Research Article

TRR
JOURNAL OF THE TRANSPORTATION RESEARCH BOARD

Transportation Research Record
1–14

DOI: 10.1177/03611981231189734
journals.sagepub.com/home/trr
Sage

Safer Streets Priority Finder: An Open-Source Tool for Vulnerable Road User Safety Analysis and Prioritization

Jessica Schoner[1], Tara Tolford[2], Theja Putta[3], Maryam Izadi[2], Rachel Finfer[3], Daniel Jatres[4], Daniel Patterson[5], Jennifer Ruley[4], Jacob Nigro[3], and Robert Stickney[6]

Abstract

Vulnerable road user traffic deaths in the United States have increased in number and proportion over the last decade. This growing disparity points to a larger need to prioritize safety for vulnerable road users. Evaluating and predicting vulnerable road user crash risk is a data-intensive and complex process. This study aimed to make safety analysis easier and more accessible by (1) developing a modeling framework with minimal data input needs, (2) converting model outputs into cost equivalents to better link the results to project scoping processes, and (3) building this functionality into an online tool and dashboard. In this paper, we present an approach to modeling vulnerable road user crash risk that uses Bayesian probability updating and Markov chain Monte Carlo simulations to blend an existing published statistical model with simple roadway and crash data inputs, which we built into an online tool and dashboard called the Safer Streets Priority Finder. We applied the tool to crash data from the City of New Orleans and describe its application for roadway safety and transit planning use cases. Overall, in most contexts, we found that this modeling approach performed as well or better than sliding window analysis and traditional high injury networks, as it goes beyond just crash history, thus enabling it to estimate crash risk even when there is no history of crashes. This performance improvement, combined with ease of use, suggests the tool could improve on one of the most common safety analysis approaches used in field of transportation planning.

Keywords

pedestrians, bicycles, human factors, safety, analysis, safety, transportation safety management systems

Over the last decade, pedestrian and bicyclist traffic deaths have increased by over 50%, while overall traffic deaths decreased by 7.9% (*1*). Vulnerable road users (VRUs), particularly pedestrians and bicyclists, comprise an increasing fraction of all traffic deaths, growing from 12.6% in 2003 to 19.5% in 2018 (*1*), despite walking and bicycling combined representing less than 12% of all trips made in 2017 (*2*). Moreover, safety outcomes have continued to worsen in recent years: between the first halves of 2019 and 2022, the rate of pedestrian fatalities increased at a staggering pace of 18%, which was nine times higher than the rate of population growth. The Governors Highway Safety Association projected that about 3,434 pedestrians lost their lives in traffic accidents in the United States during the first half of 2022. This represents a 5% increase compared with the same period in 2021 and equates to 168 more pedestrian fatalities compared with the previous year (*3*).

This growing disparity points to a greater need to prioritize safety for VRUs. Designing safer roads that prioritize the needs of all users, but especially VRUs, is essential for reducing the overrepresentation of these

[1]Safe Streets Research & Consulting, LLC, Silver Spring, MD
[2]Department of Planning and Urban Studies, University of New Orleans Transportation Institute, New Orleans, LA
[3]Toole Design Group, LLC, Silver Spring, MD
[4]City of New Orleans, New Orleans, LA
[5]Cambridge Systematics, Inc., New York, NY
[6]New Orleans Regional Transit Authority, New Orleans, LA

Corresponding Author:
Maryam Izadi, mizadi@uno.edu

users in traffic crashes. This can be achieved through infrastructure improvements and innovative road design strategies that create safer and more inclusive road environments. This makes it essential to integrate safety measures for pedestrians and vehicles in areas where pedestrian crashes are more frequent (*4*, *5*). In response, jurisdictions have embraced initiatives (such as Vision Zero) to reduce or eliminate serious crashes. These initiatives rely on understanding the safety issues faced by pedestrians and bicyclists, and proactively and systemically addressing those issues across our roadway networks.

Although bicyclist and pedestrian crashes are too common and increasing year-to-year, they are statistically rare. This makes evaluating and predicting VRU crash risk a data-intensive and complex process. This study aimed to make safety analysis for VRUs easier and more accessible by

(1) Developing a modeling framework that can estimate risk on individual network segments using widely available inputs, such as U.S. Department of Transportation's (U.S. DOT) existing Pedestrian Fatality Risk Map (*6*);
(2) Converting outputs into cost equivalents to better link the results to planning and project scoping processes; and
(3) Building this functionality into an online tool and dashboard that makes the results accessible to practitioners without sophisticated statistical or geospatial skills.

This paper is organized as follows. First, we provide a brief scan of research related to evaluating crash risk on a network for VRUs (for the purpose of this paper limited to pedestrians and bicyclists). Next, we describe the modeling approach and methodology that were developed and built into a free and open-source web tool and dashboard called the Safer Streets Priority Finder (SSPF; Schoner et al. [*7*]), including a brief validation of the tool's output. We present an illustrative application and use case for this tool in the City of New Orleans. Finally, we share conclusions and recommendations for future work.

Literature Scan

Over the last two decades, many U.S. cities and states have enacted plans and policies to encourage active transportation. However, despite efforts to promote and support walking and bicycling, between 2009 and 2020, the number of pedestrian and bicycle fatalities nationwide has continued to trend upward, with 2020 surpassing previous records to become the deadliest and costliest year for both cyclists and pedestrians since 1990 (*8*). In response, jurisdictions have embraced initiatives to reduce or eliminate serious crashes involving VRUs. A key requirement of such initiatives is better understanding the safety needs of VRUs, which in turn calls for analytic tools for evaluating and predicting crash risk. A growing body of research has emerged documenting methods for crash analysis and predictive modeling. This literature scan summarizes current analytic methods and modeling approaches relevant to active transportation safety analysis and identifies gaps in knowledge and barriers to practice, which the resulting tool seeks to address.

Previous analyses have identified the locations of past crashes within a given geographic area, typically focusing on either top crash corridors overall or "hot spots" (typically intersections) where crashes have occurred with high frequency (*9–11*). Hot spot analysis is useful for understanding the scope of the problem, can highlight specific, problematic locations and/or roadway elements with high crash volumes, and is a relatively simple analytic technique. Statistical tools may be applied to confirm crash clusters are significant (*12*, *13*).

Hot spot analysis can identify locations (intersections, neighborhoods, roadway classifications, etc.) with a disproportionate share of crashes within a dataset (*14–18*). However, it can fail to reveal corridorwide problems and tends to be less effective for pedestrian and bicyclist crashes because they are relatively rare and there may be insufficient data to identify high-frequency crash locations. Bayesian statistical techniques can help to address this limitation by representing prior uncertainty about model parameters with a probability distribution, providing more nuanced inferences about limited data (*19*). Moreover, such analyses are based entirely on previous crash history and spatial proximity and may not consider systemic factors that are likely to contribute to crash incidence in the future.

