Identification of Hazardous Sites
Part 1

February 16, 2022

Instructor: Srinivas Geedipally
Texas A&M Transportation Institute
Srinivas-g@tti.tamu.edu

Purpose

- Identifying high risk locations or sites that experience more crashes than one would expect
- In other words, given the characteristics of the site, it experiences more crashes than sites having the same characteristics

Network Screening

- Also called blackspot identification
- Identify potentially hazardous sites across the network
- Identify a smaller subgroup of sites from the entire road network for detailed investigation or site diagnostics
- Assess the conditions of an existing infrastructure to assist transportation agencies in their long-range plan

Network Screening

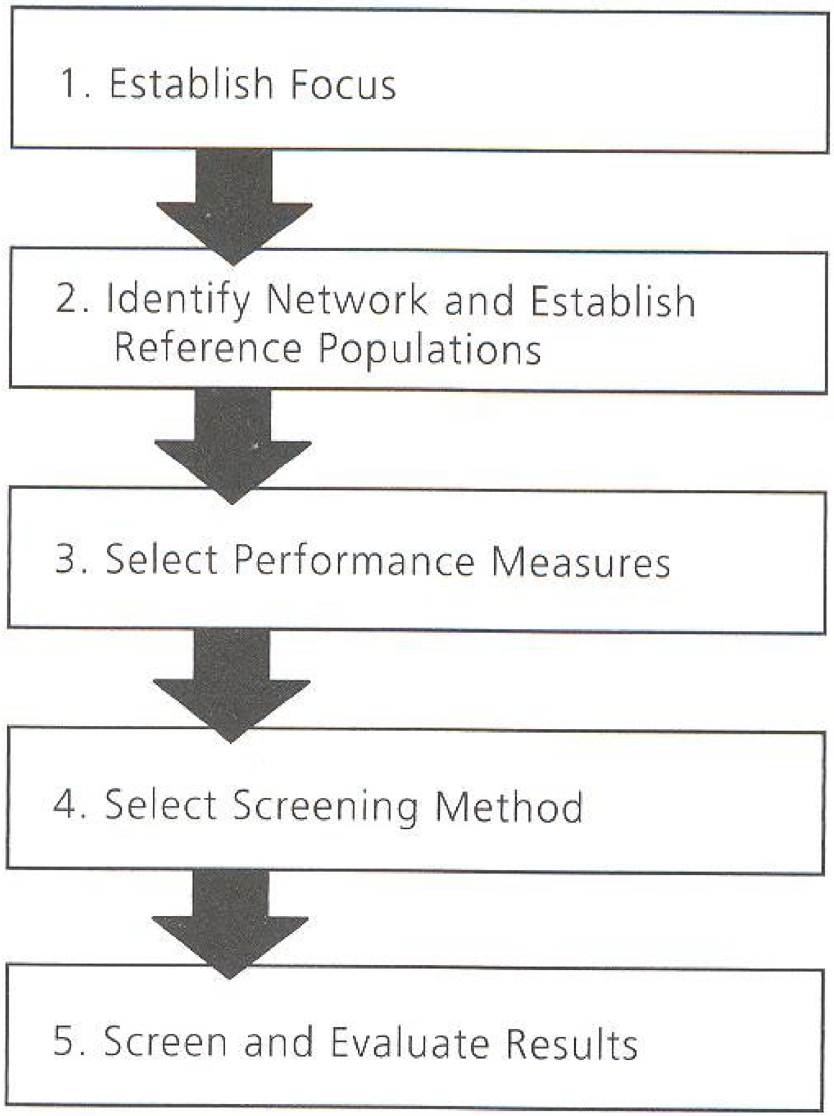
Reactive approaches

- rely on historic crash data

Proactive approaches

- rely on high-risk geometric and operational characteristics

Network Screening Process - HSM



Identify purpose and target specific crashes

Types of sites or facilities (i.e., segments, intersections, etc.)
Identify control group with sites that have similar characteristics (if needed)

Use one or more selection methods (discussed next)

Four methods: ranking, sliding window, peak searching and continuous risk profile.

