Using naturalistic driving data to model crash risk at a city level

Marta Faria¹, Patrícia Baptista¹, Tiago Farias¹, Sara Ferreira²

¹Instituto Superior Técnico, Universidade de Lisboa
²FEUP, Universidade do Porto

ana.marta.faria@tecnico.ulisboa.pt
October 26th, 2018

Outline

1. Motivation
2. Objectives
3. Methodology
4. Results
5. Conclusions
6. Acknowledgements

Motivation

• Transportation sector:
  • Major role in economy and society
  • Impacts people quality of life, by making places accessible and bringing people together
  • Depends heavily on oil resources
  • Negative impacts regarding air pollution and global warming
  • 1.2 million people die in road traffic crashes every year

• Transportation sector challenges:
  • Energy consumption
  • Pollutants emissions
  • Road safety

Motivation

• Transportation sector:
  • Major role in economy and society
  • Impacts people quality of life, by making places accessible and bringing people together
  • Depends heavily on oil resources
  • Negative impacts regarding air pollution and global warming
  • 1.2 million people die in road traffic crashes every year

• Transportation sector challenges:
  • Energy consumption
  • Pollutants emissions
  • Road safety
Motivation

Road safety:
- 25,700 road fatalities in the EU in 2014
- EU set an ambitious target to halve road deaths by 2020
- Human related factors are the main cause of road accidents

Contributing factors to vehicle crashes

Source: HSM, 2009
Source: European Commission, 2016

How to address the transportation sector challenges regarding energy consumption, emissions and safety performance?

- Alternative vehicle technologies (BEV, hybrids, etc.)
- Diversification of energy pathways (biofuel, electricity, hydrogen, etc.)
- Vehicle safety and the safety of infrastructure
- Regulatory interventions
- Promote driving behavior changes (education, training, feedback, reinforcement, etc.)

Objectives:
- characterize driving behavior at a street level;
- explore the association between driving behavior and crash occurrence.

Methodology

DATA SOURCES
- Municipalities
- Onboard data acquisition EID device
- Google Maps (Geocoding API)
- ANN

DATABASE DEVELOPMENT
- Hierarchical street level
- Coordinates (longitude, latitude), acceleration, slope, VSP, etc.
- Street and city
- Accidents data

DATA ANALYSIS
- SAS® Enterprise Guide®-L3 software
- VSP methodology
- SPSS software
- Binary logistic regression

OUTPUT
- Driving behavior characterization at a street level
- Association between driving behavior and crash occurrence
Methodology

On-board data acquisition – i2D device
- On-board monitor assistant (i2d – intelligence to drive)
- Non-intrusive system with specific dedicated sensors such as GPS and OBD
- Monitors the main variables that characterize vehicle usage (location, vehicle dynamics, road grade and engine management)
- Calculates accurate relevant indicators to characterize the impact of vehicle use

Sample
- 47 drivers monitored for a period of at least six months between 2014 and 2015
- The drivers used their own vehicles:
  - 61% diesel compliant with Euro 4 or newer
  - From small city cars to sport utility vehicles (SUVs) with the family cars being the most common (40% of the vehicles)
- More than 8,000h of driving data, of which 3,500h in the city of Lisbon

<table>
<thead>
<tr>
<th>Gender</th>
<th>Age (years)</th>
<th>Driving experience (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>26</td>
<td>12.1</td>
</tr>
<tr>
<td>Female</td>
<td>20</td>
<td>13.3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Euro standard</th>
<th>Number of vehicles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Euro 3</td>
<td>10</td>
</tr>
<tr>
<td>Euro 4</td>
<td>20</td>
</tr>
<tr>
<td>Euro 5</td>
<td>19</td>
</tr>
<tr>
<td>Euro 6</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fuel type</th>
<th>Number of vehicles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diesel</td>
<td>29</td>
</tr>
<tr>
<td>Gasoline</td>
<td>16</td>
</tr>
<tr>
<td>Gasoline/LPG</td>
<td>1</td>
</tr>
</tbody>
</table>
Methodology

Data source

- Google Maps Geocoding API
  - To associate each pair of GPS coordinates on a second-by-second basis to a street name
- Lisbon municipality
  - Hierarchical street levels that define street function:
    - Level 1 – arterial streets;
    - Level 2 – minor arterial streets;
    - Level 3 – distributor and collector streets;
    - Level 4 – local streets.
- National Authority for Road Safety (ANSR)
  - Crash data (2010-2015): crash type, injury severity, number of deaths and injuries, crash location (street, municipality and district), etc.

Methodology

Street level data analysis

- Driving behavior characterization at a street basis considering variables such as:
  - Average speed
  - Average acceleration
  - VSP distribution
- Around 50% of the Lisbon streets were characterized with both driving behavior and crash data
- Streets with less than 60 sec. of driving data were not included in the analysis

VSP methodology (Vehicle Specific Power)

- Fuel consumption and emissions obtained from measured data are grouped into bins, according to the power requirements per unit of mass of the vehicle, on a second by second basis
- Specific power can be directly measured and the emissions quantification can be done independently of the physical parameters of the vehicle (vehicle type, weight, engine size, etc.)

\[
\text{VSP} = \frac{\text{Power}}{\text{Mall}}
\]

\[
\text{Power} = \frac{\text{ æ } \times \text{ æ } \times \text{æ} \times \text{æ}}{\text{ æ } \times \text{æ } \times \text{æ} \times \text{æ} \times \text{æ} \times \text{æ}}
\]
Methodology

VSP methodology (Vehicle Specific Power)

- Divided in 14 modes (data with similar W/kg is grouped):
  - Modes 1 and 2: deceleration or negative road slopes
  - Mode 3: Idling
  - Modes 4 to 14: the higher the speed and acceleration (more power demand), the higher the mode.