In response to the limitations of simple crash frequency analysis, many cities have turned to the development of high injury networks (HINs) as a conceptual tool to characterize and develop effective crash reduction programs for urban environments (*20–23*). The goal of an HIN is to identify a limited subset of streets where serious crash density is highest. Sliding window analysis is a key analytic tool for the development of an HIN, used to identify crash clusters within flexibly sized "windows" representing overlapping segments along a street network rather than crashes at discrete locations, using previous crash history and linear proximity, while potentially capturing some elements of systemic risk. Sliding windows allow us to generalize the locations of crashes, reflecting the stochasticity in where crashes occur, while still acknowledging that locations along corridors tend to share characteristics. The result is a flexible measure of crash density that allows for a more consistent evaluation of crash injury distribution, which can greatly aid in project prioritization (*24*).

The systemic safety approach recognizes the limitations of looking at crash history alone; instead, networks are screened for conditions linked to crash outcomes—whether any one of those locations specifically have had crashes. Systemic evaluation involves the identification of which built environment and contextual factors influence crash frequency and severity and the degree of this influence (*25*, *26*). As a result, the entire roadway network may be screened to determine where crashes are more likely to occur, even if those locations show no crash history. A variety of studies have identified potentially significant factors, including traffic characteristics, land use attributes, transit access, and socioeconomic characteristics (*27*–*36*).

However, several of the tools available for systemic safety analysis have heavy data input requirements or require sophisticated geospatial or statistical techniques; both become barriers for agencies and jurisdictions aiming to reduce or eliminate deaths and serious injuries on their roadways (*37*). Operationalizing findings from systemic analyses that identify significant factors presents additional challenges. Development of several analytic tools to assess and address traffic risk (e.g., safety performance functions, crash modification factors, areawide crash rates) requires exposure data, which most jurisdictions currently lack (*38*, *39*).

Mansfield et al. combined facility-level factors (transportation system), neighborhoodwide data, and pedestrian fatality records from the Fatality Analysis Reporting System (FARS) between 2012 and 2016 to anticipate pedestrian fatality risk at census-tract-level across the United States; they utilized the results to develop a statistical model of fatal pedestrian crashes nationwide and a map viewer of the estimated number of fatal pedestrian crashes in each census tract (*6*). This model considers various factors to estimate risk for pedestrian fatalities, including vehicle miles traveled (VMT) density by roadway functional classification, intersection density, employment density, residential population density, activity mix index, and sociodemographics. The model was validated against an HIN for Los Angeles, showing the model applied in this research combined with HIN can significantly capture high-risk areas, and formed the foundation of the SSPF. However, this model does not link the results to specific network locations, which is a missing element that would allow practitioners to prioritize projects.

Overall, pedestrian and bicycle safety analysis has advanced considerably in recent years, providing a variety of complementary tools for analyzing crash data, and a robust suite of research identifying potentially significant factors that could be used to develop predictive crash models. However, some analytic approaches lack specificity at a subcorridor level or do not account for built environment factors and cannot provide significant insight into the root causes of crash risk. Moreover, performing more sophisticated analyses often involves large data requirements and requires adoption of advanced data mining and analytic techniques, for which local jurisdictions lack capacity. Finally, several of the analyses reviewed, although technically innovative, lack clear methods for linking model outputs to planning processes: to be useful to practitioners, analytic approaches need to be not only accessible to agency staff without a background in advanced statistics and modeling, but provide an output that can be directly applied in project identification, prioritization, and development processes. Inputs for cost–benefit analyses, such as a "no-build" cost value, are needed to help local and state officials and practitioners to make defined, targeted decisions around small-area and corridor-level investments with the greatest potential to prevent serious injuries and fatalities for VRUs, and to provide end users with a means of calculating associated costs of inaction where risks have been identified.

Methodology

As described in this paper's introduction, our team's goal for this project was to develop a method that could estimate risk on individual network segments using widely available data, to link the outputs to crash costs, and to embed any advanced geospatial processing in a user interface to make the analysis more accessible to practitioners, particularly those in jurisdictions without advanced in-house roadway safety analytics capabilities. U.S. DOT's Pedestrian Fatality Model (PFM) satisfied some, but not all, of our project's objectives (*6*); it estimates pedestrian fatality risk, using publicly available data with coverage of the entire United States, using an output unit that links to cost. The inputs reflect a broad range of factors associated with pedestrian deaths, including roadway conditions (functional class, VMT density, intersection density), land use characteristics (destinations, density), demographic/economic characteristics (race, ethnicity, age), and exposure correlates (transit/walking commuting) (*6*). However, the tract-level model could not account for variation within each tract; nor could it resolve risk for the busy arterials that define tract boundaries.

The project team therefore developed a method that built on the PFM, translating its tract-level outputs into segment-level estimates. Sliding window analysis, the precursor to developing an HIN, generalizes risk along corridors that tend to share similar characteristics—a nod to systemic analysis, even if the specific roadway risk factors are not analyzed. Our approach blends the PFM outputs with a sliding window process to transpose outputs onto the network. By combining these two approaches, we built an analytical framework that incorporates systemic pedestrian environment issues with observed crashes. For the model to be effective and appropriate, it needed to 1) be able to function with a limited amount of observed crash data, 2) be able to incorporate existing and future work into its process, 3) be scalable at the national level, and 4) provide information and a clear, understandable output for users with limited technical capabilities. To achieve these goals, the project team determined a Bayesian approach would be most suitable. Among other reasons, a Bayesian approach exhibits the following traits:

- Conducive for merging multiple models into one;
- Allows for implementation of prior information in the model, rather than a reliance on purely observed data;
- Has no limit on the amount of data needed to create an estimate;
- Allows for easy updating of results as more information is gained over time (i.e., as crashes are observed); and
- Allows for adjustment by tuning parameters, and relatively easy expansion of the model in future work.

Sliding Window Analysis

Sliding window analysis was used to summarize crashes on the network by mode and severity. Windows were set to 0.5 mi (800 m), and were stepped through the network in 0.1-mi (160-m) increments. These window sizes were chosen based on the authors' previous work conducting similar analysis. The window segments containing counts of crashes by mode and severity along them formed the unit of analysis for the model described in the next section. Furthermore, a severity-weighted crash score was calculated for each mode for building a traditional HIN, and the output was provided separately from the modeled outputs in the SSPF interface.

Model Formulation

The Bayesian network model (network model or NM) was formulated to estimate the expected rate of crashes by mode and severity for each window segment, given the following inputs:

1. The observed crashes by mode and severity on each window segment and in each tract,
2. PFM output for the area containing the segment, and
3. National crash rates per mile for the functional class within the window segment.

The latter two pieces of information were used to calculate our Bayesian priors. The observed crashes in each tract and on the window segment were then used to update this information and estimate a posterior likelihood of crashes occurring on the window segment. The NM itself contained two submodels. Both submodels utilized compound distributions—a gamma-Poisson and a beta-binomial. In both instances, conjugate priors were used to ensure the approach was both updatable and tractable.