Conduct screen analysis and evaluate results

Screening Methods

- Different methods are available
  - Each method has its own advantages and disadvantages
- If sites are improperly identified:
  - the "true" high-risk sites may not be treated – safety is not improved; or
  - the relatively "safe" sites may be identified for further investigation – safety funds are wrongly invested

Observed crash methods

Crash Frequency

- Sites ranked by crash frequency

Advantages

- Simplest method
- High crash frequency sites are necessarily detected

Disadvantages

- Bias towards high volume sites
- Does not consider exposure or long-term mean

Crash Frequency

Data used for examples 8.1 – 8.8.

Intersection	ADT on major st	ADT on minor st	Fatal crashes	Injury crashes	PDO crashes	Total crashes
1	37191	16705	0	18	26	44
2	19999	390	0	6	9	15
3	25608	2530	0	9	14	23
4	27223	8463	0	10	19	29
5	20336	7013	0	2	4	6
6	7524	1875	0	5	12	17
7	31646	17502	0	10	30	40
8	32117	9888	0	20	39	59
9	25224	17258	0	16	31	47
10	6856	5509	0	6	7	13
11	15025	4400	0	18	30	48
.	.	.	.	.	.	.
.	.	.	.	.	.	.
49	22577	390	0	5	8	13
50	18536	2935	0	10	4	14

Crash Frequency

Exercise 8.1

Using dataset 8.1, identify hazardous intersections based on the crash frequency method.

Calculate the average crash frequency in the reference population using Eq. 5.2:

$$\mu_r = \frac{1}{N}\sum_{i=1}^{N} x_i = \frac{1024}{50} = 20.48$$

Determine the threshold:

$$\mu(threshold) = 2 \times \mu_r = 40.96$$

The intersections with a total crash frequency greater than the threshold are identified as hazardous locations. Based on the crash frequency method, intersections 1, 8, 9, 11, 13, and 46 are detected.

Crash Frequency

Intersection	ADT on major st	ADT on minor st	Fatal crashes	Injury crashes	PDO crashes	Total crashes
1	37191	16705	0	18	26	44
2	19999	390	0	6	9	15
3	25608	2530	0	9	14	23
4	27223	8463	0	10	19	29
5	20336	7013	0	2	4	6
6	7524	1875	0	5	12	17
7	31646	17502	0	10	30	40
8	32117	9888	0	20	39	59
9	25224	17258	0	16	31	47
10	6856	5509	0	6	7	13
11	15025	4400	0	18	30	48
.	.	.	.	.	.	.
.	.	.	.	.	.	.
49	22577	390	0	5	8	13
50	18536	2935	0	10	4	14

Crash Rate

- Sites ranked by ratio between crashes and exposure

Advantages

- Commonly used by agencies
- Includes traffic exposure

Disadvantages

- Bias towards low volume sites
- Does not consider long-term mean
- Non-linear relation of crashes with exposure

Crash Rate

- Individual sites

$$R_i = \frac{C_i \times 10^6}{N \times 365 \times AADT_i \times L_i}$$

- Reference group

$$R_r = \frac{\sum_i C_i \times 10^6}{N \times 365 \times \sum_i (AADT_i \times L_i)}$$

Crash Rate

Exercise 8.2

Using dataset 8.1, identify hazardous intersections based on the crash rate method.

Calculate the average crash rate in the reference population using Eq. (8.2).

$$R_r = \frac{\sum_i C_i \times 10^6}{N \times 365 \times \sum_i (AADT_i)} = \frac{1024 \times 10^6}{2 \times 365 \times 1,177,116} = 1.19$$

Determine the threshold

$$R(threshold) = 2 \times R_r = 2.38$$

The intersections with a total crash rate greater than the threshold are identified as hazardous locations. Based on the crash rate method, intersections 6, 11, 13, and 31 are detected.

Rate Quality Control

- Sites higher than threshold are identified.
- Developed by industrial engineers for quality control purposes.

Advantages

- Consider the randomness of crashes
- Includes traffic exposure
- Establish a threshold

Disadvantages

- Complex method
- Does not consider long-term mean
- Non-linear relation of crashes with exposure

Rate Quality Control

- Site's crash rate

$$R_i = \frac{C_i \times 10^6}{N \times 365 \times AADT_i \times L_i}$$

- Site's critical rate

$$R_c = R_r + \left[p \times \sqrt{\frac{R_r \times 10^6}{AADT \times N \times 365 \times L}} + \frac{10^6}{2 \times AADT \times N \times 365 \times L} \right]$$

where R_c is the critical crash rate, R_r is the crash rate for the reference population, and, p is the P-value (=1.036, 1.282, 1.645, and 2.326 for a level of confidence of 85%, 90%, 95%, and 99%, respectively).

