Presentation for the ITE Western District Annual Meeting, June 2017, San Diego, CA

Intersection Safety Performance: A Case Study on the Application of Predictive Models

Sanjay Paul, Ph.D., P.E., PTOE, PTP
Lee Engineering, LLC.

ABSTRACT

Traffic engineers and safety advocates have developed many analysis techniques over the years to identify crash patterns and contributing factors in order to better develop and implement engineering solutions to make the roads safer. The Highway Safety Manual (HSM) that has been recently developed by the American Association of State Highway and Transportation Officials (AASHTO) represents the state-of-the-art techniques to quantify safety conditions on the roadway facilities. Methods described in the HSM can be successfully implemented in various project stages including; network screening to identify location with promise, evaluate impact of design changes, as well as to evaluate the effects of specific engineering safety countermeasures.

HSM is a data intensive procedure which requires information on geometry, traffic control and operational conditions, pedestrian activities, and other land use, and it is common for agencies not to have all of the required data readily available. This paper describes the step-by-step procedure to implement the HSM techniques and evaluate the intersection safety performance. It also describes the steps starting from data collection, data processing, application of the predictive models, estimation of potentials for safety improvements, and application of advanced statistical tools for ranking the locations to get the most bang for the bucks. This paper is a practice ready paper which provides step-by-step guidance to the practitioners to follow and implement the HSM predictive models and evaluates intersection safety performances.

INTRODUCTION

Road traffic crashes are national burden to both our society and economy. Statistics show approximately 35,000 people lost their lives, with an additional 2.4 million people injured every year in the United States, along with a significant amount of property damage crashes resulting in about $871 billion dollars of annual economic loss. This is not a problem one can title an act of God and then move on. To alleviate this curse, traffic engineers and safety advocates have developed many analysis techniques over the years to identify patterns and crash characteristics, and developed and implemented series of engineering solutions to make the roads safer.

Traditional Descriptive Methods

Traditionally, descriptive analysis is used to evaluate the safety risk at any facility. The descriptive methods includes crash frequency, crash rate, and/or equivalent property damage only (PDO). The crash frequency deals with the raw total number of crashes, where the analyst prioritize locations based on the total number of crashes. This method tends to be skewed towards the locations with high volume locations, as the crash probability is higher at high traffic locations. The crash rate normalizes the number of crashes by exposure and uses either number of vehicles use the facility (vehicles per mile length of segment or entering vehicles at any intersection) or vehicle miles traveled (VMT). This technique is skewed toward low volume areas as the denominator is volume and it yields higher values for lower traffic locations. The equivalent PDO method is similar to rate based method which converts all types of crashes onto PDO using different factors and then calculates the crash rate. This method is biased towards the locations with high severity crashes.

Advanced Predictive Methods

The HSM developed by AASHTO overcomes the drawbacks of the traditional approaches and represents the most current state-of-the-art techniques to quantify the safety performance of roadway facilities that includes both highways and surface streets and for both segments and intersections. This method takes all geometric characteristics and traffic operation features into account, and quantify safety conditions on any roadway facility. The traditional descriptive methods deals with the crashes that already occurred whereas the HSM method has the ability to predict possible number of crashes for any facilities before they actually occur. Therefore, this technique is proactive rather than reactive. Models developed in the HSM can be successfully implemented in various project stages including; network screening to identify location with promise, evaluate impact of design changes, as well as to evaluate the effects of specific engineering safety countermeasures.

PROJECT BACKGROUND AND MOTIVATION

The City of Phoenix is the sixth most populous city in the nation with more than 1.5 million people. The city is approximately 520 square miles and the roadway network has over 5,000 intersections (right part of Exhibit 1 shows a sketch map of city of phoenix). The HSM concepts and predictive models were implemented on approximately 1,600 intersections which were divided into four major intersection types based on the functional classification (Arterial-Arterial, Arterial-Collector, Arterial-Local, and Collector-Collector) of the intersecting roadways. Then each intersection type was divided into four sub-types based on the intersection geometry and presence of signal at the intersection. The four sub-types are: Four-Leg Signalized (4SG); Three-Leg Signalized (3SG); Four-Leg Stop-Controlled (4ST); and Three-Leg Stop-Controlled (3ST) intersections. The motivation of categorizing each major type into aforementioned four categories was that the HSM provides models for these type of intersections only. Left part of the **Exhibit 1** shows the number of intersections in each category which were analyzed.

