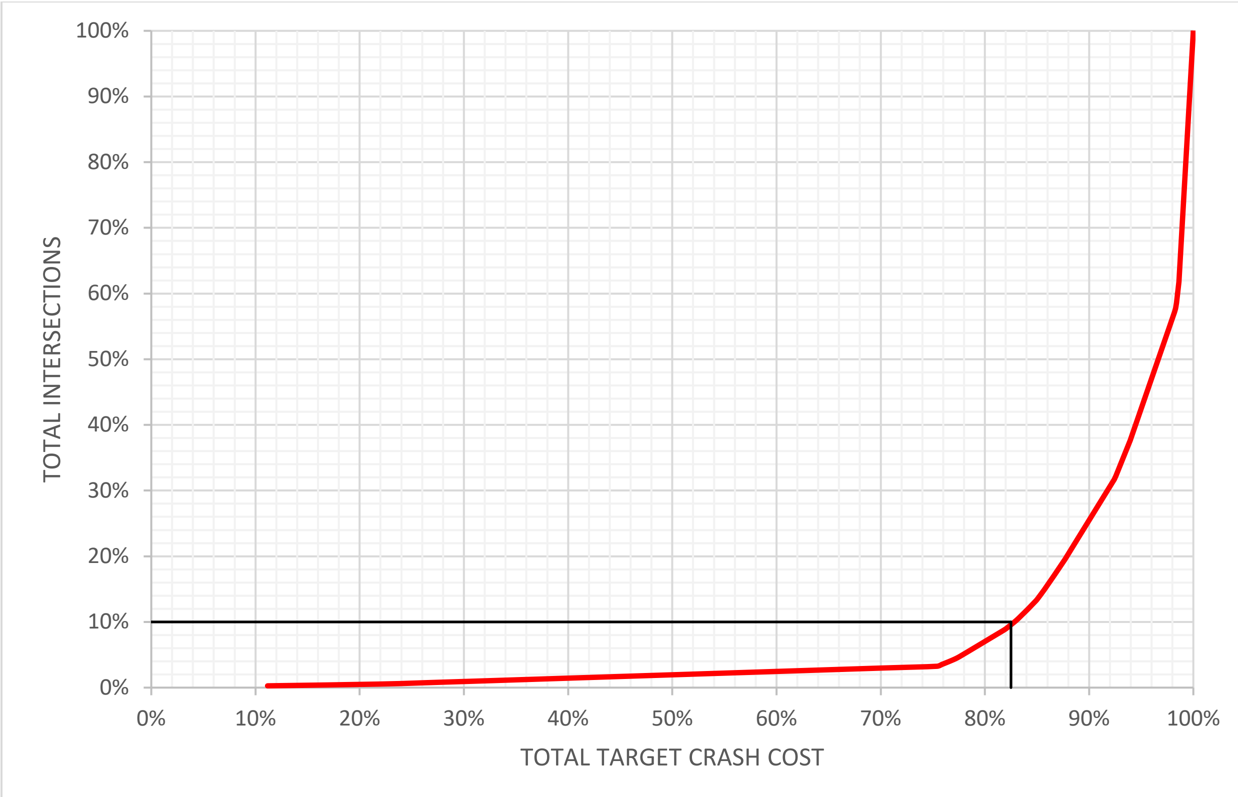

Figure 6. Relationship between cumulative number of intersections and cumulative target crash cost in Minnesota.

Difference in Candidate Intersections

As discussed in the methodology section, the candidate intersections with the most target crashes in a three year period were compared to other were two-way stop-controlled intersections to identify any notable differences in the characteristics of the highest priority candidate intersections to the other intersections. Specifically, in California there were 252 candidate intersections that experienced three or more target crashes in a three-year period. These candidate intersections were compared to those that experienced two or less target crashes in the same three-year period. Similar comparisons were not explored in the Minnesota data because only 34 intersections experienced three or more target crashes during the three-year period and therefore is too small of a grouping from which to draw inferences.

This information is presented for the benefit of FHWA and their partners when relevant in any of the three datasets. (Note that only notable differences are discussed.) The implementing agencies may also find this useful for the preliminary screening of intersections or to initiate systemic improvements.