Rate Quality Control

Exercise 8.3

Using dataset 8.1, identify hazardous intersections based on the rate quality control method.

Calculate the critical rate for each site based on Eq. (8.3). For intersection 1, with a level of confidence of 95%, the critical crash rate is:

$$R_c = R_r + \left[p \times \sqrt{\frac{R_r \times 10^6}{AADT \times N \times 365 \times L}} + \frac{10^6}{2 \times AADT \times N \times 365 \times L} \right]$$

$$= 1.19 + \left[1.645 \times \sqrt{\frac{1.19 \times 10^6}{53,896 \times 2 \times 365}} + \frac{10^6}{2 \times 53,896 \times 365} \right] = 1.49$$

For intersection 1, the crash rate is 1.12. As it is lower than the critical crash rate, it is detected as nonhazardous intersection.

Based on the RQC method, intersections 6, 8, 11, 13, 16, 17, 21, 26, 29, 31, 36, and 46 are detected.

Equivalent Property Damage Only (PDO)

- Sites ranked based on assigned weights to different crash severities

Advantages	• Simplest method • Considers crash severity
Disadvantages	• Bias towards high-speed sites • Does not consider exposure or long-term mean

Equivalent PDO

Intersection	ADT on major st	ADT on minor st	Fatal crashes	Injury crashes	PDO crashes	Total crashes
1	37191	16705	0	18	26	44
2	19999	390	0	6	9	15
3	25608	2530	0	9	14	23
4	27223	8463	0	10	19	29
5	20336	7013	0	2	4	6
6	7524	1875	0	5	12	17
7	31646	17502	0	10	30	40
8	32117	9888	0	20	39	59
9	25224	17258	0	16	31	47
10	6856	5509	0	6	7	13
11	15025	4400	0	18	30	48
.	.	.	.	.	.	.
.	.	.	.	.	.	.
49	22577	390	0	5	8	13
50	18536	2935	0	10	4	14

Equivalent PDO

Exercise 8.4

Using dataset 8.1, identify hazardous intersections based on the EPDO method. Use the following weights (AASHTO, 2010).

Severity	Weight
Fatal (K)	542
Injury (A/B/C)	11
PDO (O)	1

Calculate the EPDO score for each site based on Eq. (8.5). For intersection 1, the EPDO score is:

$$EPDO\ score = \sum_{s} f_s \times C_s = 542 \times 0 + 11 \times 18 + 1 \times 26 = 224$$

The average EPDO score in the reference population ($EPDO\ score_r$) is calculated as 119.52.

Determine the threshold:

$$EPDO\ score(threshold) = 2 \times EPDO\ score_r = 239.04$$

The intersections with EPDO score greater than the threshold are identified as hazardous locations. Based on the EPDO method, intersections 8, 13, 33, and 46 are detected.

Severity Index Method

- Considers the EPDO score but standardizes it based on the total number of crashes at the site

Advantages	• Considers crash severity • Comparisons can be easily made across sites
Disadvantages	• Bias towards high-speed sites • Do not consider exposure or long-term mean

Severity Index Method

Exercise 8.5

Using dataset 8.1, identify hazardous intersections based on the SI method.

Calculate the SI for each site based on Eq. (8.6). For intersection 1, the SI is:

$$SI = \frac{EPDO\ score}{C} = \frac{224}{44} = 5.09$$

The average SI in the reference population (SI_r) is calculated as 5.62.

Determine the threshold:

$$SI(threshold) = 2 \times SI_r = 11.24$$

The intersections with SI greater than the threshold are identified as hazardous locations. Based on the SI method, intersections 13 and 33 are detected.