Intersection Classification		Number of Intersections	Total
Based on Roadway Functional Class	Based on Intersection Geometry and Operation		
Arterial-Arterial Intersection	3SG	36	364
	3ST	15	
	4SG	292	
	4ST	21	
Arterial-Collector Intersection	3SG	54	735
	3ST	120	
	4SG	402	
	4ST	159	
Arterial-Local Intersection	3SG	34	209
	3ST	0	
	4SG	175	
	4ST	0	
Collector-Collector Intersection	3SG	8	298
	3ST	106	
	4SG	47	
	4ST	137	
Total Number of Intersections		1,606	

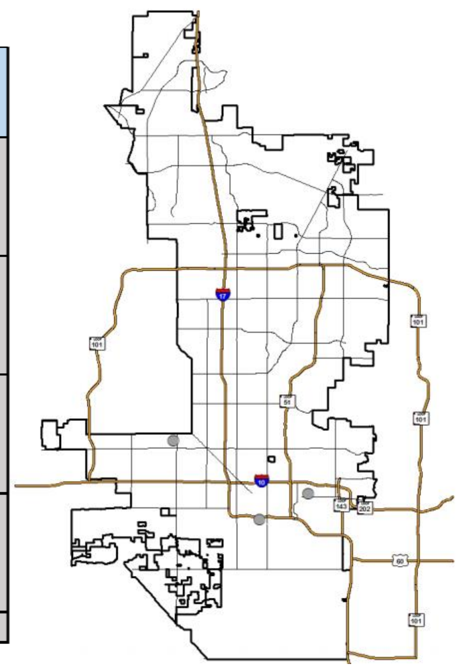

Exhibit 1: Number of intersections by different roadway geometry and control type (left); and City of Phoenix map (right)

Motivation and Objectives

Literature review revealed a number of projects that implemented the HSM developed predictive models and those are for freeway or state highway systems. Arizona Department of Transportation also completed a safety planning study where they implemented the HSM predictive methods on interstate 10 (on an urban corridor in metro Phoenix area). There was no study performed in Arizona and also there was no calibration factors for the urban intersections in the region. This gap acted as a barrier for the agencies to use the HSM methods in safety assessments. City of Phoenix took the initiatives and funded this project to implement the HSM models. The objectives of the study are as follows:

- ✓ to develop calibration factors for different types of crashes at different types of urban intersections;
- ✓ to develop and evaluate different statistical tools and select the best tool to identify the intersections with highest potential for safety improvements;
- ✓ to rank the locations based on their potential which would be used to develop projects and invest the limited available funds efficiently.

IMPLEMENTATION OF PREDICTIVE MODELS

The application of predictive models and evaluation of the safety performance of existing intersections could be categorized into three major efforts: data collection and processing, application of the predictive models, and evaluate relative performance to rank the locations. As mentioned before, the HSM predictive method is a data-intensive process which requires information on geometric characteristics, traffic volume, operational condition, pedestrian activities, and land use information in the vicinity (within 1,000 feet of the intersection).

Data Requirements, and Collections

A list of data items that is required and collected to implement the HSM predictive models to evaluate the intersection safety performance is as follows:

- Total entering volumes at the intersection by approach. Usually three or more years of traffic and crash data yields more reliable results. But one year of data would also provide the performance of the intersection. For missing year's traffic data, linear interpolation could be applied. Usually agencies have a program to collect traffic volume data once in every three years.
- Intersection type in terms of number of approaches: 3 leg signalized, 3 leg stop-controlled, 4-leg signalized, and 4-leg stop controlled.
- Presence of lighting at the intersection.
- Maximum number of lanes that the pedestrians need to cross on any approach.

