DEVELOPING DECISION SUPPORT TOOLS FOR THE IMPLEMENTATION OF BICYCLE AND PEDESTRIAN SAFETY STRATEGIES

Deo Chimba, PhD., P.E., PTOE
Associate Professor
Civil Engineering Department
Tennessee State University
3500 John A. Merritt Blvd, Nashville, TN 37209
Phone: 615-963-5430
Email: dchimba@tnstate.edu
STATEMENT OF THE PROBLEM

• Tennessee Department of Transportation (TDOT) has an extensive road safety audit program which uses criteria based on the ratio of crashes to average daily traffic but does not target locations with a high number of bike/pedestrians crashes since there are no bicycle and pedestrian counts.

• A robust methodology is not currently available to identify bicycle and pedestrian high-crash locations in Tennessee.

• The challenge is allocating funds, from TDOT’s Highway Safety Improvement Program (HSIP), equitably among rural and urban areas in a way that is most effective at reducing bicycle and pedestrian fatalities and incapacitating injuries.
Research Questions?

- Are there spatial variations in pedestrian and bicycle crashes?
- How do spatial variations in pedestrian and bicycle crashes associate with socioeconomic and demographic factors?
- What framework can be adopted to implement bicycle and pedestrian safety strategies?
Develop Data-Driven Policy Framework

- Develop criteria and conditions for the systematic identification of bicycle and pedestrian high-crash locations in the state.
- These criteria will rate each crash-prone location based on injuries and fatalities, coupled with exposure.
- Develop a systematic framework and rating system for future years’ so that the analysis can be replicated in the future with less effort.
- Prioritize funding for improvements. To support the development of a data-driven draft policy for prioritizing and maximizing the effectiveness of HSIP fund allocation.
- The policy framework will be developed with the support of TDOT staff.
- This policy framework will direct current and future decision makers at TDOT and other agencies in the prioritization of funding.
Study Approach

- Developed a framework to identify bicycle and pedestrian high crash locations for safety improvement prioritization focusing on Population, Demographic and Socioeconomic Spectra in Tennessee.

- Research approach comprised in-depth analysis using a combination of existing data, literature review, GIS, cluster analysis, and advanced statistical modeling to examine and identify bicycle and pedestrian high-crash locations.

- Relevant data from each of the selected study locations was integrated into a Geographic Information System (GIS).

- The data included crashes, roadway geometry, population, demographics and economic, and traffic.

- The study used the gathered data and information to develop safety performance functions (SPF) to identify magnitude and characteristics of variables associated with pedestrian and bicycle safety hazardous locations (black spots).

- From the SPF, the research developed tool to evaluate the expected number of crashes at block and county levels for given set of population, demographics and socioeconomic data in Tennessee.
Data
Three types of data were used;

- Crash data
- Socioeconomic data
- Demographic data
TDOT Crash Database-TRIMS
Socioeconomic and Demographic data

- TIGER Products
  (Topologically Integrated Geographic Encoding and Referencing)

- https://www.census.gov/geo/maps-data/data/tiger-data.html
Data

- Obtained from TDOT traffic crash database
- 5 years 2008-2012 data: 5,845 pedestrian crash records
- 5 years 2008-2012 data: 2,185 bicycle crash records

Crash data

- US census bureau, 2006-2010 America Community Survey
- Block group data for Tennessee
- Income, Car ownership, poverty status, Transport mode to work

Socioeconomic data

- US census bureau, 2006-2010 America Community Survey
- Block group data for Tennessee
- Population counts, age, race

Demographic data
## Crash Data Statistics

<table>
<thead>
<tr>
<th>Year of Crash</th>
<th>Pedestrian</th>
<th>Bicyclist</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>1091</td>
<td>450</td>
</tr>
<tr>
<td>2009</td>
<td>1101</td>
<td>405</td>
</tr>
<tr>
<td>2010</td>
<td>1185</td>
<td>385</td>
</tr>
<tr>
<td>2011</td>
<td>1241</td>
<td>487</td>
</tr>
<tr>
<td>2012</td>
<td>1227</td>
<td>458</td>
</tr>
<tr>
<td><strong>Grand Total</strong></td>
<td><strong>5845</strong></td>
<td><strong>2185</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type of Crash</th>
<th>Pedestrian</th>
<th>Bicyclist</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fatal</td>
<td>389</td>
<td>33</td>
</tr>
<tr>
<td>Incapacitating Injury</td>
<td>1109</td>
<td>279</td>
</tr>
<tr>
<td>Non- Incapacitating Injury</td>
<td>4051</td>
<td>1603</td>
</tr>
<tr>
<td>Prop Damage (over)</td>
<td>118</td>
<td>115</td>
</tr>
<tr>
<td>Prop Damage (under)</td>
<td>178</td>
<td>155</td>
</tr>
<tr>
<td><strong>Grand Total</strong></td>
<td><strong>5845</strong></td>
<td><strong>2185</strong></td>
</tr>
</tbody>
</table>