Model 1. The first submodel (NM-GP) used a gamma-Poisson distribution to estimate the number of crashes occurring in each census tract, using the PFM output as a prior, which factors in variables related to both roadway risk and exposure. This model was applied separately for each mode and severity (note that PFM output—estimate of fatal pedestrian crashes—was used as a prior for all modes and severities). The model was formulated as follows:

Let

- yt = number of observed crashes in the past No years in census tract t;
- $pfmt$ = PFM output of annual fatal pedestrian crash rate p in census tract t;
- No = number of years over which crashes have been observed—typically 5; and
- Nw = number of years' equivalency used to weight PFM-based prior—equal to No for fatal and serious injury crashes, and $\frac{No}{5}$ for lower severity crashes.

The prior distribution for each tract was defined by shape parameter, $\alpha 0$, and rate parameter, $\beta 0$, as follows:

- $\alpha 0 = pfmt * Nw$
- $\beta 0 = Nw$

Then these values were updated to incorporate observed crashes,

- $\alpha = \alpha 0 + yt$
- $\beta = \beta 0 + N0$

Our final distribution was described, $\lambda \sim (\alpha, \beta)$

Model 2. The second submodel (NM-BB) used a beta-binomial distribution to allocate the tract's crashes to window segments, using national crash rates by functional class as a prior, as a proxy for roadway design elements associated with risk of both crashes occurring and the crashes' outcome being severe (e.g., motor vehicle travel speeds, number of lanes, motor vehicle annual average daily traffic). This model was applied separately for each mode and severity. Note that national crashes by functional class included fatal crashes only owing to data availability. National fatal pedestrian crash rate per mile for different functional classes of roads were calculated using 2015 to 2019 FARS data and 2016 mileage data from FHWA Highway Statistics. Window segments that coincided with the boundary of a census tract used tract-level data pooled across all adjacent tracts. The NM-BB model was formulated as follows:

Let

- pw = probability p that a crash within census tract t happens on window w (versus all other windows in the tract);
- cw = number of observed crashes, c, in tract t on window w;
- cw' = number of observed crashes, c, in tract t on all other windows besides window w;
- Cf = number of fatal crashes nationally, C, happening on functional class f;
- Cf' = number of fatal crashes nationally, C, happening on all other functional classes besides f; and
- mf = mileage, m, of functional class, f, in the United States.

The prior distribution for each window segment was defined by the number crashes, grouped into two mutually exclusive and exhaustive outcome states: "hits" (crash rate on this functional class) $\alpha 0$ and "misses" (crashes per mile occurring on other functional classes) $\beta 0$.

- $\alpha_0 = \frac{C_f}{m_f}$
- $\beta_0 = \frac{C_{f'}}{m_f}$

Then these values were updated to incorporate observed crashes as follows:

- $\alpha = cw + \alpha 0$
- $\beta = cw' + \beta 0$

Our final distribution was described, $\varphi \sim B(\alpha, \beta)$

Model Combination and Sampling. Given these two distributions, we generated random samples for each mode and severity on each window segment parameterized as follows:

- $\theta = Pois(\lambda)$, which estimates expected rate of crashes in the tract per year, and
- Crashes = $B(\theta, \varphi)$, given the expected number of crashes in the tract per year, estimates how many of them happen on this window segment.

Sampling was undertaken using R programming language and a package called RStan (*40*, *41*). The core functionality of the tool was implemented in R with the user interface developed using the Shiny package (*42*). RStan generated 2,000 samples across four Markov chain Monte Carlo simulations from each distribution for each unique combination of alpha and beta values in the study area. The results were the mean value across all samples from the binomial distribution. They represented an expected rate of crashes per year on each window segment, repeated for each mode and severity separately. Dividing by the window length in miles produced an annual crash rate per mile for each sliding window.

This crash rate represented the crash density, which was then joined to the original street segments. This was done by first dicing the roads into short segments of 160 m (step size) in length. Each short segment was then assigned the highest crash density associated with all the sliding windows coinciding with it. The short window crash density was then joined back to the original street segments by calculating the average crash density of its short segments weighted by their length.

Severity-Based Crash Cost

Established methodologies already exist for estimating the monetary value of deaths, injuries, and damaged property from traffic crashes, as well as the monetary cost of economic and societal impacts from crashes (*43*, *44*). The project team calculated national costs for crashes following the guidance set forth in Chapter 6 of FHWA's *Crash Costs for Highway Safety Analysis* (2018) (*45*). Costs were applied to the estimated number of crashes by severity, following the KABCO scale (i.e., Killed, Injury A, Injury B, Injury C, or Property damage only). The tool applies a default discount rate of 3% to reflect today's value of costs projected over a 5-year time horizon. Both the crash costs and the discount rate can be customized by tool users.

Calibration

Initial results from this model showed that the intranetwork relative distribution looked reasonable (i.e., the segments with the highest model outputs also had the most observed crashes and were consistent with local transportation practitioners' understanding of high-risk portions of the network). However, the magnitude of crashes and corresponding costs, if summed across the network, tended to be higher than the observed values over an equivalent period (two to three times higher in urban areas, and much higher in rural areas). To account for this difference, we scaled the output crash estimates so that the total crash cost from the model outputs for each mode in the entire study area was the same as the observed crash cost.

Validation

To evaluate the utility of the Safer Streets model relative to HIN for correlating with future (i.e., out-of-sample) crashes and identifying priority investment areas, our team developed a validation methodology that compared analysis outputs run on 2010 to 2014 crash data with crashes observed between 2015 and 2019. The model outputs were compared with those of the sliding window analysis to observe whether the model outperformed the sliding window analysis in identifying the segments that had crashes from the validation set. The input set was run through the tool's sliding window analysis and model components. The validation set was put through the sliding window process in the tool. The results from each of these processes were joined to the road network and segmented by sliding windows in the same manner. This minimized any biases introduced by the sliding window process in the way crashes are counted. For each street segment, this created three outputs for every unique combination of mode and severity:

1. Sliding window analysis using 2010 to 2014 crashes,
2. Safer Streets model using 2010 to 2014 crashes, and
3. Sliding window analysis using 2015 to 2019 crashes.

Street segments were sorted based on the outputs from either 2010 to 2014 sliding window analysis (Outcome 1) or the Safer Streets model (Outcome 2). For various percentile groupings of streets on these two outcomes, we summed the proportion of Outcome 3 that fell within this subset of the network. For example, in the top 5th percentile of the network according to Outcome 1, what percent of the 2015 to 2019 crashes falls on this network subset? We repeated this calculation for 20 different percentile thresholds (ranging from the 5th to the 100th in 5% increments) across four crash groupings: severe pedestrian crashes (Killed plus Injury A, or K + A), all pedestrian crashes, severe bicycle crashes, and all bicycle crashes; and across three different jurisdictions: the City of New Orleans, LA; Lincoln Parish, LA; and the City of Lowell, MA. This validation exercise helped us understand how the tool performs with various data sample sizes.

Table 1 summarizes the results. Columns 2 and 3 show the cumulative outcome findings for the 10th percentile network, Columns 3 and 4 show the same for the 25th percentile network, and Columns 5 and 6 contain cumulative distribution diagrams showing all results from 5% to 100% of the network. Across the board, we saw that for every subset of the network, the Safer Streets model's subset contained an equal or larger share of the out-of-sample crashes than the sliding window analysis. For example, for New Orleans severe pedestrian crashes, 17% of Outcome 3 was contained in the top 10% of the network based on the 2010 to 2014 sliding window score (Outcome 1), whereas the top 10% based on the Safer Streets model (Outcome 2) contained 37% of the Outcome 3 score. In Lincoln Parish, where our data contained fewer crashes owing to both the rural context and missing geolocations on some crash records, the sliding window networks contained very few of the severe pedestrian or bicycle crashes.