Google earth is potential sources of information.

- Intersection control type e.g. signalized or unsignalized.
- Number of approaches with left turning lanes.
- Number of approaches with right turning lanes.
- Number of approaches with left-turning signal phasing.
- Types of left-turn signal for each approaches (e.g. protected, permissive, etc.)
- Presence of red-light violation cameras at the intersection.

Agency Traffic Management Center (TMC) or Google earth are good sources of the data.

- Number of approaches with right-turn on-red prohibited. TMC is a potential source for the information.
- Pedestrian activities in terms of total crossing pedestrian volume. Pedestrian volume count is not readily available at the agencies. TMC has idea on the pedestrian activities from a very high level point and the operators could categories the intersection into the following types (which is sufficient to apply into the models):
  - High
  - Medium
  - Low
  - Medium-High
  - Medium-Low

- Number of bus-stops and in the vicinity (1,000 feet). A geographic information system (GIS) file from the transit agency would provide the location information. A buffer circular area centering the intersection with 1,000 feet radius would provide the information on number of bus-stops near the intersection.
- Presence of schools in the vicinity (within 1,000 feet). Google earth/map could be used to extract presence of schools within 1,000 feet of the intersection.
- Number of alcohol sales establishments in the vicinity (within 1,000 feet). Google earth/map could be used to extract presence of alcohol sales stores within 1,000 feet of the intersection.
- Furthermore, type of roadway intersection could be useful to develop detail calibration factors. The definition of roadway functional classification may be tied to the roadway volume with different thresholds defined by the agencies.

Application of Predictive Models

In general the application of HSM predicted method includes the following process:

- Identify the intersection type, apply appropriate safety performance function (SPF), and estimate number of predicted crashes for locations of similar types. A SPF is an equation used to estimate or predict the expected average crash frequency per year at a location as a function of traffic volume and in some cases roadway or intersection characteristics. The following SPFs for the urban intersections are readily available in HSM:
  - ✓ Multi-vehicle Fatal and Injury crashes (Table 12-10 provides model parameters for 3ST, 3SG, 4ST, and 4SG intersections)
  - ✓ Multi-vehicle Property Damage Only crashes (Table 12-10 provides model parameters for 3ST, 3SG, 4ST, and 4SG intersections)
  - ✓ Single-vehicle Fatal and Injury crashes (Table 12-12 provides model parameters for 3SG, and 4SG intersections)
  - ✓ Single-vehicle Property Damage Only crashes (Table 12-12 provides model parameters for 3ST, 3SG, 4ST, and 4SG intersections)
  - ✓ Pedestrian-vehicle related crashes (Table 12-14)
  - ✓ Bicycle-vehicle related crashes (Table 12-17)

- Apply the crash modification factors (CMFs) based on the site-related data as mentioned in the *Data Requirements and Collections* section. A catalog of CMFs is provided in HSM as well as in the CMF Clearinghouse website *(http://www.cmfclearinghouse.org/*). The CMFs represent the effect of the intersection attributes towards the predicted number of crashes at that location. A CMF of 1.56 for the presence of nine or more alcohol establishments within 1,000 feet of the intersection represents that the vehicle-pedestrian crashes will be 56% higher at this particular intersection compare to an intersection without any alcohol establishments. A CMF adjusted crash number take account the geometric characteristics and operational conditions.

- Estimate CMF adjusted crash numbers for each type of crashes and compare with the observed number of crashes, and develop calibration factors. A calibration factor is a ratio of estimated CMF adjusted crash number to observed crash frequency for the same time period. These factors enables the SPFs that were developed using national crash data to be used in any regionals. The beauty of the calibration factors is that a set of calibration factors provides the agencies the opportunity to use the SPFs in all sorts of safety analysis in the region for both existing facilities and for future scenarios. In this project, a set of calibration factors are developed that covered each type of severity level and intersections type (as described in Exhibit 1). The calibration factors are necessary because the data used to build SPF may not necessarily reflect the exact conditions in another state/ jurisdiction and the CMFs do not cover the human behavioral aspects.