![Graph showing the number of crashes over years for Pedestrians and Bicyclists](image)
Spatial distribution of crashes
<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population density (1000 per sq. Mile)</td>
<td>1.62</td>
<td>2.53</td>
<td>0.00</td>
<td>89.44</td>
</tr>
<tr>
<td>Population below 15 years of age (%)</td>
<td>19.02</td>
<td>7.76</td>
<td>0.00</td>
<td>59.33</td>
</tr>
<tr>
<td>Population from 15 to 64 years of age (%)</td>
<td>66.98</td>
<td>8.36</td>
<td>11.80</td>
<td>100.00</td>
</tr>
<tr>
<td>Population commuting to work by private cars (%)</td>
<td>95.84</td>
<td>5.81</td>
<td>0.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Population commuting to work by walking (%)</td>
<td>0.83</td>
<td>2.89</td>
<td>0.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Median household income (&quot;000&quot; $)</td>
<td>45.42</td>
<td>24.35</td>
<td>0.00</td>
<td>247.36</td>
</tr>
<tr>
<td>Housing units with no vehicles (%)</td>
<td>6.94</td>
<td>9.47</td>
<td>0.00</td>
<td>83.97</td>
</tr>
</tbody>
</table>
Cluster Analysis

Legend
- Pedestrian Crashes
- Significant Crash Cluster
- Davidson County
- Block Group Boundary
Where are high risk census block groups?
Where are these clusters?
Developing Safety Performance Functions (SPFs)

- \( y_i \) number of crashes occurring in a certain period at a site \( i \)
- \( \lambda_i \) is the Poisson parameter for site \( i \), which is equal to site expected number of crashes at a period, \( E (y_i) \).
- Poisson assumes the mean = Variance

**Poisson**

\[
P(y_i) = \frac{\exp(-\lambda_i) \cdot \lambda_i^{y_i}}{y_i!}
\]

**Negative Binomial**

\[
P(y) = \frac{\tau(y+\alpha^{-1})}{\tau(\alpha^{-1})\tau(y+1)} \left[ \frac{1}{1+\alpha \mu} \right]^{1/\alpha} \left[ \frac{\alpha \mu}{1+\alpha \mu} \right]^y
\]

\[ \mu = E(y_i) = \exp(X_i \beta) \]

For crash data the mean ≠ Variance

- \( \text{VAR } (y_i) > E (y_i) \) — Overdispersion
- \( \alpha \) is the overdispersion factor
- \( \mu \) is the mean of crashes
# Pedestrian Crashes

## What are the associated factors-Block Group?

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Z</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population density (1000 per sq. mile)</td>
<td>0.117</td>
<td>7.77</td>
<td>0.000</td>
</tr>
<tr>
<td>Population below 15 years of age (%)</td>
<td>-0.008</td>
<td>-2.08</td>
<td>0.037</td>
</tr>
<tr>
<td>Population from 15 to 64 years of age (%)</td>
<td>0.014</td>
<td>3.76</td>
<td>0.000</td>
</tr>
<tr>
<td>Population commuting to work by private cars (%)</td>
<td>-0.038</td>
<td>-7.12</td>
<td>0.000</td>
</tr>
<tr>
<td>Population commuting to work by walking (%)</td>
<td>0.0298</td>
<td>2.34</td>
<td>0.019</td>
</tr>
<tr>
<td>Median household income (&quot;000&quot; $)</td>
<td>-0.0108</td>
<td>-7.34</td>
<td>0.000</td>
</tr>
<tr>
<td>Housing units with no vehicles (%)</td>
<td>0.0308</td>
<td>8.86</td>
<td>0.000</td>
</tr>
<tr>
<td>Constant</td>
<td>-4.4198</td>
<td>-7.14</td>
<td>0.000</td>
</tr>
</tbody>
</table>

| Population Exposure alpha                                    | 1.586       |

Transportation Research Center

for Livable Communities
What are the associated factors-County?

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>z</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population below 15 years of age (%)</td>
<td>-0.0281</td>
<td>-0.91</td>
<td>0.362</td>
</tr>
<tr>
<td>Population from 15 to 64 years of age (%)</td>
<td>0.0231</td>
<td>0.91</td>
<td>0.364</td>
</tr>
<tr>
<td>Population of White (%)</td>
<td>-0.0461</td>
<td>-2.08</td>
<td>0.038</td>
</tr>
<tr>
<td>Population of African American (%)</td>
<td>-0.0368</td>
<td>-1.6</td>
<td>0.109</td>
</tr>
<tr>
<td>Population of Hispanic (%)</td>
<td>0.0546</td>
<td>1.64</td>
<td>0.101</td>
</tr>
<tr>
<td>Population commuting to work by private cars (%)</td>
<td>-0.0705</td>
<td>-1.13</td>
<td>0.257</td>
</tr>
<tr>
<td>Population commuting to work by walking (%)</td>
<td>-0.2909</td>
<td>-1.64</td>
<td>0.102</td>
</tr>
<tr>
<td>Median household income (&quot;000&quot; $)</td>
<td>-0.0025</td>
<td>-1.91</td>
<td>0.056</td>
</tr>
<tr>
<td>Housing units with no vehicles (%)</td>
<td>0.0848</td>
<td>2.37</td>
<td>0.018</td>
</tr>
<tr>
<td>Constant</td>
<td>1.9170</td>
<td>0.3</td>
<td>0.768</td>
</tr>
</tbody>
</table>