Methodology

Binary Logistic regression

- Driving behavior data on a street basis
- Combining driving data from several drivers and hours
- Focusing on variables that are most likely associated with crashes
  - Speed, positive and negative acceleration, % of time per VSP mode, ...
- Dependent variable
  - Binary variable representing the occurrence of accidents: 1 – Yes, 0 – No
- Based on police recorded crash data for a six year period (2010-2015)
- Level 1 streets excluded from the analysis
- No streets without accidents

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level</td>
<td>Hierarchical street level (from 2 to 4)</td>
</tr>
<tr>
<td>length</td>
<td>Total length of the street in km</td>
</tr>
<tr>
<td>var_speed</td>
<td>Speed variance</td>
</tr>
<tr>
<td>max_speed</td>
<td>Maximum speed (average of all speed values above the 95th percentile)</td>
</tr>
<tr>
<td>avg_pos</td>
<td>Average positive acceleration (m/s²)</td>
</tr>
<tr>
<td>avg_neg</td>
<td>Average negative acceleration (m/s²)</td>
</tr>
<tr>
<td>max_pos</td>
<td>Maximum acceleration (average of all positive acceleration values above the 95th percentile)</td>
</tr>
<tr>
<td>max_neg</td>
<td>Maximum deceleration (average of all negative acceleration values below the 5th percentile)</td>
</tr>
<tr>
<td>perc_A3</td>
<td>Percentage of time spent in acceleration above 3 m/s² (%)</td>
</tr>
<tr>
<td>perc_D3</td>
<td>Percentage of time spent in deceleration below -3 m/s² (%)</td>
</tr>
<tr>
<td>perc_A4</td>
<td>Percentage of time spent in acceleration above 2 m/s² (%)</td>
</tr>
<tr>
<td>perc_D4</td>
<td>Percentage of time spent in deceleration below -2 m/s² (%)</td>
</tr>
</tbody>
</table>
Results

Average speed

- Level 4 streets: average speed 60% lower than in level 1 streets
- Level 2, 3 and 4 streets:
  - same posted speed limit (50 km/h)
  - variations of -15% and -10% in speed
- More local streets: typically lower number of lanes

Driving behavior characterization

- Average positive and negative accelerations increase by more than 30%
- Similar values for level 2, 3 and 4 streets
- Maximum and minimum accelerations typically occur in level 2 streets
Results

Driving behavior characterization

- Level 1 streets → almost 50% of the time on speeds above 50 km/h
- Level 3 and 4 streets → more than 90% of the time on speeds below 50 km/h
- Level 2 streets → ~27% of the time idling
- Excess speeding → level 1 – 13%; level 2 – 17%; level 3 – 8%; level 4 – 5%

Variations in terms of power demand
- Level 2, 3 and 4 streets: – Lower power requirements – Higher % of time spent in VSP modes 2 and 3 – Influenced by intersections and interactions with pedestrians

Crash risk assessment

- Length – Exposure variable: longer streets → more likely to have accidents
- Level → higher crash risk in level 2 and 3 than in level 4 streets
- Higher % time idling, VSP1, VSP4 may represent congestion situations → higher crash risk
- Higher % of time in excess speeding → less likely to have accidents

Preliminary results

Conclusions

Driving behavior characterization

- Driving in streets of different levels leads to significant differences in driving behavior:
  - Average speed → decreasing trend from level 1 to level 4 streets
  - Average speed → 60% lower in level 4 streets than in level 1
  - Average positive and negative accelerations → increase by more than 30% for more local streets (from level 1 to level 4)

Crash risk assessment

- Higher percentage of time spent idling, in VSP mode 1 and VSP mode 4
  - More likely to have accidents
Future work

- Develop an equivalent model for weekends → compare results for weekdays and weekends
- Split weekdays data in peak and off-peak hours
- Some issues to solve:
  - Most accidents occur in Level 1 and 2 streets
    - Level 1 streets: had to be excluded from the analysis since all streets had accidents (no zeros)
    - Level 2 streets: few streets without accidents
  - Analysis on a street basis instead of section of road:
    - Few streets without accidents
    - Most safety related databases have high percentage of zeros (no accidents)

How to overcome these issues?

Suggestions?

Acknowledgements

- The authors acknowledge Fundação para a Ciência e Tecnologia (FCT) for the Doctoral (PD/B0/105714/2014) financial support, and the project grant (SusCity Project, MITP-TB/C 5/0028/2013). This work was also supported by FCT, through IDMEC, under LAETA, project ID/EMS/50022/2013 and through IN+, Strategic Project UID/EEA/50009/2013.
- The authors would also like to acknowledge Autoridade Nacional de Segurança Rodoviária (ANSR) for providing the accidents data.
- Acknowledgements are also due to:

Thank you for your attention!