Exhibit 2 shows the calibration factors developed for the City of Phoenix that could be used for the region.

Intersection Classification		Calibration Factors	
Based on Roadway Functional Class	Based on Intersection Geometry and Operation	F&I Crash	PDO Crash
Arterial-Arterial Intersection	3SG	2.23	2.10
	3ST	1.88	1.68
	4SG	2.61	2.62
	4ST	1.64	1.55
Arterial-Collector Intersection	3SG	1.58	1.61
	3ST	1.14	1.27
	4SG	1.54	1.39
	4ST	0.99	1.06
Arterial-Local Intersection	3SG	1.26	3.21
	3ST	--	--
	4SG	1.46	1.46
	4ST	--	--
Collector-Collector Intersection	3SG	1.58	2.07
	3ST	1.02	1.34
	4SG	1.90	1.64
	4ST	0.90	1.01

Exhibit 2. Calibration Factors for the City of Phoenix

- Apply the calibration factors and estimate the predicted number of crashes for the intersections in the subject region. The predicted number of crashes is the estimate of average crash frequency which is forecast to occur as a site using a predictive model. Exhibit 3 (top part) shows the equation to estimate the predicted number of crashes.

Intersection Predicted Number of Crashes:

$$N_{Predicted} = \overset{\text{Calibration Factor}}{CF} * \overset{\text{Crash Modification Factors}}{\text{Combined CMF}} * \overset{\text{Safety Performance Function}}{\boxed{(a + b*\ln(AADT_{Major}) + c*\ln(AADT_{Minor}))}}$$

Site-Specific Empirical Bayes:

$$N_{Expected} = w * N_{Predicted} + (1-w) * N_{Observed}$$

$w = f\,(1/k)$; k is overdispersion parameter

Exhibit 3. Equation to Estimate the Predicted Number of Crashes

- Apply Empirical Bayes methodology to capture the effects of regression-to-mean, and estimated the expected number of crashes. The regression-to-mean is the natural variation in crash data. If regression-to-mean is not accounted for, a site might be selected for study when the crashes are at a randomly high fluctuation, or overlooked from study when the site is at a randomly low fluctuation while the expected number of crashes is the estimate of long-term expected crash frequency of a site, facility, or network under a given set of AADT in a given year. This step requires the over-dispersion factor which is provided in the HSM for each developed SFP in it. It also requires the observed and predicted number of crashes for the specific site. Exhibit 3 (lower part) shows the equation to estimate the site specific expected number of crashes.

- Estimate the expected and predicted number of crashes by different collision manners using the proportions found in the local data for both severity levels.

- Calculate the excess number of crashes. The excess number of crashes represents the difference between the predicted number of crashes and the expected number of crashes at a site. The excess number of crashes also represents the potential for safety improvements of the sites. Exhibit 4 shows a graphical representation of safety performance function, predicted number of crashes (after combined CMF and calibration factor adjustments), observed number of crashes, expected number of crashes, and excess number of crashes for specific AADTs.