The magnitude of the outcomes for both models was much lower than the in-sample metrics used to characterize HIN (e.g., 70% of K + A crashes on 12% of the network). In our top 10% networks, we see the Safer Streets model capturing 0% to 41% of the out-of-sample crashes, and sliding window analysis capturing 0% to 38%. This was expected, both because out-of-sample predictions are harder than predicting input data, and because roadway investments in higher risk areas over time should cause the relative share of crashes on this subset to decline. Nonetheless, the Safer Streets model outperformed the sliding window analysis for severe crashes in all three test jurisdictions.

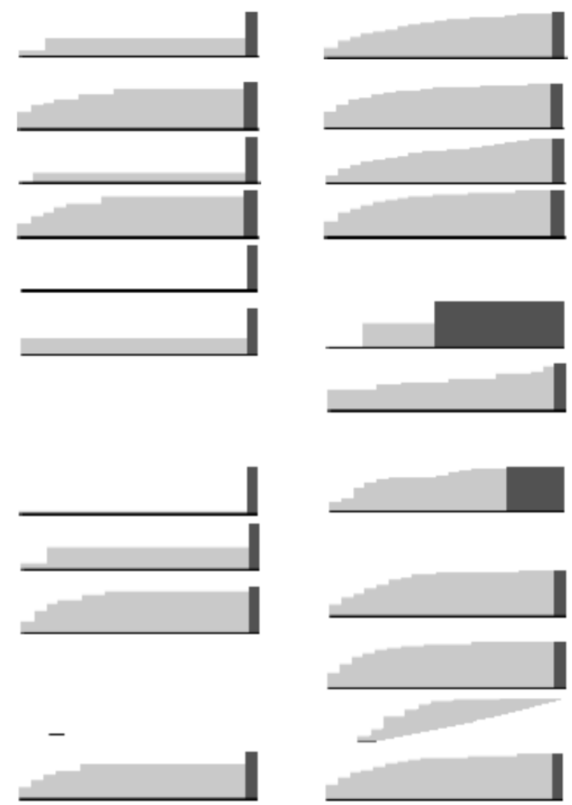
Further validation is needed to compare the magnitude of model outputs to observed crashes over time and explore differences by mode and severity, but at this stage, the Safer Streets model appears to perform comparably or slightly better than sliding window analysis, indicating that model output might be used alongside or instead of traditional HINs for informational purposes to screen for locations that may benefit from safety investment.

Use Case: City of New Orleans, LA, and New Orleans Regional Transit Authority

The project team developed the SSPF to address an identified need to better understand both the likelihood of serious crashes involving people walking and bicycling on the City of New Orleans' Road network, as well as a

Out-of-Sample Validation Results for Sliding Window Analysis and Safer Streets Model in Three Test Jurisdictions

	What percent of out-of-sample crashes overlap with the top 10% of the network?		What percent of out-of-sample crashes overlap with the top 25% of the network?		Cumulative distribution: y% of out-of-sample crashes overlap with the top x% of the network (dark gray = 100%)	
	Sliding window, %	Safer Streets model, %	Sliding window, %	Safer Streets model, %	Sliding window	Safer Streets model
New Orleans, LA pedestrian K + A crashes	17	37	42	60		
Pedestrian all crashes	53	54	67	75		
Bicycle K + A crashes	23	34	23	54		
Bicycle all crashes	46	53	75	77		
Lincoln Parish, LA pedestrian K + A crashes	0	0	0	56		
Pedestrian all crashes	38	48	38	56		
Bicycle K + A crashes*	NA	NA	NA	NA		
Bicycle all crashes	9	29	9	73		
Lowell, MA pedestrian K + A crashes	17	41	50	71		
Pedestrian all crashes	51	55	74	84		
Bicycle K + A crashes	3	29	3	72		
Bicycle all crashes	41	50	63	74		



*Sliding window and Safer Streets model ran on 2010 to 2014 data; out-of-sample crashes came from 2015 to 2019 data. Out-of-sample validation for Bicycle K + A crashes in Lincoln Parish was not possible because there were no valid geocoded fatal or Injury A bicycle crashes observed in the 2015 to 2019 data.
Note: NA = not available.

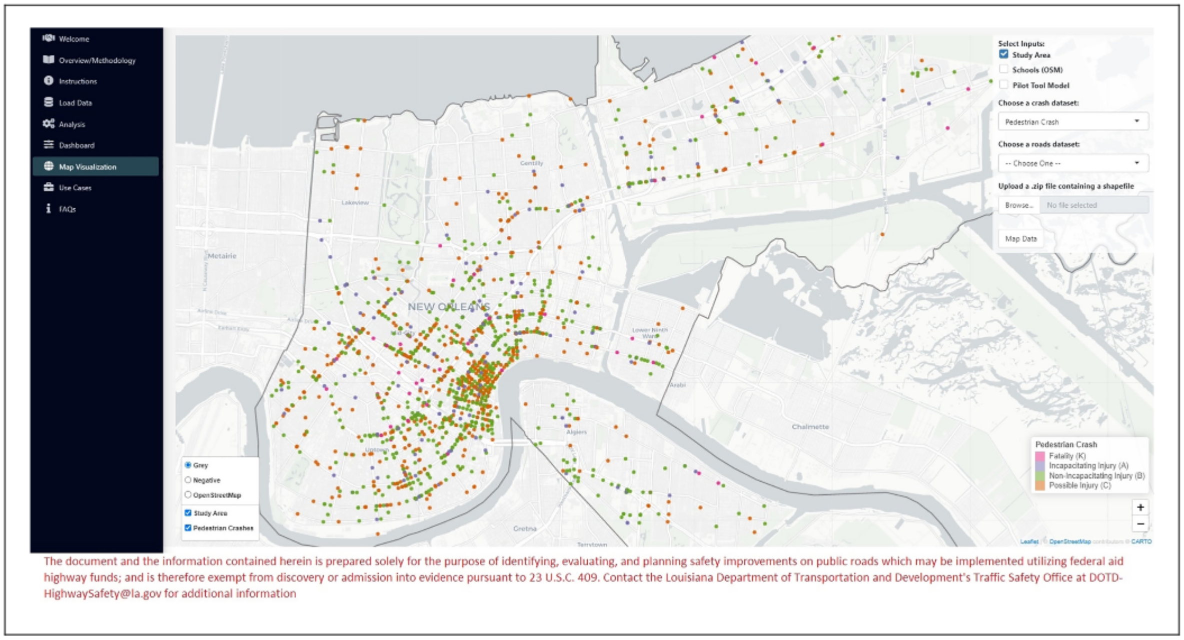


Figure 1. Pedestrian crashes in the City of New Orleans mapped by severity within the tool interface.

means by which to quantitatively rank segments (and thus, potential interventions) and estimate the costs of those crashes, in the context of a robust suite of pending and proposed Complete Streets interventions and implementation of a transit network redesign. This section highlights how the City of New Orleans and New Orleans Regional Transit Authority (RTA) have initially applied the tool to address outstanding safety questions.