Volume

The California data included average AADT for the mainline and for the cross street. The mainline and cross-street AADTs for the priority candidate intersections were compared to the remaining intersections and presented in Table 15. The priority candidate intersections have higher average mainline and cross street AADT. This is expected as more volume, both for the mainline and the cross street present more opportunity for crashes. Notably, the higher average mainline volume may also reduce the number of adequate gaps for the minor approach.

Table 15. Comparison of volume characteristics of priority candidate intersections to other intersections in California.

Variable	Priority Candidate Intersections (3 or more target crashes)	Other Intersections
Major AADT	Minimum 625 Average 20155 Maximum 80333	Minimum 80 Average 10617 Maximum 11400
Minor AADT	Minimum 10 Average 1258 Maximum 7500	Minimum 10 Average 397 Maximum 10900

Speed

California includes design speed information in the HSIS roadway inventory data. The mainline design speed of the two groups of intersections were compared. Both groups include intersections with design speeds that range from 25 mi/hr to 70 mi/hr. However, as would be expected, the priority intersections included more intersections with design speeds 60 mi/hr or greater (56 percent of the priority intersections versus 48 percent for the other intersections). A similar finding is present in the functional class description. The majority (62 percent) of the priority intersections are on principal arterials. Therefore, the priority intersections are on high-speed mainline roadways.

Terrain

A large notable difference in the two groups of intersections in the California data is for terrain. The terrain is flat for 73 percent of the intersections that experienced three or more target crashes compared to 45 percent of the remaining intersections. That is, those that are experiencing more target crashes are actually at flat intersections. This finding may seem counterintuitive as flat terrain should provide more opportunity to identify oncoming vehicles on the mainline. However, the flat terrain may present higher speeds and potentially some difficulties in judging the approaching speed of oncoming vehicles.

Approach Legs

The candidate intersections includes intersections with three approaches and intersections with four approaches. Notably, in California, those intersections that experienced three or more target crashes included more four-legged intersections. Specifically 66 percent of those intersections were four-legged intersections compared to just 27 percent of the other intersections.

CONCLUSIONS AND DISCUSSION

This report presents a method for State or local agencies to screen two-way stop-controlled intersections and develop a first prioritization of these intersections for deployment of a vehicle-to-infrastructure Stop Sign Gap Assist (SSGA) system. This effort was based on two States (California and Minnesota). This effort included several assumptions for ease of analysis and to demonstrate the approach including the system effectiveness, vehicle penetration rates, and flat crash levels. Additionally, three system cost scenarios were presented. All of these assumptions were inputs to the analysis and can be changed as more refined inputs are available.

Based on the analysis conducted, the following process is proposed for agencies in identifying potential two-way stop-controlled intersections for the installation of V2I SSGA systems:

Step 1. Identify Two-way Stop-Controlled Intersections

This analysis used two States that maintain an intersection inventory that provided information on the intersection control. Without such an inventory, agencies can use a sign inventory, local knowledge, or manual review of aerial maps or photo logs to identify these intersections and determine the control for the intersection approaches.

Step 2. Attribute Crashes to Intersections

For each intersection, attribute crashes to the intersection. This is generally done by identifying crashes within a 250 ft radial distance of the intersection, although the process varies by agency (e.g., some use 150 ft in urban areas) and the method should reflect agency practices for similar efforts. The target crashes for the SSGA system are crossing path crashes at the intersection. However, a larger radial distance is used to account for inaccuracies in location referencing.

Step 3. *Remove Intersections Improved in the Last Three Years of Planned for Improvement*

Two-way stop-controlled intersections that experience frequent crashes may be identified for signalization or other changes in traffic control that would remove the intersection for consideration for SSGA. This step will likely require an agency to seek additional information beyond what is available in a roadway inventory including signal warrant studies, transportation improvement program documents, HSIP project lists, and local knowledge.

