William W. Hay Railroad Engineering Seminar

Railroad Grade Crossing Accident Analysis at Microscopic and Macroscopic levels

Ray Benekohal
Professor
University of Illinois at Urbana-Champaign

Juan Medina
Postdoctoral Research Associate
University of Illinois at Urbana-Champaign

Date: Friday, November 14, 2014
Time: Seminar Begins 12:15
Location: Newmark Lab, Yeh Center, Room 2311

Sponsored by RAILTEC UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
RAILROAD GRADE CROSSING ACCIDENT ANALYSIS AT MICROSCOPIC AND MACROSCOPIC LEVELS

Rahim F. Benekohal
Juan C. Medina

William W. Hay Railroad Engineering Seminar
Nov 14, 2014 – University of Illinois at Urbana-Champaign
Outline

1 – Introduction
2 – Research Questions
3 – A Micro Approach
4 – State of Practice – The U.S. DOT Model
5 – A Combined Model
7 – Accuracy of Predictions / Rankings
8 – Future Research
Introduction

Grade crossing accidents are more likely to be more severe, more costly, and to involve a fatality than other highway crashes (NCHRP 755, 2013)

Fatalities at Public Grade Crossings

Fatal accidents have remained at ~10% of accidents for the last 25 years

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Incidents</td>
<td>6526</td>
<td>4979</td>
<td>3489</td>
<td>3085</td>
<td>1916</td>
</tr>
<tr>
<td>Fatal Incidents</td>
<td>648</td>
<td>512</td>
<td>340</td>
<td>322</td>
<td>210</td>
</tr>
<tr>
<td>Percent Fatal Incidents</td>
<td>9.93</td>
<td>10.28</td>
<td>9.74</td>
<td>10.44</td>
<td>10.96</td>
</tr>
</tbody>
</table>

* Source: NCHRP 755
Introduction

Crossings remain a significant hazard. Latest trends do not show accident decline (NCHRP 755)

Accidents at Grade Crossings (2003 – 2012)

**Nationwide Data**

**Illinois Data**
Introduction

Distribution of accidents in Illinois

Similar general patterns, but difficult to predict specific locations due to low frequencies
Introduction

Distribution of accidents at public crossings in Illinois (2003 – 2012)

- Accidents occurred at ~13% of crossings
  (similar nationwide trend ~12%)
- Highest accident frequency = 9
- Only 48 locations with 3 or more accidents
- Highly skewed distribution, as expected

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Locations in the analysis</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>5874</td>
<td>87.22%</td>
</tr>
<tr>
<td>1</td>
<td>685</td>
<td>10.17%</td>
</tr>
<tr>
<td>2</td>
<td>128</td>
<td>1.90%</td>
</tr>
<tr>
<td>3</td>
<td>31</td>
<td>0.46%</td>
</tr>
<tr>
<td>4</td>
<td>11</td>
<td>0.16%</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>0.06%</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>0.01%</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>0.01%</td>
</tr>
<tr>
<td>Total</td>
<td>6735</td>
<td>100.00%</td>
</tr>
</tbody>
</table>
Research Questions

- Can we find trends, contributing factors by analyzing individual high accident locations? (micro approach)

- If there are trends, can we incorporate them in a macro model?
  - In current models?
  - In new models?

- If we can incorporate them in a model, can we improve:
  - Accident predictions?
  - Ranking of locations?
A Micro Approach – Single Crossing

- Micro-level approach is being developed
- The main goal is to spot local trends/contributing factors
- Simple, easy-to-use, and complements info for site inspections and prelim analysis
- 3 steps so far:
  - Sketch of crossings with key info
  - Tree structure to spot trends
  - Additional information (e.g. surroundings, land use, nearby ramps)
A Micro Approach – Single Crossing

• Sketch of crossings with key info:

Nagle Ave and Avondale Ave, Near I-90, Chicago
A Micro Approach – Single Crossing

- **Variables observed:**

  - Highway User Type (typveh)
  - Action of Motorist (motorist)
  - Highway User Direction (vedir)
  - Time Table Direction (trndir)

- **K = Pedestrian**
  - A-J = Motorized vehicle
  - M = Other

- **Action of Motorist (motorist):**
  - 1 = Drove around or thru the gate
  - 2 = Stopped and then proceeded
  - 3 = Did not stop
  - 4 = Stopped on crossing
  - 5 = Other

