Combination of Microscopic and Macroscopic Models for Analysis of Accidents at Railroad Grade Crossings

RAHIM F. BENEKOHAL
JUAN C. MEDINA

GLXS 2014 – University of Illinois at Urbana-Champaign
Background

- Grade crossings accidents are often predicted based on models developed on national/state level trends (macro model)

- Current state-of-practice is using the U.S. DOT accident model:
  
  (FRA’s Web Accident Prediction System – WBAPS)

- Useful tool that may provide acceptable results

- Are WBAPS estimates accurate enough? Can we do better?
Background

- Two modeling options:
  - Micro scale: Detail analysis of contributing factors at high accident crossings
  - Macro scale: Models built on national/state level trends and contributing factors

- Each option has usefulness and limitations
- Goal is to combine results from micro and macro analysis
- Illinois data is used to illustrate potential benefits of combined method
A Micro Approach

- Micro-level approach is being developed

- The main goal is to spot local trends/contributing factors

- Simple, easy-to-use, complements info for site inspections and prelim analysis

- 3 steps so far:
  - Sketch of crossings with key info
  - Tree structure to spot trends (dynamic structure is under development)
  - Additional information (e.g. surroundings, land use, nearby ramps)
A Micro Approach

- Sketch of crossings with key info:
A Micro Approach

- Tree structure to visualize trends: A first example

Angle Issue?
A Micro Approach

A Second example:

Unusual trend spotted at the bottom of the tree:

- 4 of 5 accidents involved drivers older than 80
- The remaining driver was 61 years old
A Micro Approach

A Second example (Cont...)

- Crossing surroundings indicated high concentration of assisted living communities
- Countermeasures should incorporate population type
- Should aspects like this be included at macro level?
A Combined Macro Model

- Combination of micro results into improved macro models currently under study

- Variables such as crossing angle can be tested and added to improve accuracy

- Example to compare U.S. DOT model and a combined model based on:
  - The overall accident predictions (absolute predictions)
  - Ranking high-accident locations
The U.S. DOT Model

- Initial collision model was developed based on the data from 1975 database

- Most coefficients remain unchanged since 1980
  - (“Rail-Highway Crossing Hazard Prediction – Research Results”)

- A new term is added to reflect frequency of accidents in recent years

- Weighted average of the initial and new terms is computed

- To compute the predicted No. of accidents: the predicted value is multiplied by a constant that changes by the device type and 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>0.003646</td>
<td>(\exp{-0.1553at})</td>
<td>0.2</td>
<td>(d + 0.2 \times 0.0470)</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>(\exp{0.190015-1})</td>
</tr>
</tbody>
</table>

Three Main Steps in U.S. DOT Model

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

Accident history (N accidents in T years)

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

3. Final collision prediction (A) = B * Normalizing Constant

Source: Accident Prediction and Resource Allocation Procedure Normalizing Constants. 2010
Observations on the U.S. DOT Model

- Does not provide relevant information about local contributing factors

- Final prediction is mostly based on past history:

\[
\frac{T_0}{T_0 + T}(a) + \frac{T}{T_0 + T}\left(\frac{N}{T}\right)
\]

For a 5-year analysis: \( T_0 \geq T \) if \( a < 0.15 \)

For most high accident locations \( a > 0.15 \)

Example: AADT=3000, Trains=50, 3 lanes, 3 tracks \( a=0.22 \)

- In summary, prediction similar to past history for high accident locations, limited role of \( a \)
A Combined Macro Model

- Using data from Illinois, we tested different regression models, found good fit for:
  - Zero-inflated negative binomial (ZINB)
- 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 device categories as U.S. DOT formula
A Combined Macro Model

- Significant contributing variables (ZINB Models) – example active warning devices:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Category</th>
<th>Active Warning Devices</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>0.0089</td>
<td></td>
</tr>
<tr>
<td>total_train</td>
<td>-</td>
<td>✓</td>
<td>0.0226</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>✓</td>
<td>0.0044</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></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept (zero model)</td>
<td>-</td>
<td>✓</td>
<td>0.0438</td>
<td></td>
</tr>
<tr>
<td>total_train (zero model)</td>
<td>-</td>
<td>✓</td>
<td>0.0275</td>
<td></td>
</tr>
</tbody>
</table>
Accuracy of Combined Macro Model

Overall absolute predictions:

- ZINB Model

- Adjustment based on accident history (similar to $B$ in U.S. DOT model)
Accuracy of Combined Macro Model

- 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>Active</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>
</tbody>
</table>
Conclusions

- 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:
  ◦ We got positive feedback from practitioners
  ◦ Future developments are promising

- Combination of macro and micro analysis showed accident prediction can be improved
Conclusions

- US DOT Model trends from case study:
  ◦ Initial underestimation of accidents (a), then almost complete reliance of history (B)
  ◦ Significant overestimation of absolute predictions

- Macro regression models using recent data provided improvements:
  ◦ Better accuracy, more reliable results (based on current data)
  ◦ Basis for justify rankings, absolute predictions are sound

- Future activities:
  ◦ Improve the accident predictions based on the combination approach
  ◦ Develop probabilistic analysis, and dynamic tree
Thanks!

Questions?

Rahim F. Benekohal (rbenekoh@Illinois.edu)
Juan C. Medina (jcmedina@Illinois.edu)