Step 4. Determine a Method to Identify Target Crashes in Crash Data

Target crashes for this application are those resulting from a vehicle on the stop-controlled minor approach colliding with a vehicle on the perpendicular uncontrolled major approach. These are crashes where the driver of the vehicle on the minor approach entered the intersection without an adequate gap in approaching traffic. These crashes include several types of crossing path crashes because the minor road vehicle may have been going straight, turning left, or turning right. Based on the analysis conducted here, this should be defined in the crash data using information on the involved vehicles including the number of vehicles involved (two or more to exclude single vehicle crashes), the direction of travel for the involved vehicle (to include crashes involving at least a vehicle from the minor road and a vehicle from the major road), and some information on either the movement preceding the crash or the accident type to remove crashes resulting from unusual maneuvers such as U-turns.

Step 5. Calculate Three-Year Average Annual Target Crashes and Target Crash Costs

Using the three most recent years of available crash data, calculate the average number of target crashes at each intersection and the average annual cost of the target crashes. The research team suggests that agencies include all severities in their screening efforts and apply the crash costs presented in this report (or their own agency developed costs) by severity to calculate the costs.

Step 6. Combine Intersections into Related Groups (Optional)

If desired, the agency could use several variables to group intersections including number of legs, land use, and approach lanes. Groups of 30 intersections or more is a reasonable base. The purpose of this step is to identify groups that may need separate consideration, particularly if separate funds are available for certain function classes such as rural two-lane roads.

Step 7. Develop Prioritized List

The analysis here developed a prioritized list based on a three-year average of target crash frequency. The list could also be prioritized by the monetized cost of the target crashes or subdivided by the groups identified in step 6.

This method is based on the reported crashes and operational and geometric data available in a roadway inventory. The agency would use this list as an initial step in their efforts. The next step in prioritization would likely involve a detailed review (including field collection and observations) of intersections that the agency intends to move forward. The costs for individual intersection systems would be compared to the monetized benefit of the crashes that the system is expected to prevent. The ability of the system to prevent crashes will increase every year as more and more vehicles are equipped.

There are additional considerations that an agency may have in prioritizing intersections for these systems that could be incorporated into the initial prioritization efforts. The largest consideration is the agency's existing future plans for the intersection. For example, if the intersection is part of a planned large-scale improvement such as a large corridor improvement program, the agency may remove the intersection from consideration for the system or consider how the system implementation could be scheduled as part of other construction at the intersection. Other considerations may include equity by district or region.

REFERENCES

1) Council, F., Zaloshnja, E., Miller, T., and Persaud, B. Crash Cost Estimates by Maximum Police-Reported Injury Severity Within Selected Crash Geometries, Washington D.C, Federal Highway Administration, 2005, FHWA-HRT-05-051.
2) Thomson, K. and C. Monje, Guidance on Treatment of the Economic Value of a Statistical Life (VSL) in U.S Department of Transportation Analyses-2015 Adjustment, U.S Department of Transportation, Memorandum, June 17, 2015.
3) Najm, W.G., Smith, J.D., and Yanagisawa, M. Pre-Crash Scenarios Typology for Crash Avoidance Research. Washington, D.C.: U.S. Department of Transportation, National Highway Traffic Safety Administration, 2007. DOT HS 810 767.
4) Nujjetty, A. P., Mohamedshah Y. M., Council, F., M. Guidebook for Data files-California, Washington D.C, Highway Safety Information System, Federal Highway Administration, 2014.
5) Nujjetty, A. P., Mohamedshah Y. M., Council, F., M. Guidebook for Data files-Minnesota, Washington D.C, Highway Safety Information System, Federal Highway Administration, 2014.

APPENDIX A: DATA ELEMENTS

HSIS Data Elements for Analysis: California

Total Variables: 60

*Note – Description: SAS Variable Name

Accident Subfile (Total: 16)

1. Type of Collision: ACCTYPE
2. Collision ACCYR: ACCYR
3. Accident Case Number: CASENO
4. County Route: CNTY_RTE
5. Time of Accident: HOUR
6. Intersection/Ramp Accident Location: INT_RMP
7. Intersection Crash: INTER
8. Light Condition: LIGHT
9. Location Type: LOC_TYP
10. Milepost: MILEPOST
11. Total number of vehicles: NUMVEHS
12. Road Surface: RDSURF
13. Collision Severity: SEVERITY
14. Motor Vehicles Involved: VEH_INVL
15. Type of Vehicle at Fault DOT: VTYPE_AT_FAULT_DOT
16. Weather: WEATHER, WEATHER 1