Background

In 2012, FHWA designated the City of New Orleans as a pedestrian safety focus city, which led to the drafting of a pedestrian safety action plan. In 2015, FHWA updated this designation to include a focus on bicycle safety. Although the action plan identified crash hot spots that focused on severe and fatal injuries, these findings have resulted in prioritized implementation for 20 intersections. However, this effort has not led to a systemic approach to improving safety outcomes or an ability to identify and prioritize specific roadway segments and appropriate countermeasures that will have the greatest impact on human lives. Meanwhile, City is currently engaged in the implementation of a rapid-build protected bikeway network and is interested in developing evidence-based tools for project prioritization. City previously conducted preliminary assessments toward the development of an HIN, but found this method failed to account for factors likely to impact future risk (particularly in the context of ongoing, rapid changes to roadway networks).

RTA also collaborated on the development of this tool, with the intent that it can be used to evaluate a) areas where pedestrian safety enhancements along the transit network are most likely to benefit transit riders, and b) to explore additional analytic uses of the model itself, such as mapping and evaluating crashes involving transit vehicles to identify and prioritize systemic safety issues.

Application

The tool accepts users' crash data and street network data and allows them to be mapped to standardized formats the models can then use. Minimal data preprocessing is required for the crash data to ensure that the dataset includes a unique ID for each record, the crash year, the severity level (mapped to the KABCO scale of severity), and the mode of involved parties (i.e., pedestrians, bicyclists, or "other"). Each above-listed attribute should be assigned to a single column. Figures 1 and 2 illustrate an example of the user interface at interim steps in the process, specifically, visualizing the pedestrian and bicycle crash severities and their locations for New Orleans from 2015 to 2019.

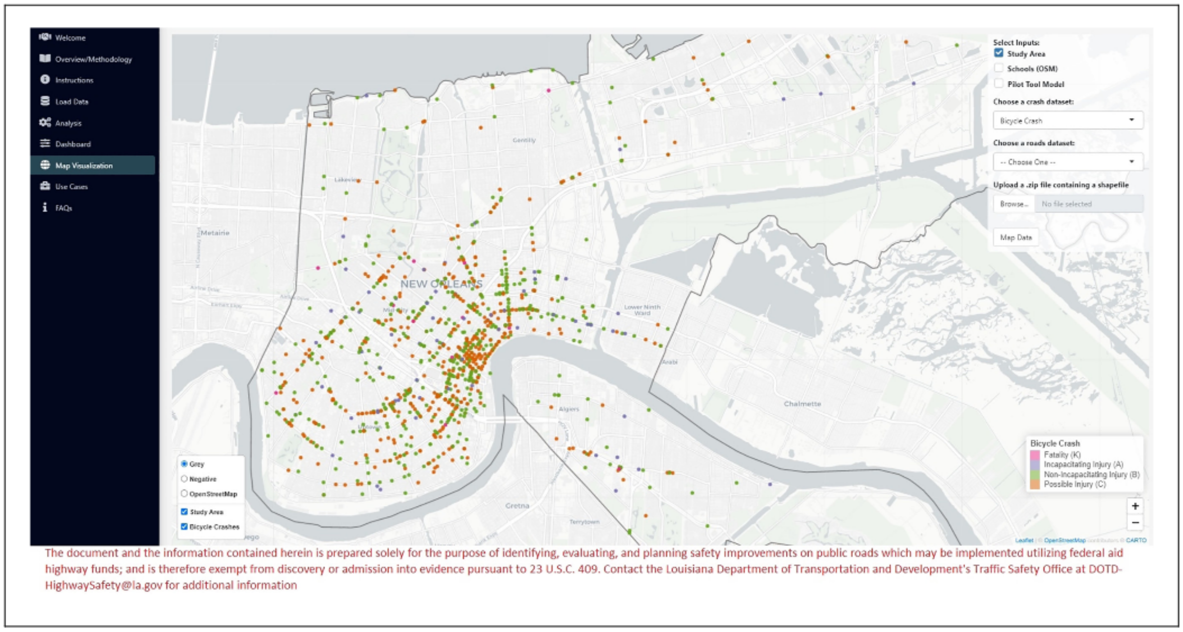


Figure 2. Bicycle crashes in the City of New Orleans mapped by severity within the tool interface.

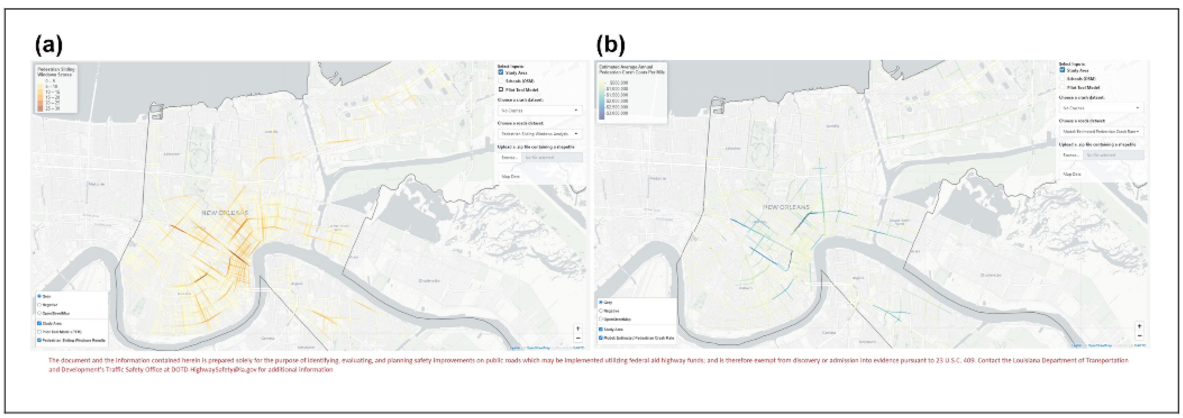


Figure 3. (*a*) Sliding window analysis and (*b*) Safer Streets model output for pedestrian crashes.

The tool's sliding windows analysis feature provides an initial look at historical crash density per mile by mode (shown for New Orleans in Figure 3a), whereas the Safer Streets model shows the estimated crash cost per mile, calculated as described in this paper (Figure 3b). This snapshot of areas where crashes previously occurred, and their concentration, weighted by severity, provides planners with an instant visualization of the road network to compare with past crash analyses.