- **Highway User Direction (vedir):**
  - 1 = north, 2 = south, 3 = east, 4 = west
  - (Depend on the actual Geometrical direction on the map)

- **Time Table Direction (trndir):**
  - 1 = north, 2 = south, 3 = east, 4 = west
  - (Depend on the actual Geometrical direction on the map)
A Micro Approach – Single Crossing

• Variables observed:

  - Circumstance of Accident (typacc)
    - 1 = rail equipment struck highway user
    - 2 = rail equipment struck by highway user

  - Vehicle Driver Age and Gender (drivage, drivgen)
    - (drivage): Numerical value of the age
    - (drivgen): 1 = male, 2 = female

  - Weather Condition and Visibility (weather, visiblty)
    - (weather):
      - 1 = clear, 2 = cloudy, 3 = rain,
      - 4 = fog, 5 = sleet, 6 = snow
    - (visiblty):
      - 1 = dawn, 2 = day,
      - 3 = dusk, 4 = dark
A Micro Approach – Single Crossing

173887/G
9 Accidents

Motorized
Vehicle

Drove around or
trough the gate

Vehicle
Northbound

Vehicle
Southbound

Other
(Specified in Text)

Pedestrian

Other
(Acton Unknown)

Pedestrian
Southbound

Train
Eastbound

Train
Westbound

Train struck
highway user

Train struck
by highway user

Train struck
highway user

Train struck
by highway user

Train struck
highway user

Train struck
by highway user
A Micro Approach – Single Crossing

A Second Example:

South Chicago
119 St and Wood St
(near I-57)

- 7 accidents in 10 years
- All accidents on main tracks
A Second Example:

No clear trends
A Third Example:

- Chestnut and Lehigh Ave, Glenview, IL

- 4 of 5 accidents involved drivers older than 80

- The remaining driver was 61 years old
A Micro Approach – Single Crossing

A Third Example (Cont…)

- High concentration of assisted living communities
- Could countermeasures incorporate population type?
- Potential age-related solutions (signs, marking)?
State of Practice

Let’s look at a common model to predict accidents at grade crossings…

- FRA Web Accident Prediction System (WBAPS)
- This system uses the DOT accident prediction formula (developed from 1975 database)
- Accident predictions are an input to resource allocation model
Presenter:

Dr. Juan C. Medina
Postdoctoral Research Associate
The U.S. DOT Model

- Initial collision model was developed based on the data from 1975 database
- Most coefficients remain unchanged since 1980
- An accident history term was added to reflect safety trends in recent years
- Weighted average of the initial and accident history terms is computed
- The computed value is multiplied by “normalizing constants” that vary by year
Three Main Steps in U.S. DOT Model

1. Initial collision prediction (a) = \( K \times EI \times MT \times DT \times HP \times MS \times HT \times HL \)

Example for Flashing Lights:

<table>
<thead>
<tr>
<th>Crossing Category</th>
<th>Formula Constant K</th>
<th>Exposure Index Factor EI</th>
<th>Main Tracks Factor MT</th>
<th>Day Thru Trains Factor DT</th>
<th>Highway Paved Factor HP</th>
<th>Maximum Speed Factor MS</th>
<th>Highway Type Factor HT</th>
<th>Highway Lanes Factor HL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flashing Lights</td>
<td>0.0083646</td>
<td>( e^{0.1681 t + 0.2 \frac{0.3683}{0.2}} )</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>( e^{0.1380(1-t)} )</td>
<td></td>
</tr>
</tbody>
</table>

Three Main Steps in U.S. DOT Model

2. Second collision prediction \((B) = \frac{T_0}{T_0 + T} (a) + \frac{T}{T_0 + T} \left( \frac{N}{T} \right)\)

\[ T_0 = \frac{1.0}{(0.05 + a)} \]

Accident History Term
(N accidents in T years)

3. Final collision prediction \((A) = B \times \text{Normalizing Constant}\)

Source: Accident Prediction and Resource Allocation Procedure Normalizing Constants. 2010
Other Alternatives… A Combined Model?

