Validation of Simulation Models Using Vehicle Trajectories

TRB Annual Meeting
January 11, 2015
Overview

- Project Objectives and the Project Team
- State of Practice for Aggregate Calibration
- Trajectory Datasets and Data Collection Methods
- Insights from Mining Vehicle Trajectories
- Disaggregate Trajectory Validation
- Aggregate Trajectory Validation
- Trajectory Validation Computational Engines
Current approaches to develop, calibrate, and validate simulation tools are based on use of aggregate-level field data.

Microsimulation models, while simulating the detailed position (trajectories) of vehicles on a subsecond level, are for the most part not validated at that level.

Much more realistic analysis tools can be developed if validated against vehicle trajectory data over a variety of operational conditions.
• Compile existing and collect new vehicle trajectory datasets

• Develop a methodology for trajectory-level AMS tool validation

• Develop a computational engine that allows observed and simulated trajectories to be analyzed, visualized, and compared to each other at the trajectory or aggregate levels

• Demonstrate the validation process using a proof of concept application
Project Deliverables

- Build methodology to validate trajectories at the aggregate or disaggregate level
- Collect new data and/or enhance/clean/analyze existing trajectory data sources
- Develop reasonable ranges for a number of performance measures based on observed trajectory data
- Build a computational engine that reads observed or simulated trajectories, performs tracing tests, allows the user to analyze the data, and reports measures
Project Team

• FHWA: James Colyar, John Halkias, Jim Sturrock, and Paul Heishmann

• Cambridge Systematics: Michalis Xyntarakis, Vassili Alexiadis, Erin Flanigan, Robert Campbell

• Partners: Dr. Vincenzo Punzo, Dr. Lilly Elefteriadou, Dr. Alex Skabardonis, Dr. Martin Treiber, Angshuman Guin

Stakeholders: Jordi Casas, Michael Mahut, Dan Morgan, Xuesong Zhou, Karl Wunderlich, Jorge Laval, Li Zhang, Rama Balakrishna, Keir Opie, Kaan Ozbay
State of Practice for Aggregate Model Calibration
Model Setup And Calibration

1. Summarize Model Calibration Criteria
2. Develop Baseline Model Network including relevant transportation facilities and modes
3. Conduct Demand Modeling for Existing Baseline Year
4. Calibrate Simulation Model

Approximate LOE 30-40%

Supporting Substeps:
## Example Guideline Calibration Criteria (Recurrent Congestion)

<table>
<thead>
<tr>
<th>Calibration Criteria and Measures</th>
<th>Calibration Acceptance Targets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traffic flows within 15% of observed volumes for links with peak-period volumes greater than 2,000 vph</td>
<td>For 85% of cases for links with peak-period volumes greater than 2,000 vph</td>
</tr>
<tr>
<td>Sum of all link flows</td>
<td>Within 5% of sum of all link counts</td>
</tr>
<tr>
<td>Travel times within 15%</td>
<td>&gt;85% of cases</td>
</tr>
<tr>
<td>Visual Audits</td>
<td>To analyst’s satisfaction</td>
</tr>
<tr>
<td>Individual Link Speeds: Visually Acceptable Speed-Flow Relationship</td>
<td></td>
</tr>
<tr>
<td>Visual Audits</td>
<td>To analyst’s satisfaction</td>
</tr>
<tr>
<td>Bottlenecks: Visually Acceptable Queuing</td>
<td></td>
</tr>
</tbody>
</table>
Example Observed vs Modeled Volumes
Example Speed Contour Diagram

- Bottleneck Start Time and Location
- Bottleneck Queue Length
- Bottleneck Duration
- Bottleneck End Time and Location
Example Bottleneck Model Calibration

Observed and simulated speeds over different times of the day and miles per point (MP) from 3:00 PM to 7:00 PM.
Example Calibration Criteria for Nonrecurrent Congestion