Composite Safety Score

- Considers different crash measures

Advantages	• Accounts for crash frequency, severity, and collision type
Disadvantages	• Requires rigorous analysis • Does not consider exposure or long-term mean

Composite Safety Score

- Site's composite safety score

$$FS = \frac{1}{5} \times \frac{CF}{\mathrm{Max}(CF)} + \frac{3}{5} \times \frac{CS}{\mathrm{Max}(CS)} + \frac{1}{5} \times \frac{CT}{\mathrm{Max}(CT)}$$

With

$$CS = 40C_K + 9C_A + 5C_B + 2C_C + C_{PDO}$$

$$CT = \sum Cost_t \times N_v$$

where FS is the composite safety score; CF is the total crash frequency at the site; CS is the total crash severity index for the site; CT is the total crash type score for the site; $\mathrm{Max}(CF)$, $\mathrm{Max}(CS)$, and $\mathrm{Max}(CT)$ are the maximum values recorded for any intersection in the network; C_s is the frequency of crashes with severity s; $Cost_t$ is the crash cost of collision type t; and, N_v is the number of vehicles involved in each crash. For more details about this method, the interested reader is referred to Qin et al. (2009).

Predicted crash methods

Potential for Improvement (PI) using Predicted Crashes

- Compares observed and predicted crashes. Statistical model is developed using the reference sites.

Advantages	• Accounts for traffic volume • Estimates a threshold
Disadvantages	• Requires rigorous analysis • May not properly capture the long-term mean of the sites (use average of the sites)

PI using Predicted Crashes

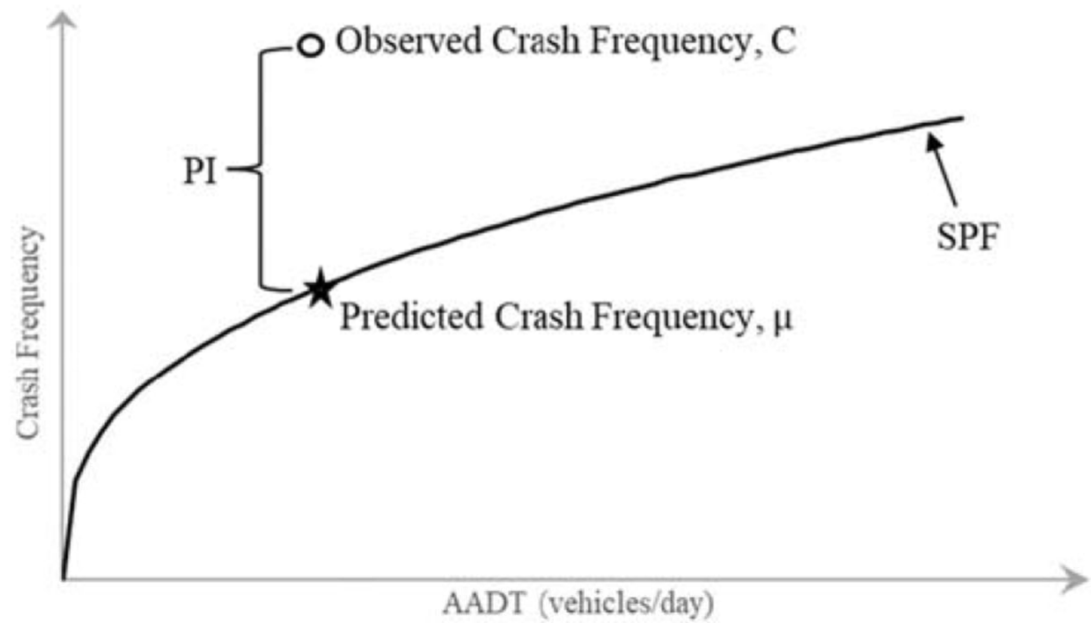


FIGURE 8.1 Graphical representation of potential for improvement using SPF.

$$PI = C - \mu$$

PI using Predicted Crashes

Exercise 8.6

Using dataset 8.1, identify hazardous intersections based on the potential for improvement using the SPF method. For simplicity, let us assume that the crashes are a function of just the major and minor street flows. Use the following functional form and parameter estimates.

$$\mu = e^{\beta_0} AADT_{maj}^{\beta_{maj}} AADT_{min}^{\beta_{min}}$$

Parameter	Estimate	Std. error	*P*-value
β_0 (Intercept)	−4.3049	1.3370	0.0013
β_{maj} (Major street AADT)	0.5969	0.1410	<0.0001
β_{min} (Minor street AADT)	0.1850	0.0614	0.0026
α (Dispersion)	0.2423	0.0610	<0.0001

Calculate the predicted crashes for each site based on the above functional form. For intersection 1, the predicted number of crashes are:

$$\mu = e^{-4.3049} 37191^{0.5969} 16705^{0.1850} = 43.6 \text{ crashes}$$

The *PI* for intersection 1 is estimated as $44 - 43.6 = 0.4$ crashes.