- Consider findings from micro analysis in building a macro model
- Used data from Illinois, explored different regression models:
  - Poisson, NB, **Zero-inflated NB**
- 5 years of data (2003-2007) used for model building
- 5 years for data (2008-2012) used for evaluation/validation of the model
- ZINB models created for the same warning devices used by U.S. DOT formula
A Combined Model

• ZINB Model Form:

\[
Accidents = \left(1 - \frac{1}{1 + e^{-(k_1 - (k_2 \cdot aadt))}}\right) \cdot e^{-k_3 + (k_4 \cdot aadt) + (k_5 \cdot \text{trains}) + (k_6 \cdot \text{lanes}) + f_{\text{angle}}}
\]

\[
f_{\text{angle}} = \begin{cases} 
  k_7 & \text{if } 0 < \text{angle} < 30 \\
  k_8 & \text{if } 30 \leq \text{angle} < 60 \\
  0 & \text{if } 60 \leq \text{angle} \leq 90 
\end{cases}
\]
A Combined Model

- Model for crossbucks based only on conflicting traffic (similar to US DOT model)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Category</th>
<th>Passive Warning Devices (e.g. crossbucks)</th>
<th>Selected for Final Model</th>
<th>Pvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-</td>
<td>✓</td>
<td>✓</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>aadt</td>
<td>-</td>
<td>✓</td>
<td>✓</td>
<td>0.0383</td>
</tr>
<tr>
<td>total_train</td>
<td>-</td>
<td>✓</td>
<td>✓</td>
<td>0.0098</td>
</tr>
<tr>
<td>cross angle</td>
<td>1 (0~29°)</td>
<td>x</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 (30~59°)</td>
<td>x</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 (60~90°)</td>
<td>x</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>total_tracks</td>
<td>-</td>
<td>x</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>traf_lanes</td>
<td>-</td>
<td>x</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>hwy_near</td>
<td>1 (0-200ft)</td>
<td>x</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 (&gt;200ft)</td>
<td>x</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Intercept (zero model)</td>
<td>-</td>
<td>✓</td>
<td>✓</td>
<td>0.35 *</td>
</tr>
<tr>
<td>total_train (zero model)</td>
<td>-</td>
<td>✓</td>
<td>✓</td>
<td>0.18 *</td>
</tr>
</tbody>
</table>

Passive Warning Devices (e.g. crossbucks)
A Combined Model

- Angle was a significant contributing factor for active warning devices

<table>
<thead>
<tr>
<th>Variable</th>
<th>Category</th>
<th>Active Warning Devices (e.g. flashing lights)</th>
<th>Selected for Final Model</th>
<th>Pvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-</td>
<td>facebook.com</td>
<td>✓</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>aadt</td>
<td>-</td>
<td>facebook.com</td>
<td>✓</td>
<td>0.0089</td>
</tr>
<tr>
<td>total_train</td>
<td>-</td>
<td>facebook.com</td>
<td>✓</td>
<td>0.0226</td>
</tr>
<tr>
<td>cross angle</td>
<td>1 (0°-29°)</td>
<td>facebook.com</td>
<td>✓</td>
<td>0.0746</td>
</tr>
<tr>
<td></td>
<td>2 (30°-59°)</td>
<td>facebook.com</td>
<td>✓</td>
<td>0.0667</td>
</tr>
<tr>
<td></td>
<td>3 (60°-90°)</td>
<td>facebook.com</td>
<td></td>
<td>N/A</td>
</tr>
<tr>
<td>total_tracks</td>
<td>-</td>
<td>facebook.com</td>
<td>x</td>
<td>N/A</td>
</tr>
<tr>
<td>traf_lanes</td>
<td>-</td>
<td>facebook.com</td>
<td>✓</td>
<td>0.0044</td>
</tr>
<tr>
<td>hwy_near</td>
<td>1 (0-200ft)</td>
<td>facebook.com</td>
<td>x</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>2 (&gt;200ft)</td>
<td>facebook.com</td>
<td></td>
<td>N/A</td>
</tr>
<tr>
<td>Intercept (zero model)</td>
<td>-</td>
<td>facebook.com</td>
<td>✓</td>
<td>0.0438</td>
</tr>
<tr>
<td>total_train (zero model)</td>
<td>-</td>
<td>facebook.com</td>
<td>✓</td>
<td>0.0275</td>
</tr>
</tbody>
</table>
A Combined Model

- Angle not important for gated crossings; distance to nearest highway intersection was significant