Using the tool, the City of New Orleans was able to generate an estimation of costs associated with future pedestrian and bicyclist injuries and deaths if no action is taken to mitigate traffic hazards. The model output includes estimated annual and 5-year societal cost per mile for crashes for either mode. This estimate can be used to quickly compare the relative impact of investment on individual segments, or segments could be normalized by length (to estimate an absolute predicted crash cost for the segment) and aggregated to calculate estimated costs for an entire corridor or proposed project area. These results, in turn, could be used as an input for cost–benefit analysis relative to the cost of one or more proposed interventions to address systemic risk. In addition, the tool allows for quick visual comparison of how

Figure 4. Safer Streets model output for bicycle crashes.

model outputs vary by mode: although some corridors stand out for a high frequency of serious crashes for a specific mode, other corridors stand out for being relatively crash-dense for all modes. These visualizations can support mode-specific planning efforts or multimodal approaches to roadway redesign. Figure 4 shows an example of a Safer Streets model output for bicycle crashes.

These distinctions can be important in engaging the community and decision makers in safety priorities as well as targeted awareness campaigns for VRUs. For example, these results have created new collaboration opportunities with key stakeholders such as the New Orleans RTA because some of the higher pedestrian crash densities in the sliding window analysis correspond with high-ridership fixed route corridors where transit users may be at higher risk of traffic injuries. The built-in map visualization tool, moreover, allows the user to upload additional shapefiles (such as transit routes or stops) to visually assess the relationship of crash outcomes to features of interest, or the results can be exported for further analysis. RTA is using this information to adjust station locations, improve connections to and from station locations, and to lobby state and local partners to improve pedestrian amenities around its stops. Figure 5 shows an example of a Safer Streets model output for pedestrians overlaid with existing transit stops in the study area.

Finally, the tool dashboard and downloadable report synthesize high-level findings about the input data, including total number of crashes by year, severity, mode, and functional classification of roadway. The dashboard also identifies a concise list, "Highest Crash Corridors by Sliding Window Analysis," for pedestrians and bicyclists, as well as graphic representation of model fit for the dataset.

These features provide a timesaving asset for citywide and neighborhood-level planning, reducing time spent on data reduction and analysis. Figure 6 shows an excerpt from the dashboard showing the distribution of crashes by mode and severity, confirming the overrepresentation of pedestrians and bicyclists among fatal and serious injury crashes.

Use Case Outcomes

The City's previous approach to identifying problematic locations focused on hot spot analyses to derive priority locations for bicycle crash reduction strategies, and highlighted intersections rather than street segments. Based on the results of this model, the City of New Orleans has identified a concise list of street segments where changes in the built environment to improve walking and bicycling safety are most likely to have a measurable impact on future outcomes. In this case, several of these corridors have already recently undergone safety-oriented changes or are slated for future improvement. Thus, this analysis provides a useful baseline for future evaluation of project impacts. For other corridors, the model results

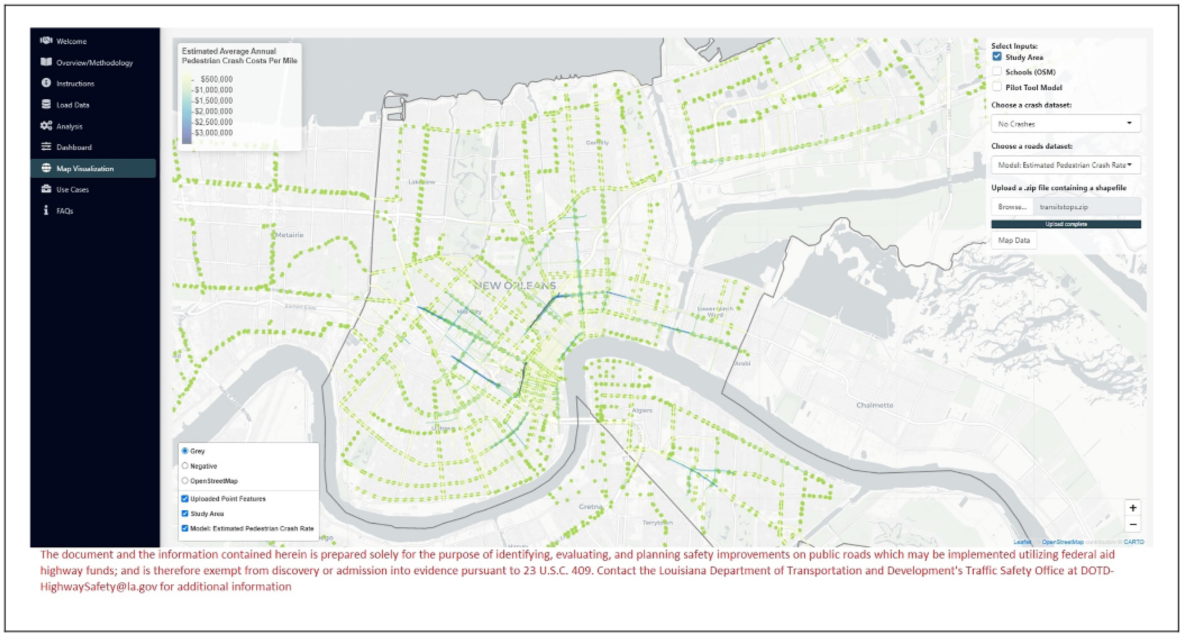


Figure 5. Map visualizer showing outputs for pedestrian risk model, overlaid with transit stop point features.

Figure 6. Dashboard excerpt displaying distribution of crashes by mode and severity.

provide a roadmap for selecting and proposing proven crash countermeasures, allocating or securing funding, and engaging the community, local leadership, and elected officials.

Meanwhile, RTA is using transit ridership as an additional input to identify agency priorities in discussion with City departments about future investments. Streets identified by the tool as having the highest level of risk for pedestrians coincided with some of the most important and heavily trafficked routes in the RTA system. These results have highlighted areas to advocate for safety improvements for RTA riders and are being integrated into an updated framework for ensuring safe and equitable pedestrian access to transit.

As future routes and improvements are planned, pedestrian walksheds (i.e., area within a defined walking range of a specific location) will be overlaid with crash maps to ensure stops are placed in a manner that reduces walkability barriers. Additionally, RTA plans to use the tool to visualize and analyze crashes involving transit vehicles and incorporate the tool results into operator training exercises. Having the tool to contextualize crashes involving transit vehicles could help the agency better discern whether a crash occurred from operator behavior or from high-risk road design.

Conclusions

The SSPF accepts simple end-user input data to build two different safety analyses with varying levels of complexity and reliance on crash history versus other risk factors. Our validation efforts to date have shown that the Safer Streets model performed at least as well at identifying locations with fatal and serious injury crashes as the simpler sliding window analysis, and in many cases outperformed it, indicating that with relatively lower data and technical demands, even small jurisdictions may be able to use the SSPF to elevate the quality and value of transportation network analysis for walking and bicycling compared with typical current practice. Neither of these analyses have been validated against other types of safety analyses (e.g., safety performance functions). This tool therefore is a good entry point to safety analysis—particularly among agencies where such work is not conducted on a routine basis owing to limited staff, data, and so forth—but should not replace any of the more advanced systemic analysis work agencies are already doing.

Outputs from the network model are expressed as costs, to better link the impact of crashes to the planning process. During tool development, we applied calibration factors to bring the total aggregate modeled crash cost by mode in the validation test study area into alignment with observed 5-year crash costs in the study area. These costs are still not a perfect representation of risk on the network: they are an approximation of aggregate, long-term trends, all things being equal. However, used cautiously and with an understanding of their limitations, the costs on the network can be summed or combined across corridors or study areas to estimate the potential opportunity for safety improvements.