If sites with *PI* greater than 20 crashes are considered to be hazardous locations, then intersections 8, 11, 13, and 46 are detected.

Level of Service of Safety

- Each site is assigned a LOSS category based on the difference between observed and predicted crashes.

Advantages

- Accounts for variance in crash data
- Estimates a threshold

Disadvantages

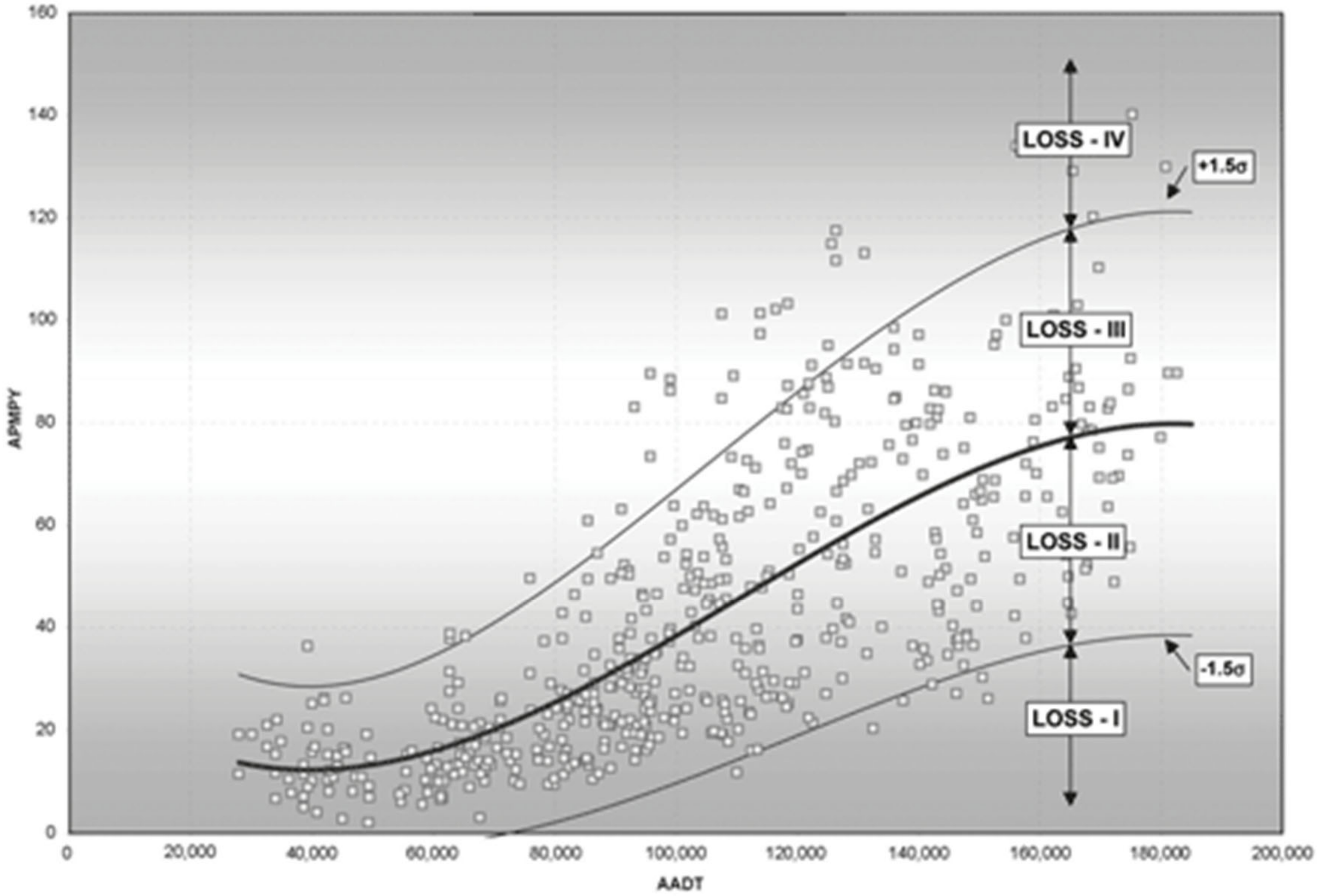
- Requires rigorous analysis
- May not properly capture the long-term mean of the sites (use average of the sites)

Level of Service of Safety

TABLE 8.2 LOSS categories (Kononov and Allery, 2003; AASHTO, 2010).

