Transport Demand Modeling with Limited Data: Transfer and Calibration of a Model from another Geographical Region

Dominik Ziemke, Kai Nagel, Chandra Bhat
hEART Conference 2014 | 11 September 2014
Motivation

• Goal: Create Transport Demand Representation for a given region
  • Activity-based
  • Policy-sensitive
• Problem: Data Availability
  • Trip diaries with geo-coding increasingly difficult to procure because of privacy concerns (DE, US, …)
  • No information about trip destinations
• Approach: Generate locations in activity plans somewhat randomly and remove plans/locations not consistent with measurements.
  1. For home-work-home plans
  2. For full daily plans.
MATSIM
Multi-Agent Transport Simulation (MATSim)

• Agent-based
• Activity-based
• Iterative Demand-Supply Adaptation
• Dynamic Traffic Assignment
• Microscopic Demand
  • Plan Scoring
    \[ V(i) = \sum_{act \in m} V_{perf,m} + \sum_{trav \in n} V_{trav,n} \]
  • Replanning
    (Prob. Selection or Mutation)
    \[ P(i) = \frac{e^{V(i)}}{\sum_j e^{V(j)}} \]
CADYTS
Calibration of Dynamic Traffic Simulations (Cadyts)

- Developed by Flötteröd [2010]
- Can interact with any stochastic, dynamic, and iterative transport simulation framework, e.g. MATSim
- Calibrates in a Bayesian setting
- Treats simulation as a black box
- Correction term: \( \Delta V_a(k) = \frac{y_a(k) - q_a(k)}{\sigma^2_a(k)} \)
  \[ V(i) = \sum_{act \in m} V_{\text{perf},m} + \sum_{trav \in n} V_{\text{trav},n} \]
  \[ + w \cdot \sum_{ak \in i} \Delta V_a(k) \]
- Performance evaluation in terms of behavior and real-world measurement reproduction at the same time
HOME-WORK-HOME PLANS
Home-Work-Home Plans

Idea

• Create several HWH plans per agent, each with another work location
• Use Cadyts to sort out those HWH plans which are not consistent with measurements (traffic counts)

Setup

• Population with home and work locations based on Commuter Statistics
• Only people going by car (car share by region)
• Work activity starts at 8:00 +/- 1.5h
• Work activity ends at 16:30 +/- 1.5h
• 7+1 plans per agent
### Settings and Results

<table>
<thead>
<tr>
<th>Parameter</th>
<th>W/o Cadyts</th>
<th>W/ Cadyts</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demand Elasticity</td>
<td>Yes</td>
<td>Yes</td>
<td>n/a</td>
</tr>
<tr>
<td>Number of Plans</td>
<td>10</td>
<td>10</td>
<td>n/a</td>
</tr>
<tr>
<td>Number of initial Plans</td>
<td>8</td>
<td>8</td>
<td>n/a</td>
</tr>
<tr>
<td>Flow Capacity Factor</td>
<td>0.02</td>
<td>0.02</td>
<td>n/a</td>
</tr>
<tr>
<td>Cadyts Scoring Weight</td>
<td>n/a</td>
<td>15</td>
<td>n/a</td>
</tr>
<tr>
<td>Calibration Time</td>
<td>n/a</td>
<td>0h - 24h</td>
<td>n/a</td>
</tr>
<tr>
<td>Norm. Log Likelihood</td>
<td>n/a</td>
<td>-213</td>
<td>-10*</td>
</tr>
<tr>
<td>Car Trips</td>
<td>1.35m</td>
<td>1.29m</td>
<td>n/a</td>
</tr>
<tr>
<td>Car Trips / Person</td>
<td>2.0</td>
<td>2.0</td>
<td>n/a</td>
</tr>
<tr>
<td>Avg. Trip Distance [km]</td>
<td><strong>12.7</strong></td>
<td><strong>14.9</strong></td>
<td><strong>14.7</strong></td>
</tr>
<tr>
<td>Avg. Trip Duration [km]</td>
<td><strong>25.8</strong></td>
<td><strong>34.4</strong></td>
<td><strong>31.1</strong></td>
</tr>
</tbody>
</table>

* Flötteröd et al. 2009, p.10
** Own calculations based on SrV Berlin 2008
HWH Plans without Cadyts vs. with Cadyts

Without Cadyts

With Cadyts

• HWH Plans (alone) generate morning/afternoon peak traffic, but is far away from counts
• HWH Plans + Cadyts generate morning/afternoon traffic which is much closer to counts
CEMDAP
CEMDAP

• Comprehensive Econometric Microsimulator for Daily Activity-Travel Patterns
• C. Bhat et al., University of Texas

Input

• Disaggregate Demographics (Preprocess: SPG)
• Model Specification (taken from Dallas/Fort Worth, TX)

Output

• Daily Activity-Travel Patterns for each Indiv.
Used Data

Input
• Commuter Statistics by Federal Employment Agency
• Basic Demographics
• Settlement Patterns
• Traffic Counts