<table>
<thead>
<tr>
<th>Variable</th>
<th>Category</th>
<th>Gates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Selected for Final Model</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-</td>
<td>✓</td>
</tr>
<tr>
<td>aadt</td>
<td>-</td>
<td>✓</td>
</tr>
<tr>
<td>total_train</td>
<td>-</td>
<td>✓</td>
</tr>
<tr>
<td>cross angle</td>
<td>1 (0°-29°)</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>2 (30°-59°)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 (60°-90°)</td>
<td></td>
</tr>
<tr>
<td>total_tracks</td>
<td>-</td>
<td>✓</td>
</tr>
<tr>
<td>traf_lanes</td>
<td>-</td>
<td>x</td>
</tr>
<tr>
<td>hwy_near</td>
<td>1 (0-200ft)</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>2 (&gt;200ft)</td>
<td></td>
</tr>
<tr>
<td>Intercept (zero model)</td>
<td>-</td>
<td>✓</td>
</tr>
<tr>
<td>total_train (zero model)</td>
<td>-</td>
<td>✓</td>
</tr>
</tbody>
</table>

P-value

- Intercept: < 0.0001
- aadt: < 0.0001
- total_train: < 0.0001
- total_tracks: 0.0003
- hwy_near (1): 0.0283
- hwy_near (2): N/A
- Intercept (zero model): 0.2521*
- total_train (zero model): 0.2495*
Accuracy of Prediction

• Comparison is based on two factors:

1. Absolute and relative predictions
2. Ranking of high accident locations
Accuracy of Prediction

**Overall absolute predictions:**

- \((Σ\text{ Predicted})/(\text{Total observed})\)
- True prediction of all crossings together
- Useful to check magnitude of predictions
Accuracy of Combined Macro Model

Passive Warning Devices

Cumulative Proportion of Crossings

Crossings with Gates

Cumulative Proportion of Crossings
Accuracy of Prediction

Overall relative predictions:

- $(\Sigma \text{ Predicted})/(\text{Total Predicted})$
- Useful to check if prediction curve is similar to data
Accuracy of Combined Macro Model

![Graphs showing cumulative proportion of crossings for Passive Warning Devices and Crossings with Gates. The graphs compare Field Data, US DOT Model, and ZINB Model predictions.](image-url)
# Ranking of high accident locations

<table>
<thead>
<tr>
<th>Warning Device</th>
<th>Ranking Method</th>
<th>Number of crashes predicted in top locations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Top 10</td>
</tr>
<tr>
<td>Passive (Crossbucks)</td>
<td>Data (observed)</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>US DOT Formula</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>ZINB - Average model and data</td>
<td>8</td>
</tr>
<tr>
<td>Active (Flashing Lights)</td>
<td>Data (observed)</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>US DOT Formula</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>ZINB - Average model and data</td>
<td>11</td>
</tr>
<tr>
<td>Gates</td>
<td>Data (observed)</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>US DOT Formula</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>ZINB - Average model and data</td>
<td>20</td>
</tr>
</tbody>
</table>
Summary

- Improvements for state-of-practice in accident prediction are needed:
  - Accuracy of predictions, understanding contributing factors, finding countermeasures
- Micro approach finds contributing factors that data aggregation may mask
- Combination of macro and micro analysis improved accident prediction
  - Better accuracy
  - Better ranking
Future (and Ongoing) Research

- Additional accident model forms, selection of best option
- Dynamic tree structure to automatically sort attributes and discover trends
  - Cluster crossings based on potential trends (corridors)
- Corridor Analysis
- Approaches for combining macro and micro

… More
A Micro Approach – Extension to Corridors?

• Can we extend the micro approach to corridors?
• Accidents along corridor can be grouped to identity possible trends
• GIS to locate crossings, add data with socio-economic and geographic info
Future (and Ongoing) Research

Example Corridor

- Northeast Illinois Regional Commuter Railroad

25 accidents at 8 crossings between 2003 and 2012
Future (and Ongoing) Research

Combining Micro and Macro

- One idea is to simply include variables from micro to macro models

- We could also add categories to the macro model (dummy variables, indicators, different models) and apply factors to modify predictions of some crossings

- Micro could also lead to different classification of crossings based on risk assessments
Future (and Ongoing) Research

...Continued

• Temporal analysis on the occurrence of accidents

• Verification with other datasets:
  • So far used Illinois data. How to generalize it
Questions?

Rahim F. Benekohal
Juan C. Medina

(rbenekoh@Illinois.edu)
(jcmedina@Illinois.edu)