Vehicle Subfile (Total: 8)

1. Accident Year: ACCYR
2. Contribution Factor: CAUSE
3. First, Second Associated Factor: CONTRIB1, CONTRIB2
4. Direction of Travel: DIR_TRVL

5. Movement Preceding Accident: MISCACT1
6. Vehicle at Fault: VEH_AT_FAULT
7. Vehicle Number: VEHNO
8. Violation Category: VIOL

Roadway File (Total: 5)

1. Design Speed: DESG_SPD
2. Median Type: MED_TYPE
3. Median Width: MEDWID
4. Roadway Classification: RODWYCLS
5. Terrain: TERRAIN
6. Road County Route: CNTYRTE
7. Beginning Milepost: BEGMP
8. Average Annual Daily Traffic AADT
9. ADT Date: ADT_DTE

Intersection File (Total: 31)

1. County: COUNTY
2. Intersection Description: INT_DESC
3. Intersection Effective Date: INT_DTE
4. Intersecting Route Prefix: INT_PRF
5. Intersection Population Code: INT_POPGRP
6. Intersection Milepost: INTMP
7. Cross Street County Route: INTY_RTE
8. Junction Type: JUNCTYPE
9. Intersection Light Type: LGHT_TYP
10. Milepost: MILEPOST
11. Mainline AADT: ML_AADT
12. Mainline Number of Lanes: ML_LANES

13. Mainline Left Turn Channelization: ML_LEFT
14. Mainline Signal Mastarm: ML_MAST
15. Mainline Right Turn Channelization: ML_RIGHT
16. Mainline Traffic Flow: ML_TRFLO
17. Roadway Route Number: RTE_NBR
18. Traffic Control Type: TRF_CNTL
19. Intersection Type: TYPEDESC
20. X-Street AADT: XSTAADT
21. X-Street Number of Lanes: XSTLANES
22. X-Street Left Turn Channelization: XSTRTLFT
23. X-Street Signal: XSTRTMST
24. X-Street Right Turn: XSTRTRGH
25. X-Street State Route Indicator: XSTSTRT
26. X-Street Traffic Flow: XSTTRFLO
27. Cross Street County Route: INTY_RTE
28. Intersection Milepost: INTMP
29. Mainline ADT date: ML_ADTDT
30. Cross Street ADT date: XSTADTDT
31. Cross Street State Route Indicator: XSTSTRT

HSIS Data Elements for Analysis: Minnesota

Total Variables: 73

*Note – Description: SAS Variable Name

Intersection subfile (Total: 48)

1. Route System: RTE_SYS
2. Route number: RTE_NBR
3. Calculated Beginning Milepost: BEGMP
4. Calculated Ending Milepost: ENDMP
5. Intersection Milepost: MPOFFSET
6. Intersection Type: TYPE
7. Intersection Description: DESC
8. Traffic Control Device: TRAF_DEV
9. Roadway Lighting: RDWY_LGH
10. Category Assigned by District: DIST_CAT
11. Central Office Category: CNTL_CAT
12. Date of Accident Geocoding: EFEC_DTE
13. Verbal Description of an Approach of Intersection/interchange: INT_DESC
14. Number of Routes into: NBR_RTES
15. Number of Legs into: NBR_LEGS
16. Update Date: UPT_DTE
17. Combined RTE_SYS/RTE_NBR: INT_SYNB
18. Traffic Control Devices: TRAFCNTL
19. Intersection Description-Revised: TYPEDESC
20. Leg route system: RTESYS2
21. Leg route number: RTENBR2
22. Leg milepost: MPOFSET2
23. Approach road description: RDESC