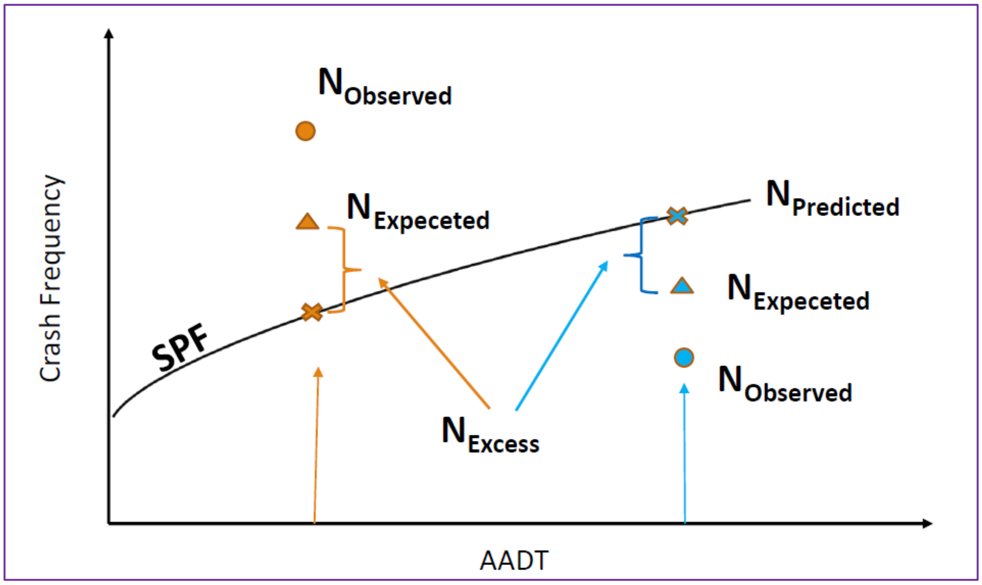

Exhibit 4: Safety performance function, and predicted, observed and expected crashes

- Provide diagnostic steps to better understand of crash patterns as well as describe potential countermeasure to mitigate unusual roadway and operational characteristics.

Network Screening & Relative Safety Performance Measures

A number of statistical tools were investigated at this stage of the study for network screening, identifying their safety performance, and ranking the locations with highest potential for safety improvements. The identified tools are as follows:

- *Excess number of crashes:* As HSM describes, the excess number of crashes is the difference between the expected number of crashes and the predicted number of crashes. This is commonly used as one of the safety performance measures to compare and rank locations. However, this tool does not relates the excess crashes with the predicted number of crashes and thus it is difficult to measure relative performance of all locations on the same scale. The Z-score statistics and Level-of-service-of-safety (described in the following sections) overcome this drawback.

- *Z-Score Statistics:* Z-Score is a ratio of the excess number of crashes to the square root of expected number of crashes for any specific location. It normalizes the excess number of crashes by the expected number of crashes and provides a better ranking for treatment for safety improvements. Exhibit 5 (on page 7) provides a sample representation of Z-scores and combined CMFs. A higher combined CMF represents a location with known problems/issues that are associated with CMFs (yellow box area in Exhibit 5). An example of possible problem could be permissive left-turns which is contributing towards more left-turn crashes. A higher Z score location represents higher potential for safety improvements for not any issues that are related to the CMFs (known) and are locations with promise for investments (blue box area in Exhibit 5).

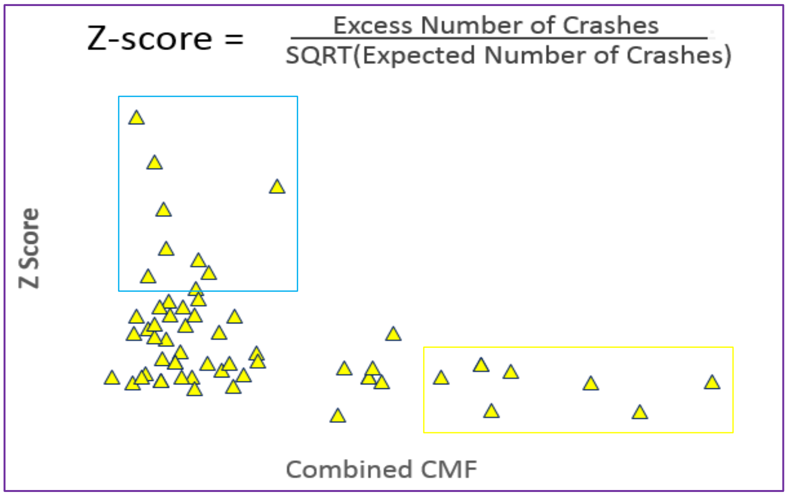

Exhibit 5: Z-score Statistics and Combined CMF (concept developed by Kittelson and Associates)

- *Level-of-Service-of-Safety (LOSS):* LOSS concept was developed by Kononov and Allery to qualitatively describe the degree of safety or unsafety of a roadway segments/components, identify the most promising locations, and allocate the limited available funds for improvements. The four zones are created taking the expected numbers as the base line and a ± 2.5 times standard deviation both sides of it. The beauty of this chart is the user can narrow it while sufficient fund is available as well as make it wider when less fund is available and need to know about a handful number of locations for improvements. Exhibit 6 demonstrates the concept of LOSS and categorizing locations based on the expected crash frequency.