We anticipate the primary users of this tool will be government agencies at the city or county level who are engaged directly in project prioritization and implementation. State DOTs and regional governments and entities could also use the tool by aggregating results across multiple counties to identify priority projects, inform funding requests and allocations, and support public engagement efforts. The tool's outputs may be used directly within the online application, substantially reducing the technical burden associated with safety analysis, enabling the product to be useful to a wide range of jurisdictions and users. The tool also offers spatial and safety analysis to nongovernmental advocacy groups, who previously may not have had these capabilities. The tool's open-source nature will afford future developers and researchers the opportunity to expand, adapt, or update components of the tool.

The project team identified multiple opportunities for future research and development, including

- Models and processes: The Bayesian model needs further validation against out-of-sample data sources and other analytical tools (e.g., safety performance functions). Future development should additionally include a model for motor vehicle crashes and/or other types of VRUs such as motorcycles. Updating the PFM with newer data as well as estimating tract-level models for bicycle- and motor vehicle crashes would likely improve model performance.
- Customization and user interface enhancements: User interface enhancements to facilitate building an HIN from either sliding window analysis or Safer Streets model output, as well as additional customization and reporting, expanding the default datasets available within the tool (such as preloading publicly available crash data); provision of functionality to sum crash costs in subareas or corridors.
- Usage guidance: Development of additional use cases in collaboration with partner agencies and end users.

Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: D. Jatres, J. Ruley, J. Schoner, R. Stickney, T. Putta; data collection: R. Finfer, J. Nigro, D.

Patterson, J. Schoner, T. Tolford, M. Izadi; analysis and interpretation of results: R. Finfer, D. Jatres, T. Putta, J. Ruley, J. Schoner, M. Izadi, R. Stickney, T. Tolford; draft manuscript preparation: R. Finfer, M. Izadi, D. Jatres, J. Nigro, T. Putta, J. Schoner, R. Stickney, T. Tolford. All authors reviewed the results and approved the final version of the manuscript.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The development of the Safer Streets Priority Finder and all associated work was supported by a grant from the U.S. Department of Transportation's Safety Data Initiative.

ORCID iDs

Jessica Schoner https://orcid.org/0000-0003-2340-8490
Tara Tolford https://orcid.org/0000-0002-9080-7590
Theja Putta https://orcid.org/0000-0003-2120-1103
Maryam Izadi https://orcid.org/0000-0003-2458-2271
Daniel Jatres https://orcid.org/0000-0001-6010-968X
Jennifer Ruley https://orcid.org/0009-0007-8561-1984
Jacob Nigro https://orcid.org/0000-0002-1654-0533

References

1. PBIC. Safety. https://www.pedbikeinfo.org/factsfigures/facts_safety.cfm. Accessed July 29, 2021.
2. NHTS. Popular Person Trips (PT) Statistics. https://nhts.ornl.gov/person-trips. Accessed July 29, 2021.
3. Governors Highway Safety Association. Pedestrian Traffic Fatalities by State (January-June 2022 Preliminary Data). *February* 2023. https://www.ghsa.org/sites/default/files/2023-02/GHSA%20Pedestrian%20Traffic%20Fatalities%20by%20State%2C%20January-June%202022%20-Preliminary%20Data.pdf. Accessed March 11, 2023.
4. Montella, A., S. Chiaradonna, A. C. de Saint Mihiel, G. Lovegrove, P. Nunziante, and R. R. Maria. Sustainable Complete Streets Design Criteria and Case Study in Naples, Italy. *Sustainability*, Vol. 14, No. 20, 2022, p. 1314.
5. Riccardi, M. R., F. Galante, A. Scarano, and A. Montella. Econometric and Machine Learning Methods to Identify Pedestrian Crash Patterns. *Sustainability*, Vol. 14, No. 22, 2022, p. 15471.
6. Mansfield, T. J., D. Peck, D. Morgan, B. McCann, and P. Teicher. The Effects of Roadway and Built Environment Characteristics on Pedestrian Fatality Risk: A National Assessment at the Neighborhood Scale. *Accident Analysis & Prevention*, Vol. 121, 2018, pp. 166–176.
7. Schoner, J., T. Putta, T. Tolford, J. Daniel, and M. Izadi. Safer Streets Priority Finder. https://www.saferstreetspriorityfinder.com. Accessed May 26, 2022.
8. BTS. Transportation Fatalities by Mode. https://www.bts.gov/content/transportation-fatalities-mode. Accessed November 18, 2020.
9. Abdel-Aty, M., and A. Pande. Crash Data Analysis: Collective vs. Individual Crash Level Approach. *Journal of Safety Research*, Vol. 38, 2007, pp. 581–587.
10. Brozen, M., and D. Shockley. Towards Best Practice Collision Analysis for Vision Zero Programs. Presented at 95th Annual Meeting of the Transportation Research Board, Washington, D.C., 2016.
11. Jahangiri, A., M. Hasani, I. N. Sener, S. Munira, J. M. Owens, B. Appleyard, S. Ryan, S. M. Turner, and S. Ghanipoor Machiani. *Data Mining to Improve Planning for Pedestrian and Bicyclist Safety*. 2019. https://vtechworks.lib.vt.edu/bitstream/handle/10919/96338/01-003_Final%20Research%20Report_Final.pdf?sequence=1&isAllowed=y. Accessed January 22, 2021.
12. Kim, K., and E. Y. Yamashita. Using Ak-Means Clustering Algorithm to Examine Patterns of Pedestrian Involved Crashes in Honolulu, Hawaii. *Journal of Advanced Transportation*, Vol. 41, No. 1, 2007, pp. 69–89.
13. Tolford, T., J. Renne, and B. Fields. Development of Low-Cost Methodology for Evaluating Pedestrian Safety in Support of Complete Streets Policy Implementation. *Transportation Research Record: Journal of the Transportation Research Board*, 2014. 2464: 29–37.
14. Morency, P. From Targeted "Black Spots" to Area-Wide Pedestrian Safety. *Injury Prevention*, Vol. 12, No. 6, 2006, pp. 360–364.
15. Jang, K., S. H. Park, S. Kang, K. H. Song, and S. Chung. Evaluation of Pedestrian Safety: Geographical Identification of Pedestrian Crash Hotspots and Evaluating Risk Factors for Injury Severity. Presented at 92nd Annual Meeting of the Transportation Research Board, Washington, D.C., 2013.
16. Schuurman, N., J. Cinnamon, V. A. Crooks, and S. M. Hameed. Pedestrian Injury and the Built Environment: An Environmental Scan of Hotspots. *BMC Public Health*, Vol. 9, No. 1, 2009, pp. 1–10.
17. Schneider, R. J., F. R. Proulx, R. L. Sanders, and H. Moayyed. United States Fatal Pedestrian Crash Hot Spot Locations and Characteristics. *Journal of Transport and Land Use*, Vol. 14, No. 1, 2021, pp. 1–23.
18. Lee, C., and M. Abdel-Aty. Comprehensive Analysis of Vehicle–Pedestrian Crashes at Intersections in Florida. *Accident Analysis & Prevention*, Vol. 37, No. 4, 2005, pp. 775–786.
19. Lynch, S. M.Basics of Bayesian Statistics. In *Introduction to Applied Bayesian Statistics and Estimation For Social Scientists* (S. M. Lynch, ed.), Springer, New York, NY, 2007, pp. 47–75.
20. DVPRC. Philadelphia Crash Analysis Standards & Recommendations. 2018. https://www.dvrpc.org/Reports/17068.pdf. Accessed November 18, 2020.
21. Saxton, J. A High-Injury Network for Atlanta: How are Severe and Fatal-Injury Crashes Concentrated on Atlanta's Streets? *gatech*. 2018. https://repository.gatech.edu/server/api/core/bitstreams/66b105b2-6967-4d18-857f-fb495370dd0e/content. Accessed December 10, 2020.