24. Lower limit: LOLIMT
25. Upper limit: UPLIMT
26. Leg/Approach Number: LEG_NBR
27. Approach direction: DIR
28. Year 1 AADT: AADT1
29. AADT Year 1: ADTYR1
30. Year 2 AADT: AADT2
31. AADT Year 2: ADTYR2
32. Year 3 AADT: AADT3
33. AADT Year 3: ADTYR3
34. Year 4 AADT: AADT4
35. AADT Year 4: ADTYR4
36. Year 5 AADT: AADT5
37. AADT Year 5: ADTYR5
38. Approach Speed Limit: AP_SPD
39. Approach Traffic Control: AP_CNTL
40. Number of Approach Through lanes during off-peak period: AP_TLOFF
41. Number of Approach Through lanes during peak period: AP_TLPEK
42. Number of Leaving Approach Through lanes during off-peak period: LV_TLOFF
43. Number of Leaving Approach Through lanes during peak period: LV_TLPEK
44. Approach bypass/turn lanes: AP_BP_TL
45. Approach Comments: AP_COMNT
46. Reference Point: REF_PNT
47. Unique identifier for each record: RECORD_ID
48. True leg milepost: LEG_TRUE_MP

Accident Subfile (Total: 19)

1. Diagram of Accident Code: ACCDIGM
2. Type of Accident: ACCTYPE

3. Year Accident Occurred: ACCYR
4. Accident Number: CASENO
5. Hour Accident Occurred: HOUR
6. Light Conditions: LIGHT
7. Location Description: LOC_NARR
8. Relation to Intersection: LOC_TYPE
9. Location Reliability: LOCN_REL
10. Modified Reference Point: MILEPOST
11. Number of Vehicles: NUMVEHS
12. Rural/Urban Pop Code: POP_GRP
13. Roadway Classification: RODWYCLS
14. Route Number: RTE_NBR
15. Route System: RTE_SYS
16. Combined Route System/Route Number: RTSYSNBR
17. Accident Severity: SEVERITY
18. Traffic Control Device: TRF_CNTL
19. Travel Direction: TRVL_DIR

Vehicle Subfile (Total: 6)

1. Accident Number: CASENO
2. First Contributing Factor: CONTRIB1
3. Sequence of Event: EVENT1, EVENT2, EVENT3, EVENT4
4. Action Prior to Accident: MISCACT1
5. Direction Vehicle was Traveling: VEH_DIR
6. Relative Vehicle Number: VEHNO

APPENDIX B: SAFETY PERFORMANCE FUNCTION AND EMPIRICAL BAYES (EB) CALCULATION

Safety Performance Function for 4-legged intersections

$$Target3yrs = MLAADT^{\beta_1} \times XSTAADT^{\beta_2} \times e^{(\beta_3 \times MTN + \beta_4 \times ROLL + \beta_5 \times SPD50PLUS + \beta_6 \times LANE4 + \beta_7 \times LANE6PLUS + \beta_8)}$$

Coefficient	Description	Estimated value	Standard Error
β_1	Mainline AADT	0.469	0.061
β_2	Cross street AADT	0.495	0.033
β_3	Indicator for mountainous terrain	-0.242	0.142
β_4	Indicator for rolling terrain	-0.298	0.103
β_5	Indicator for mainline posted speed of 50mph or higher	0.408	0.100
β_6	Indicator for 4 lanes on the mainline	-0.401	0.110
β_7	Indicator for 6 or more lanes on the mainline	-0.622	0.222
β_8	Intercept	-7.812	0.571
k	Overdispersion parameter	1.641	0.133

Safety Performance Function for 3-legged intersections

$$Target3yrs = MLAADT^{\beta_1} \times XSTAADT^{\beta_2} \times e^{(\beta_3 \times DIVIDED + \beta_4 \times SPD50PLUS + \beta_5)}$$

Coefficient	Description	Estimated value	Standard Error
β_1	Mainline AADT	0.760	0.061
β_2	Cross street AADT	0.494	0.034
β_3	Indicator for divided roadway for the mainline	-0.314	0.119
β_4	Indicator for mainline posted speed of 50mph or higher	0.269	0.099
β_5	Intercept	-11.547	0.565
k	Overdispersion parameter	2.790	0.259

The variables are defined as follows:

- Target3yrs is the predicted number of crashes at each intersection (crashes/3 years)
- MLAADT is the average annual daily traffic entering the intersection from the mainline (both directions, veh/day)
- XSTAADT is the average annual daily traffic entering the intersection from the cross street (both directions, veh/day)
- MTN is an indicator variable for mountainous terrain (=1 if the intersection is in a mountainous area, =0 otherwise)
- ROLL is an indicator variable for rolling terrain (=1 if the intersection is in a rolling area, =0 otherwise)
- SPD50PLUS is an indicator variable for posted speed of 50 mph on the mainline (=1 if the posted speed limit on the mainline is 50 mph or higher, =0 otherwise)
- LANE4 is an indicator variable for intersection with 4 lanes on the mainline (=1 if the mainline has 4 lanes, =0 otherwise)
- LANE6PLUS is an indicator variable for intersection with 6 more lanes on the mainline (=1 if the mainline has 6 more lanes, =0 otherwise)
- DIVIDED is an indicator for divided roadway on the mainline (=1 if the mainline is divided, =0 otherwise)

Empirical Bayes (EB)-adjusted number of expected crashes:

$$N_{expected} = w * N_{predited} + (1 - w) * N_{observed}$$

Where:

- $N_{expected}$ is the EB-adjusted number of expected crashes
- $N_{predicted}$ is the number of crashes predicted by the Safety Performance Function
- w is SPF weight, accounting for the accuracy of the SPF prediction:

$$w = \frac{1}{1 + k * N_{predicted}}$$

- k is the dispersion parameter of the SPF model

APPENDIX C: HIGHEST CRASH FREQUENCY INTERSECTIONS IN CALIFORNIA

No.	Intersection ID	Intersection Description	3-Year Target Crashes	3-Year Total Crashes
1	15939	ARBOLEDA DR	17	31
2	16641	MCCABE RD. F	13	15
3	13360	PHILLIPS ST	12	19
4	9751	JENSEN AVE	12	16
5	16082	STODDARD ROAD LT	12	24
6	8746	PURISIMA ROAD, LT	11	17
7	13396	KIMBALL AVE	11	26
8	16046	HENRY MILLER	11	17
9	9246	AVE 144	11	13
10	15696	KASSON RD LT RIVER RD	11	14
11	16225	7TH ST. F	10	17
12	10268	ROAD 284, WORTH RD	10	15
13	7194	FRAZER LAKE RD	10	17
14	5294	FILBERT ST	9	12
15	17239	BLAYLOCK DR. RT OLD	9	22
16	13597	NB ON & OFF RAMPS-RTE15	8	18
17	16227	5TH ST. F	8	16
18	13128	GILBERT ST	8	10
19	10522	ROAD 196 (SOUTH)	8	12
20	3597	ORANGE AVE	8	14
21	17276	21ST ST	8	24
22	8134	ESPINOZA RD RUSSELL RD	7	54
23	705	INDIANOLA CUTOFF	7	11
24	10757	ROSE AVE	7	11
25	15561	E ST (OKDL)	7	28
26	10266	ROAD 224, WESTWOOD DR	7	9
27	17270	15TH ST	7	25
28	7380	MILPITAS BLVD	7	60

No.	Intersection ID	Intersection Description	3-Year Target Crashes	3-Year Total Crashes
29	12133	FERN ST LT	7	13
30	7674	SANTA LUCHIA CNYN RD LT	7	11
31	384	HARTMANN RD - RT 104	7	11
32	3612	ACACIA RD	6	10
33	2655	ALEXANDER/THIRD ST RT&L	6	9
34	13383	J ST	6	13
35	4935	KELLEY RDG/MINERS RANCH	6	13
36	12933	HILL VIEW RD-RT	6	7
37	8388	DEPOT ST	6	21
38	7798	BEAR CREEK RD	6	9
39	10148	BETHEL AVE	6	9
40	9559	AVE 256	6	6
41	8998	NEBRASKA AVE	6	7
42	9047	LACEY BLVD	6	8
43	8359	LUCY BROWN LN	6	11
44	14658	FINE RD	6	8
45	9450	AVE 312, RIGGIN AVE	6	20
46	9235	AVE 56	6	8
47	7888	WRIGHT RD	6	8
48	7143	NORTH KNOLL RD - LT	6	20
49	8189	RAMP INTERSECTION	6	7
50	223	RTE 29 FAS 1039 RT	5	10
51	2242	BROADWAY RT/SOUTHERN AV	5	7
52	11317	ROSE AVE	5	15
53	6043	W SELBY LN-E SELBY LN	5	17
54	2675	BRANSTETTER LN-LT.	5	9
55	17350	MERCED AVE	5	5
56	13368	SUNKIST ST	5	8
57	8295	CASSERLY RD CARLTON RD	5	11

