Development of the Idaho Statewide Travel Demand Model Trip Matrices Using Cell Phone OD Data and Origin Destination Matrix Estimation

Ben Stabler
*Parsons Brinckerhoff*

Follow this and additional works at: [https://pdxscholar.library.pdx.edu/trec_seminar](https://pdxscholar.library.pdx.edu/trec_seminar)

Let us know how access to this document benefits you.

**Recommended Citation**

Stabler, Ben, "Development of the Idaho Statewide Travel Demand Model Trip Matrices Using Cell Phone OD Data and Origin Destination Matrix Estimation" (2014). *TREC Friday Seminar Series*. 30. [https://pdxscholar.library.pdx.edu/trec_seminar/30](https://pdxscholar.library.pdx.edu/trec_seminar/30)

This Book is brought to you for free and open access. It has been accepted for inclusion in TREC Friday Seminar Series by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.
Development of the Idaho Statewide Travel Demand Model Trip Matrices Using Cell Phone OD Data and Matrix Estimation

Portland State University, October 24, 2014
Topics

- Idaho Statewide Travel Model
- Cell Phone OD Data
- OD Matrix Estimation
- Validation
- Discussion
Statewide travel demand model

- In spring 2013, started building the Idaho Transportation Department (ITD) statewide travel demand model (STDM)
- Why? – as part of ITD’s data-driven, performance-based Investment Corridor Analysis Planning System (ICAPS)
- The two key requirements for the model are to forecast link level (road segment) auto and truck traffic, including external traffic
What is a travel demand model

- A series of mathematical equations that represent how choices are made when people travel
- Combines a network (supply) with population and employment by location (demand for travel)
Network

- Network for:
  - Routing trips
  - Generating travel time and distances between locations
  - Accumulating forecasted trips on roadway segments to estimate volumes

- Started from ITD’s GIS system so LRS coding is maintained

- Stitched-in MPO networks and FHWA’s network for areas beyond the state
Zone system

- All land use coded at the zone level
  - Uses MPO land use forecasts in order to be consistent
- Zones are the origin and destination of all travel in the model
- Developed the 4000+ zones in conjunction with MPOs and ITD District Planners
Zone System

- MPO zones ~3200
- Non-MPO ~350
- Buffer area ~600
- Remaining US & CA ~55
- Total zones 4200+
Travel demand (i.e. trips)

- How do we get an estimate of the travel demand for the entire model region? Two approaches in this project:
  - Phase 1 - we used cell phone origin-destination location data to synthesize travel demand
    - Not a forecast, but useful for estimating travel
  - Phase 2 - estimate models based on surveys and other data that forecast travel based on land use
    - Will have an activity-based person travel model and a FAF/Transearch disaggregation-based freight model
External travel demand

- External travel is travel coming in and/or going out of the study area.
- Very difficult to collect external travel data.
- Typically estimate external travel based on traffic counts, as opposed to land use.
- Cell phone OD data is emerging as a promising data set for external travel estimates.
- Cell phone OD data will be used for external travel estimates in both phase 1 and in phase 2.
Phase 1 model

Cell phone OD matrices and traffic counts

OD Matrix Estimation

Auto trips

Auto

Truck trips

Truck

Network Assignment

Volumes by user class, LOS measures

System Users

Auto

Truck

TREDIS Economic Model
Phase 2 model

Socioeconomic and Transearch data

Person Transport

- Long distance
  - Non-work trips
- Short distance
  - Commute and non-work trips

Freight Transport

- Long distance
- Short distance
  - Commodity flows
  - Truck trips

Performance Measures

Network Assignment

Volumes by user class, LOS measures

TREDIS Economic Model

System Users

- Tourism*
- Employment Retail
- Long haul Agriculture*
- Short haul Agriculture*
Cell Phone OD Data

- AirSage converts cell phone time and location data into trip OD data
- Has exclusive agreement with Verizon and others to aggregate and sell the cell phone location data
- Extracts the time and location of the cell phone every time it talks to the network - email, texts, phone calls, GPS, etc.
- Identifies cell device usual home and work location based on the cluster of points identifying where the phone “sleeps” at night and “works” during the day
Cell Phone OD Data

- Trips are coded with respect to the home and work anchor locations:
  - Home-based work
  - Home-based other
  - Non-home-based
  - Resident versus visitor
- AirSage expands the sampled trips to better match the population using various Census data sets
Cell Phone OD Data Request

- Calendar: Average weekday for the month of September 2013
- Markets: Resident HBW, HBO, and NHB; Visitor NHB
- Time period: Daily
- Zones: 750 x 750 super zone matrices to reduce cost
- Price: Quite reasonable
- License: Data licensed only for the project; derivative products can be used for other purposes though
Disaggregation to Model Zones and Initial Network Assignment

- Matrices disaggregated from 750 zones to 4000+ zones using each model zone’s share of super zone population and employment
- Results in daily raw cell phone flows between model zones for four markets
- Assign (or route) cell phone “trips” through the network using free flow travel time as the routing criteria
- Compare trip lengths to check results
  - Statewide model network trip lengths were joined to the trip records
Cell Phone HBW and Census JTW Trip Lengths