LOSS	Condition	Description
I	$\sigma < C < (\mu - 1.5\sigma)$	Low potential for crash reduction
II	$(\mu - 1.5\sigma) \leq C < \mu$	Low to moderate potential for crash reduction
III	$\mu \leq C < (\mu + 1.5\sigma)$	Moderate to high potential for crash reduction
IV	$C \geq (\mu + 1.5\sigma)$	High potential for crash reduction

Level of Service of Safety



Level of Service of Safety

Exercise 8.7

Using dataset 8.1, identify hazardous intersections based on the LOSS method. Use the same functional form and parameter estimates as in Exercise 8.6.

Calculate the standard deviation of predicted crashes using Eq. (8.12). For intersection 1, the standard deviation is estimated as follows:

$$\sigma = \sqrt{\alpha\mu^2} = \sqrt{0.2423 \times 43.6^2} = 21.5 \text{ crashes}$$

The LOSS for intersection 1 falls under category III.

The sites with LOSS IV are considered as hazardous locations, so intersections 11, 13, and 46 are detected.

PI using expected crashes (Empirical Bayes)

- Combines the site's crash history with the predicted crash frequency from the crash-frequency model.

Advantages	• Accurate estimation • Estimates a threshold • Considers long-term mean
Disadvantages	• Requires rigorous analysis

PI using expected crashes (Empirical Bayes)

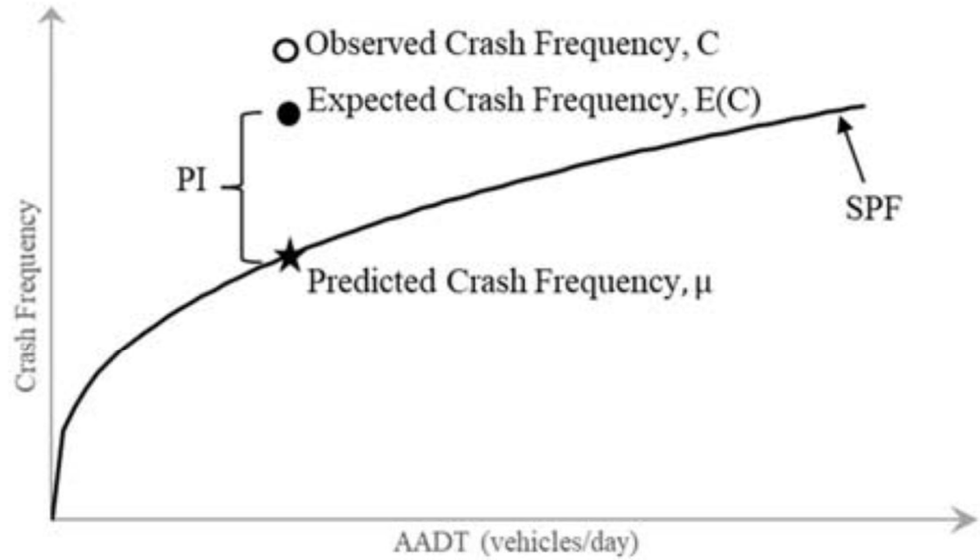


FIGURE 8.2 Graphical representation of potential for improvement using EB method.

$$PI_{EB} = E(C) - \mu$$

$$E(C) = w \times \mu + (1 - w) \times C$$

$$w = \frac{1}{1 + (\alpha \times \mu)}$$

Combined Criteria

Combined threshold

- each method's thresholds must be met

Individual threshold

- at least one method's threshold is met

Individual threshold and minimum criteria

- sites are ranked in descending order using one method and the site is detected once it reaches the minimum thresholds set for other criteria

Thank you