No.	Intersection ID	Intersection Description	3-Year Target Crashes	3-Year Total Crashes
58	7579	OCEAN VIEW AVE	5	7
59	16182	MUSSEY GRADE RD - RT T	5	11
60	11895	15TH ST. E	5	7
61	9938	SAMPLE RD (RIGHT)	5	6
62	9227	EXCELSIOR AVE	5	8
63	10761	GUNDRY AVE	5	13
64	9051	FREMONT AVE	5	6
65	10256	ROAD 152, BLISS LANE	5	8
66	12595	SHEEP CREEK RD	5	10
67	8382	RUSSELL AVE	5	12
68	17244	CHRYSLER WAY LT OLD	5	28
69	9040	KASAS AVE	5	11
70	10655	AVENAL CUTOFF RD(EAST)	5	13
71	6872	VINE HILL LT/MUELLER RT	5	5
72	11578	WOOD RD	5	12
73	12589	BELLFLOWER ST - RT	5	8
74	9464	AVE 368	5	13
75	9187	FILBURN ST.	5	7
76	16272	PAUMA RESV. RD LT. T	5	12
77	8401	CONCEPTION AVE	5	16
78	17359	CITRUS DR LT	4	11
79	10815	BROAD AV	4	16
80	10948	PALM AVE LT	4	9
81	9595	KERN ST	4	4
82	12121	RICE RD	4	7
83	7486	DOLAN RD, RT	4	34
84	2456	NORTH ST L&R	4	4
85	13379	H ST	4	10
86	9140	PANAMA LANE (EAST)	4	6

No.	Intersection ID	Intersection Description	3-Year Target Crashes	3-Year Total Crashes
87	7403	ORCHARD DR. (RT)	4	6
88	16562	SECOND ST. F	4	7
89	5751	L-WEST ZINFANDEL LN	4	10
90	11318	CEDAR AVE	4	18
91	8457	CRANE ST	4	4
92	7515	BOYSEN AVE	4	11
93	9125	W TULARE AVE(WEST)	4	6
94	9531	LINDA VISTA CONNECTOR	4	15
95	13932	KIRBY RD	4	10
96	9054	JACKSON AVE	4	11
97	11319	RAMONA ST	4	20
98	6815	LLANO RD - LT	4	19
99	15997	PALM AVE	4	7
100	4393	RIOSA ROAD EAST- RT.	4	6
101	10162	FRANKWOOD AVE	4	6
102	13902	RABBIT SPGS RD	4	5
103	14975	SHAWS FLAT RT	4	7
104	15105	BROADWAY	4	4
105	15935	KIBBY RD	4	7
106	17237	HOLLAND DR. RT OL	4	31
107	17217	MEMPHIS ST LT/SEABRIDGE	4	11
108	17233	NEWMAN AVE	4	33
109	7545	OSO FLACO LAKE ROAD	4	8
110	5687	CAIFORNIA ST	4	13
111	13361	BELMONT AVE	4	13
112	12658	GARNET AV-RT & GREENSPO	4	20
113	9008	CONEJO AVE	4	13
114	7240	E14TH ST AT 163RD AVE R	4	9
115	6758	CHURCH ST	4	6

No.	Intersection ID	Intersection Description	3-Year Target Crashes	3-Year Total Crashes
116	15720	JEFFERSON ST (MOD)	4	4
117	8453	PAJARO ST	4	16
118	4923	ARBOL AVE	4	7
119	14745	SPERRY RD (PATTERSON)	4	6
120	11743	226TH ST	4	7
121	6710	MIDWAY RD	4	6
122	7488	STRUVE RD N	4	14
123	15562	SIERRA AVE - LT	4	32
124	3375	EAGLES NEST RD	4	5
125	15253	SANDY MUSH RD	4	5
126	11922	GOLDEN VIEW	4	6
127	10332	2ND AVE	4	10
128	11805	BALCOM CYN RD LT	4	9
129	16114	HILTON HEAD RD RT	4	10
130	17429	CATALPA AVE - RT	4	7
131	13595	CAJON BLVD-RT	4	12