Population Exposure
alpha

0.11
Developed Crash Prediction Model (SPF)

$\mu = \exp [\ln(P) - 0.028A + 0.023B - 0.046C - 0.037D + 0.055E - 0.071F - 0.291G - 0.003H + 0.085I + 1.917]$

Where:
- $\mu$: Number of pedestrian crashes
- $P$: Population of a County
- $A$: Population below 15 years of age (%)  
- $B$: Population from 15 to 64 years of age (%)  
- $C$: Population of White (%)  
- $D$: Population of African American (%)  
- $E$: Population of Hispanic (%)  
- $F$: Population commuting to work by private cars (%)  
- $G$: Population commuting to work by walking (%)  
- $H$: Median household income ("000" $)  
- $I$: Housing units with no vehicles (%)
Prediction accuracy

RSquare = 0.9628
Integrating SPFs with Access Database

• **User friendly**: Unlike crash prediction models that are expressed in form of complicated equations and time consuming; this tool simplifies this process.

**It is built in form of a database**: With huge amounts of data now available, local and national agencies are now building their database.

• It helps users to gain more insight into the **relationships** between crashes and sociodemographic factors by varying the values of contributing factors.
Flow Chart of the Decision Support Tool

TIGER From ACS file → ArcGIS Geodatabase (.mdb) → Clean Integrated file → Output

Crash Data From E-TRIMS TDOT → ArcGIS Geodatabase (.mdb) → Access Database → Build Queries & SQL Script
Interface of Decision Support system

View County Crashes

STATE: TN
COUNTY ID: 9
COUNTY: Carroll
ACTUAL CRASHES: 11
SEARCH COUNTY: Carroll
PREDICTED CRASHES: 10
<table>
<thead>
<tr>
<th>Category</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>State</td>
<td>TN</td>
</tr>
<tr>
<td>County</td>
<td>Carroll</td>
</tr>
<tr>
<td>Population</td>
<td>28644</td>
</tr>
<tr>
<td>Population with age &lt; 15 (%)</td>
<td>18.3</td>
</tr>
<tr>
<td>Population with age from 15 to 64 (%)</td>
<td>64.2</td>
</tr>
<tr>
<td>Population of white (%)</td>
<td>84.7</td>
</tr>
<tr>
<td>Population of black (%)</td>
<td>9.9</td>
</tr>
<tr>
<td>Population of hispanic (%)</td>
<td>1.9</td>
</tr>
<tr>
<td>Population commuting to work by private mode</td>
<td>98.2</td>
</tr>
<tr>
<td>Population commuting to work by walking (%)</td>
<td>0.5</td>
</tr>
<tr>
<td>Median household income (&quot;000&quot; $)</td>
<td>36</td>
</tr>
<tr>
<td>Housing units without vehicles (%)</td>
<td>5.8</td>
</tr>
</tbody>
</table>
Let’s look at it

..\TOOL\Decision Support Tool.accdb

Applications?
Conclusions and Recommendations

- Implement design practices that accommodate Pedestrian and bicycle needs;
  - Sidewalks
  - Bike lanes
  - Shared lanes

- Reduce speeds on;
  - Roadways serving as boundaries of Block groups
  - Roadways crossing high crash Block groups
  - Shared lanes
  - School speed limits*

- Future direction:
  - Consider more variables; Roadway, Vehicle, Driver
  - Collect actual pedestrian volumes
  - Represent resulting crashes on map
Deliverables

Conference proceedings and presentations


4. Kidando, E, **Musinguzi, A** and Chimba, D. “Bayesian hierarchical analysis of pedestrian crashes and socio-demographic factors” Presented at the 2nd Summer Conference on Livable Communities, Kalamazoo, MI, 7/23/2015. **Award of best Student poster presentation**


**Papers under peer review**

6. **Musinguzi, A** and Chimba D. “Using kernel density to evaluate dependence of pedestrian crashes on demographic and socioeconomic factors”.

7. **Musinguzi, A** and Chimba, D “Adaptive neuro-fuzzy inference system (ANFIS) approach for pedestrian injury analysis”.


Transportation Research Center for Livable Communities