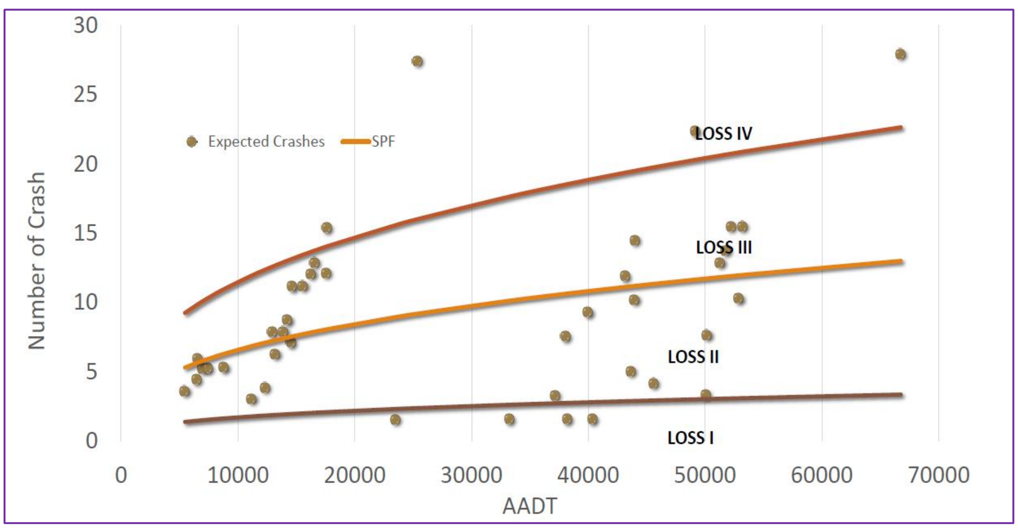

Exhibit 6: Level-of-Service-of-Safety Concept to Classify Locations with highest potential for Safety Improvement (concept developed by Kononov & Allery)

The characteristics of the four LOSS zones are as follows:

- ✓ LOSS-IV zone has high potential for safety improvement
- ✓ LOSS-III zone has less than expected performance with rooms for safety improvements

- ✓ LOSS-II zone has better than expected safety performance
- ✓ LOSS-I zone has low potential for safety improvements

A combination of these tools were used and a number of promising locations with highest potential for safety improvement were identified for the city.

CONCLUSIONS

Since the development of HSM a number of studies implemented the models but there was no calibration factors for this region in Arizona which acted as a barrier to use this advanced tools for safety assessments. As the calibration factors may vary significantly between regions, it is critical for each agency to develop their own calibration factors which allows them to take full advantage of the HSM predictive methods. This study guided the city personnel, directed on data collections, implemented the HSM predictive models, and developed the calibration factors for the city which could be used for any agencies in the region. The project also developed an excel-based tool which is a customized version of the AASHTO excel-file and it accommodates five years of data and perform all necessary analyses for the city to use. The study also introduced new techniques to measure relative safety performance and rank the locations. The ranking is a critical components which allows the agencies select locations with highest promise for safety improvements and ensures efficient use of the limited funds. This paper provides step-by-step procedure and guidance for the practitioners to implement the HSM predictive models in any agencies intersection safety assessments.

ACKNOWLEDGEMENTS

The author expresses sincere gratitude and gives credits to the other team members including Dr. Jim Lee of Lee Engineering, Dr. Andrew Kwasniak of Tatum Group, and Melinda Cruz and Kerry Wilcoxon of City of Phoenix for their involvements and contributions.