- Freeway bottleneck locations. Should be on a modeled segment consistent in location, design, and attributes of the representative roadway section
- Duration of incident-related congestion. Duration where observable within 25 percent
- Extent of queue propagation. Should be within 20 percent
- Diversion flows. Increase in ramp volumes where diversion is expected to take place
- Arterial breakdown when incident. Cycle failures or lack of cycle failures
Available Data and Collection Methodology
### Most Relevant Trajectory Datasets – U.S.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Duration (Coverage)</th>
<th>Resolution (Hz)</th>
<th>Sample Size (vehicles)</th>
<th>Segment Length (ft)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NGSIM US101</strong></td>
<td>45 min</td>
<td>10</td>
<td>6,101</td>
<td>2,100</td>
</tr>
<tr>
<td><strong>NGSIM US-I80</strong></td>
<td>45 min</td>
<td>10</td>
<td>5,648</td>
<td>1,600</td>
</tr>
<tr>
<td><strong>NGSIM Lankershim</strong></td>
<td>30 min</td>
<td>10</td>
<td>2,450</td>
<td>1,600</td>
</tr>
<tr>
<td><strong>NGSIM Peachtree</strong></td>
<td>30 min</td>
<td>10</td>
<td>2,337</td>
<td>2,100</td>
</tr>
<tr>
<td>Naturalistic Driving Study (NY, PA, NC, IN, FL, WA)</td>
<td>3,700 veh-years</td>
<td>1</td>
<td>2,600 vehs</td>
<td>Entire trips from origin to destination</td>
</tr>
<tr>
<td><strong>Data Collected in this project</strong></td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>
Video detection holds high promise and is used with increasing success to detect vehicle movement.

GPS is ubiquitous but may lack accuracy.

WAAS GPS is becoming cheaper and cheaper and allows for lane inference through map-matching.

Inertial Measurement Units (IMU’s).

On-Board-Diagnostics (OBD).

LIDAR can collect positions of all surrounding vehicles.

Radar detectors are routinely used to measure gaps.
Insights from Trajectories
Speed Vs Acceleration

NGSIM Speed vs Acceleration For Each Vehicle and Frame (0.1sec)
Reconstucted Data by M. Montanino, and V. Punzo (8/0 4:00 to 4:15PM)

U.S. Department of Transportation
Federal Highway Administration
Number of Lane Changes Per Hour/Lane/Mile (Simulation)

A = 300 ft on-ramp
B = 300 ft 4 feet freeway segment immediately downstream a lane drop
Flow Per Hour Per Lane (Simulation)

A = 300 ft on-ramp
B = 300 ft 4 feet freeway segment immediately downstream a lane drop
Disaggregate Trajectory Validation
One-to-One Trajectory Comparison

- Similarity measure: sum of Euclidian distances between all corresponding pairs of point locations in the two trajectories
- Possible measures: use logarithms, Dynamic Time Warping, Multidimensional EDIT distance
- Calibrated trajectories differ no less than 20% from observed depending on car-following model used
Methodology

• In a simulator let a probe vehicle move freely but move all surrounding vehicles according to observed trajectories to record deviations between the probe’s observed and simulated trajectory.

• One-to-one trajectory validation can lead to conclusions only when observed and simulated trajectories belong to the same driver type.

• Extended Floating Car Data data can be used to validate car-following models but NGSIM type of data are required for lane-changing.

Figure 1: Comparison of simulated and empirical trajectories. The model parameters are calibrated according to Table 1 for the mixed error measure (10).
Aggregate Trajectory Validation
Aggregate Trajectory Validation

- How do we statistically compare point, one-dimensional, or two-dimensional aggregate trajectory measures?

- Two dimensional comparison measure:

\[
RMSNE = \sqrt{\frac{1}{N} \sum_{n=1}^{N} \left( \frac{m_{n}^{\text{sim}} - m_{n}^{\text{obs}}}{m_{n}^{\text{obs}}} \right)^{2}}
\]
Possible Validation Measures

- Car-following validation
  - Speed versus max acceleration/deceleration
  - Speed versus gap
- Lane-changing validation
  - Number of lane changes per mile
  - Time between successive lane changes
  - Gap distribution before a lane change
  - How far ahead vehicles make mandatory lane changes?
  - Time-space diagram of lane-changes
Trajectory Computational Engines and Visualization Tools
• Trajectories combine spatial, temporal, and relational data and are harder to query than tabular-only data

• Transportation
  – VTAPE (Dr. Scott Washburn, University of Florida)
  – Trajectory Explorer (Dr. Jorge Laval, GaTech)
  – SHRP L04 (Dr. Xuesong Zhou, Arizona State)
  – SHRP 2; Urban Reliability Analysis with GPS data (Nie)
  – Computer Science (Moving Objects Databases)
  – PostGIS (PostgreSQL with spatial extensions)
Conclusions

- Top-down aggregate calibration and validation methodologies can result in over fitting.
- Bottom-up calibration and validation of car-following and lane changing models under various conditions can enhance current calibration/validation methods.
- A methodology for validating simulated trajectories against observed ones at the aggregate and disaggregate level will be developed.
- Trajectory data will be collected and existing trajectory databases will be used to derive statistical/behavioral reasonableness checks.
Thank You