22. San Francisco Department of Public Health. Vision Zero High Injury Network: 2017 Update - A Methodology for San Francisco, CA. 2017. https://www.sfdph.org/dph/files/EHSdocs/PHES/VisionZero/Vision_Zero_High_Injury_Network_Update.pdf. Accessed November 13, 2020.
23. High Injury Network. City of Los Angeles. https://geohub.lacity.org/datasets/4ba1b8fa8d8946348b29261045298a88_0/explore?location = 34.019200%2C-118.405852%2C11.15. Accessed October 27, 2020.
24. Fitzpatrick, K., R. E. Avelar, and S. M. Turner. *Identification of High Pedestrian Crash Locations*. Report No. FHWA-HRT-17-107. United States, Federal Highway Administration. Office of Safety Research and Development, Washington, D.C., 2018.
25. Thomas, L., L. Sandt, C. Zegeer, W. Kumfer, K. Lang, B. Lan, Z. Horowitz, et al. *Systemic Pedestrian Safety Analysis*. Project No. 17-73. Transportation Research Board. 2018.
26. Miranda-Moreno, L. F., P. Morency, and A. M. El-Geneidy. The Link Between Built Environment, Pedestrian Activity and Pedestrian–Vehicle Collision Occurrence at Signalized Intersections. *Accident Analysis & Prevention*, Vol. 43, No. 5, 2011, pp. 1624–1634.
27. Yu, C. Y. Built Environmental Designs in Promoting Pedestrian Safety. *Sustainability*, Vol. 7, No. 7, 2015, pp. 9444–9460.
28. Manepalli, U. R. R., and G. H. Bham. Identification of Crash-Contributing Factors: Effects of Spatial Autocorrelation and Sample Data Size. *Transportation Research Record: Journal of the Transportation Research Board*, 2013. 2386: 179–188.
29. Aziz, H. A., S. V. Ukkusuri, and S. Hasan. Exploring the Determinants of Pedestrian–Vehicle Crash Severity in New York City. *Accident Analysis & Prevention*, Vol. 50, 2013, pp. 1298–1309.
30. Schneider, R. J., M. C. Diogenes, L. S. Arnold, V. Attaset, J. Griswold, and D. R. Ragland. Association Between Roadway Intersection Characteristics and Pedestrian Crash Risk in Alameda County, California. *Transportation Research Record: Journal of the Transportation Research Board*, 2010. 2198: 41–51.
31. Ossenbruggen, P. J., J. Pendharkar, and J. Ivan. Roadway Safety in Rural and Small Urbanized Areas. *Accident Analysis & Prevention*, Vol. 33, No. 4, 2001, pp. 485–498.
32. Osama, A., and T. Sayed. Evaluating the Impact of Connectivity, Continuity, and Topography of Sidewalk Network on Pedestrian Safety. *Accident Analysis & Prevention*, Vol. 107, 2017, pp. 117–125.
33. Harvey, C., and L. Aultman-Hall. Urban Streetscape Design and Crash Severity. *Transportation Research Record: Journal of the Transportation Research Board*, 2015. 2500: 1–8.
34. Thomas, L., B. Lan, R. L. Sanders, A. Frackelton, S. Gardner, and M. Hintze. Changing the Future? Development and Application of Pedestrian Safety Performance Functions to Prioritize Locations in Seattle, Washington. *Transportation Research Record: Journal of the Transportation Research Board*, 2017. 2659: 212–223.
35. Clifton, K. J., C. V. Burnier, and G. Akar. Severity of Injury Resulting from Pedestrian–Vehicle Crashes: What Can We Learn from Examining the Built Environment? *Transportation Research Part D: Transport and Environment*, Vol. 14, No. 6, 2009, pp. 425–436.
36. Coleman, H., and K. Mizenko. *Pedestrian and Bicyclist Data Analysis [Traffic Safety Facts]: Research Note*. Report No. DOT HS 812 205. United States, National Highway Traffic Safety Administration, Washington, D.C.https://rosap.ntl.bts.gov/view/dot/36726. Accessed October 29, 2020.
37. Montella, A., M. Aria, A. D'Ambrosio, and F. Mauriello. Data-Mining Techniques for Exploratory Analysis of Pedestrian Crashes. *Transportation Research Record: Journal of the Transportation Research Board*, 2011. 2237: 107–116.
38. Turner, S., I. N. Sener, M. E. Martin, S. Das, R. C. Hampshire, K. Fitzpatrick, and L. J. Moln. *Synthesis of Methods for Estimating Pedestrian and Bicyclist Exposure to Risk at Areawide Levels and on Specific Transportation Facilities*. Report No. FHWA-SA-17-041. United States, Federal Highway Administration. Office of Safety, Washington, D.C., 2017. https://rosap.ntl.bts.gov/view/dot/36098. Accessed October 12, 2020.
39. Lyon, C., and B. Persaud. Pedestrian Collision Prediction Models for Urban Intersections. *Transportation Research Record: Journal of the Transportation Research Board*, 2002. 1818: 102–107.
40. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, 2013.
41. Stan Development Team. *RStan: The R Interface to Stan*. R Package Version 2, No. 1, 522. 2016. https://mc-stan.org/rstan/.
42. Chang, W., J. Cheng, J. Allaire, Y. Xie, and J. McPherson. *Shiny: Web Application Framework for R*, R Package Version 1, No. 5. 2017. https://shiny.posit.co/.
43. Schneider, H. *Louisiana Crash Costs*. Louisiana Highway Safety Research Group, 2020. https://carts.lsu.edu/admin/uploads/2019_Louisiana_Crash_Cost_Estimates_2f7f533350.pdf. Access January 26, 2021.
44. Blincoe, L., T. R. Miller, E. Zaloshnja, and B. A. Lawrence. *The Economic and Societal Impact of Motor Vehicle Crashes, 2010*. Revised No. DOT HS 812 013. U.S. Department of Transportation, National Highway Traffic Safety Administration, Washington, D.C., 2015.
45. Harmon, T., G. B. Bahar, and F. B. Gross. *Crash Costs for Highway Safety Analysis*. Report No. FHWA-SA-17-071. United States, Federal Highway Administration. Office of Safety, Washington, D.C., 2018.