Validation
• SrV 2008 Scientific Use File for Berlin

Premise
• “Do not use data that are not readily accessible.”
Synthetic Population

• Problem: „Big Municipalities“

Properties

• Home-Work Relations
• Settlement Patterns
• Car Share
• Age
• Gender
• Employment Status
• Retired people
• Students

LOR („Lebensweltlich orientierte Räume“)
FULL DAILY PLANS
CEMDAP Output -> MATSim Input

CEMDAP
• yields daily activity-travel patterns of individuals

Are we done yet?
• No context-specific estimation of model parameters
• No interaction of supply and demand („feedback“)

MATSim
• expects daily plans as input
• ≈ CEMDAP output
HWH Plans vs. CEMDAP Full Daily Plans (w/o Cadyts)

HWH Plans (as earlier)

CEMDAP Plans

- CEMDAP plans generate travel for whole day where HWH plans only reproduce morning and afternoon peaks
FINAL SETUP
CEMDAP -> MATSIM+CADYTS
Full Daily Plans + Cadyts

Idea

• Create several full daily plans per agent with CEMDAP, each with another work location
• Use Cadyts to sort out those full daily plans which are not consistent with measurements (traffic counts)

Setup

• As before (CEMDAP)
• 7+1 plans per agent
Summary of Procedure

• Run CEMDAP multiple times
• Get multiple activity patterns for each agent with different locations
• Consider output plans as Initial Suggestions for a Demand Representation
• MATSim evaluates the performance of these suggestions in terms of
  • Behavioral soundness
  • Relation to reality (via Cadyts)
## Settings and Results

<table>
<thead>
<tr>
<th>Parameter</th>
<th>W/o Cadyts</th>
<th>W/ Cadyts</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demand Elasticity</td>
<td>Yes</td>
<td>Yes</td>
<td>n/a</td>
</tr>
<tr>
<td>Number of Plans</td>
<td>10</td>
<td>10</td>
<td>n/a</td>
</tr>
<tr>
<td>Number of initial Plans</td>
<td>4</td>
<td>4</td>
<td>n/a</td>
</tr>
<tr>
<td>Flow Capacity Factor</td>
<td>0.02</td>
<td>0.02</td>
<td>n/a</td>
</tr>
<tr>
<td>Cadyts Scoring Weight</td>
<td>n/a</td>
<td>15</td>
<td>n/a</td>
</tr>
<tr>
<td>Calibration Time</td>
<td>n/a</td>
<td>0h - 24h</td>
<td>n/a</td>
</tr>
<tr>
<td>Norm. Log Likelihood</td>
<td>-219</td>
<td>-23</td>
<td>-10*</td>
</tr>
<tr>
<td>Car Trips</td>
<td>3.89m</td>
<td>2.92m</td>
<td>3.20m**</td>
</tr>
<tr>
<td>Car Trips / Person</td>
<td>3.9</td>
<td>3.4</td>
<td>3.4**</td>
</tr>
<tr>
<td>Avg. Trip Distance [km]</td>
<td>12.0</td>
<td>11.0</td>
<td>9.5**</td>
</tr>
<tr>
<td>Avg. Trip Duration [km]</td>
<td>27.0</td>
<td>22.0</td>
<td>22.3**</td>
</tr>
</tbody>
</table>

* Flötteröd et al. 2009, p.10
** Ziemke 2013, pp.59, based on SrV Berlin 2008
CEMDAP Full Daily Plans (without Cadyts vs. with Cadyts)

Without Cadyts (as earlier)

With Cadyts

- CEMDAP full daily plans calibrated by Cadyts generate travel for whole day
VALIDATION
Departure Times

Without Cadyts

With Cadyts
Trip Distance (Beeline)

Without Cadyts

Average = 12.0km

Reference = 9.5km

With Cadyts

Average = 11.0km
Trip Duration

Without Cadyts

With Cadyts

Average = 27.0min

Average = 22.0min

Reference = 22.3min
Activity Type at Trip End

With Cadyts

<table>
<thead>
<tr>
<th>Activity Type</th>
<th>Simulation</th>
<th>Survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Home</td>
<td>0.42</td>
<td>0.35</td>
</tr>
<tr>
<td>Leisure</td>
<td>0.16</td>
<td>0.15</td>
</tr>
<tr>
<td>Other</td>
<td>0.20</td>
<td>0.20</td>
</tr>
<tr>
<td>Shop</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>Work</td>
<td>0.20</td>
<td>0.20</td>
</tr>
</tbody>
</table>
Conclusion

• Generate a transport demand representation for Berlin
  • Microscopic – individuals with full daily plans
  • Based on readily accessible input data
  • A activity-travel plan generator estimated for another region (CEMDAP)
  • Updating procedure based on traffic counts (Cadyts)
• Good fit towards measurements
• Good validation results

• Individual travelers maintained over whole process

• Contribution to dealing with data without geo-coding
• Contribution to transferability of models