- Similar results within each District as well
Cell Phone OD and Boise MPO (COMPASS) Survey Trip Lengths

Coincidence Ratio and Average Trip Length Difference by Trip Category

<table>
<thead>
<tr>
<th>Trip Category</th>
<th>Coincidence Ratio</th>
<th>Average Trip Length (Miles)</th>
<th>Avg. Trip Length Difference</th>
<th>% Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>COMPASS</td>
<td>AirSage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HBW</td>
<td>0.72</td>
<td>8.65</td>
<td>8.84</td>
<td>0.19</td>
</tr>
<tr>
<td>HBO</td>
<td>0.91</td>
<td>4.94</td>
<td>5.33</td>
<td>0.39</td>
</tr>
<tr>
<td>NHB</td>
<td>0.68</td>
<td>4.19</td>
<td>6.24</td>
<td>2.05</td>
</tr>
</tbody>
</table>
Census JTW to Cell Phone Resident HBW Trips

- Would expect around 2 cell phone trips for each Census work journey
- The proportion of Census JTW to AirSage trips should be around 0.5
- This is essentially the case, as shown in the figure to the right, when summarized for all zone pairs
- The unexplained exception around 1.6 is flows to/from the state of Utah
## Trip Length Differences by District and Trip Type

<table>
<thead>
<tr>
<th>District</th>
<th>Trip Category</th>
<th>Coincidence Ratio</th>
<th>Average Trip Length (Miles)</th>
<th>Avg. Trip Length Difference</th>
<th>% Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Census JTW</td>
<td>AirSage</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Internal-Internal</td>
<td>0.89</td>
<td>15.18</td>
<td>15.03</td>
<td>-0.15</td>
</tr>
<tr>
<td></td>
<td>MPO Resident</td>
<td>0.98</td>
<td>8.08</td>
<td>8.43</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>Non-MPO Resident</td>
<td>0.89</td>
<td>28.69</td>
<td>30.60</td>
<td>1.91</td>
</tr>
<tr>
<td>2</td>
<td>Internal-Internal</td>
<td>0.81</td>
<td>22.03</td>
<td>23.00</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>MPO Resident</td>
<td>0.89</td>
<td>10.67</td>
<td>13.16</td>
<td>2.49</td>
</tr>
<tr>
<td></td>
<td>Non-MPO Resident</td>
<td>0.78</td>
<td>29.35</td>
<td>30.68</td>
<td>1.33</td>
</tr>
<tr>
<td>3</td>
<td>Internal-Internal</td>
<td>0.92</td>
<td>10.48</td>
<td>10.76</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>MPO Resident</td>
<td>0.92</td>
<td>8.92</td>
<td>9.61</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>Non-MPO Resident</td>
<td>0.78</td>
<td>24.17</td>
<td>26.74</td>
<td>2.57</td>
</tr>
<tr>
<td>4</td>
<td>Internal-Internal</td>
<td>0.78</td>
<td>20.72</td>
<td>21.01</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td>MPO Resident</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Non-MPO Resident</td>
<td>0.78</td>
<td>20.72</td>
<td>21.01</td>
<td>0.29</td>
</tr>
<tr>
<td>5</td>
<td>Internal-Internal</td>
<td>0.88</td>
<td>16.54</td>
<td>16.75</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>MPO Resident</td>
<td>0.89</td>
<td>9.95</td>
<td>11.24</td>
<td>1.29</td>
</tr>
<tr>
<td></td>
<td>Non-MPO Resident</td>
<td>0.86</td>
<td>25.33</td>
<td>24.91</td>
<td>-0.42</td>
</tr>
<tr>
<td>6</td>
<td>Internal-Internal</td>
<td>0.82</td>
<td>13.37</td>
<td>13.47</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>MPO Resident</td>
<td>0.88</td>
<td>8.85</td>
<td>9.43</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td>Non-MPO Resident</td>
<td>0.79</td>
<td>16.30</td>
<td>17.00</td>
<td>0.70</td>
</tr>
</tbody>
</table>
Observations from Initial Assignment

- Reasonable goodness-of-fit between the cell phone trip length distributions and the Census Journey-to-work and Boise MPO travel survey data sets
- Significant differences for NHB trips, especially short distance trips
- Why? Most likely a classification issue
  - Survey NHB trips are just household-based non-home-based trips, whereas the AirSage trips are everything else, including commercial vehicles
  - Very short trips in terms of distance and time may drop out of the AirSage data set as well
  - Simplified procedure to disaggregate super zone flows to model zones likely creating differences for some OD pairs
- Remember we’re comparing cell phone movements to person reported travel
Origin Destination Matrix Estimation

- Assign initial trip matrices to the daily statewide network using free flow travel time for impedance
- Adjust the trip demand matrices to minimize the difference between the estimated link volumes and traffic counts by user class
- Check difference between observed and estimated traffic volumes by user class (auto and truck) and facility type
- Repeat procedure until acceptable convergence

### Traffic Counts by Agency

<table>
<thead>
<tr>
<th>Count Source</th>
<th>Counts</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMPO</td>
<td>522</td>
</tr>
<tr>
<td>BTPO</td>
<td>434</td>
</tr>
<tr>
<td>COMPASS</td>
<td>2,811</td>
</tr>
<tr>
<td>KMPO</td>
<td>441</td>
</tr>
<tr>
<td>LCVMPO</td>
<td>500</td>
</tr>
<tr>
<td>ITD (only 10% used)</td>
<td>30,497</td>
</tr>
<tr>
<td>Total</td>
<td>35,205</td>
</tr>
</tbody>
</table>
Origin Destination Matrix Estimation (ODME)

Inputs
- Highway Network
- Trip Table

Highway Assignment

PRMSE Calculation

Closure criteria satisfied?
- Yes
- No

Matrix Estimation Loop

Run through all pre-defined groups?
- Yes
- No

OD Trip Table Adjustment

Adjustment Factors Calculation

Counts/Volume Skimming

Group Loop

Completion

Run through all pre-defined groups?
- Yes
- No

Matrix Estimation Loop
ODME Steps

• Input link level traffic counts
• Assign trip demand matrix (i.e. route trips through the network)
• Skim the sum of link traffic counts by OD
• Skim the sum of link assigned volumes (where count >0) by OD
• Calculate the ratio of count to assigned volume by OD
• Scale trip demand matrix by OD using the ratio calculated above
• Re-run assignment and repeat until converged
• Weight links by importance (i.e. larger counter = larger weight)
• Encourage solution convergence by averaging results across iterations (such as 50% this iteration + 50% previous iteration)

• Procedure borrowed from the Florida DOT
ODME Results

- %RMSE (goodness-of-fit measure) by ODME iteration
- Final %RMSE: auto 10.0%, truck 15.8%
## ODME Results

### %RMSE by Facility Type

<table>
<thead>
<tr>
<th>Type</th>
<th>Collector</th>
<th>Freeway</th>
<th>Local</th>
<th>Minor Arterial</th>
<th>Principal Arterial</th>
</tr>
</thead>
<tbody>
<tr>
<td>%RMSE</td>
<td>25.30%</td>
<td>2.80%</td>
<td>49.30%</td>
<td>15.10%</td>
<td>11.60%</td>
</tr>
</tbody>
</table>

![Graph showing %RMSE by Facility Type](image)
ODME Results

Assigned Volumes vs. Traffic Counts

$R^2 = 0.9941$

% Difference

[Map with color-coded lines indicating percentage differences]
ODME Results

- Good results in the MPOs and non-MPO areas as well
- All MPOs have similar results to COMPASS
ODME Results

- Reasonable trip length frequency results as well

HBW

NHB

HBO

Visitor
ODME Results

- County to county HBW flows for COMPASS area

<table>
<thead>
<tr>
<th>Destination</th>
<th>Ada</th>
<th>Canyon</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td>Origin</td>
<td>Adjusted Trips</td>
<td>Census JTW</td>
<td>Adjusted Trips</td>
</tr>
<tr>
<td>Ada</td>
<td>22.83%</td>
<td>19.49%</td>
<td>1.47%</td>
</tr>
<tr>
<td>Canyon</td>
<td>1.89%</td>
<td>2.83%</td>
<td>6.00%</td>
</tr>
<tr>
<td>Others</td>
<td>0.29%</td>
<td>0.83%</td>
<td>0.17%</td>
</tr>
</tbody>
</table>

- Reasonable goodness-of-fit between synthesized travel demand and limited observed data across multiple dimensions - user class, facility type, geography
ODME Criticisms

- ODME “naively” adjusts the travel demand to match the traffic counts.
- This can result in overfitting (which is where the model describes random error instead of the underlying relationships between variables).
- This means it can only be used for short-term forecasting, in which conditions are similar to today.
- ODME estimated traffic flows are a best-case scenario for goodness-of-fit since the process explicitly adjusts the input to better match the output.
- The phase 2 travel demand model, which is a function of land use, will be more sensitive to inputs, but is unlikely to match the traffic counts as well.
Discussion

- ITD wanted an ODME model in order to get components of the system (network, zones, trip matrices, etc.) up and running as early as possible in the project.
- The ODME model can be used for current year (and short term) estimates of roadway volumes by auto and truck.
- The next phase of the model will be more of a long range forecasting tool since it is a function of land use (which drives travel demand).
Discussion Continued

- The cell phone OD data is a reasonable starting point for generating statewide trip matrices.
- Used in conjunction with existing travel modeling tools and techniques, the cell phone OD data has a very promising future in our industry.
- Additional work is required to better understand how cell phone flows are different than traditional travel data sets.
- The phase 2 demand models will only replace the internally generated travel and so the cell phone trip matrices will still be used to model the external travel.
More Information

- David Coladner, David.Coladner@itd.idaho.gov
- Ben Stabler, stabler@pbworld.com
- Sujan Sikder, sikders@pbworld.com
- Bob Schulte, rjs@dksassociates.com