Integrated Transportation/Land-Use Modeling: Using publicly available data sources for advanced modeling

Rolf Moeckel
Parsons Brinckerhoff

TRB Conference
Using Census Data for Transportation Applications
University of California Irvine · 25-27 October 2011
Interaction of Land Use and Transportation

Wegener (1994)
Need for land-use modeling

• It is known that transportation and land use influence each other. The extent and direction of this influence usually is guessed only.

• Integrated modeling helps
  – decision makers,
  – staffing,
  – transparency,
  – public participation,
  – and the tradeoffs between these...
Concept of integrated simulation models

- Land use
- Transportation demand
- Dynamic traffic assignment
- Environmental impacts

Feedback arrows indicate the flow of information:
- Environmental feedback from environmental impacts to land use
- Transport feedback from transportation demand to dynamic traffic assignment
Simple Land Use Orchestrator

SILO IDEA
The world of land use modeling (in the US)

- **Index**
- **Place³s**
- **What-If**

**SILO**
- Integrated with transportation model
- Simple to implement
- Easy to run scenarios
- Long run times

**PECAS**
- UrbanSim
- Data hungry

**TRANUS**

**What-If Index**

**Place³s**

**Simple to implement**

**Easy to run scenarios**

**Integrated with transportation model**

**Data hungry**

**Long run times**
Motivation

• Fully represent the land-use transport feedback cycle
• Develop a tool that can be implemented with tight schedules and/or a smaller budget
• Ability to analyze smart growth strategies
• Simulate greenhouse gas emissions
• Ability to run many scenarios in a short time frame
Design Responses by SILO

- Microsimulation model
- Policy-sensitive
- Use zones for spatial resolution but accommodate raster or parcel databases
- User-friendly GUI
- Required data shall be publicly available
- Short run times (e.g. 10 minutes for one simulation period in a large region of 5 million people)
Policy scenarios

- Zoning
- Changes in transportation costs
- Planned job/housing balance
- Mixed Development
- Transit-Oriented Developments (TOD)
- Alternative growth assumptions
Simple Land Use Orchestrator

CONCEPT
Concept

Land Use Model

- Households
  - Demography
  - Moves

- Dwellings
  - Construction
  - Demolition
  - Down-/upgrading

Accessibility

- Jobs
  - Firmography
  - Relocation

- Non-res. floorspace
  - Construction
  - Demolition
  - Down-/upgrading

Environmental impacts

Travel model
Graphical User Interface (GUI)
Data Input

• Essential data
  – PUMS data
    • HH by size, income and auto-ownership
    • Persons by age and gender
    • Dwellings by type and price
  – Non-res floorspace by type and price
  – Forecasts of pop/emp growth for study area

• Desirable (can use national averages instead)
  – Preferences for location choice
  – Demographic transition probabilities
Output Indicators

- Population by age, hh size, auto ownership
- Employment by type
- Dwelling vacancy
- Non-residential floorspace vacancy
- Change in housing costs
- Land-use consumption
Prototype application for MetCouncil
Simple Land Use Orchestrator

SYNTHETIC POPULATION
Required micro data

• As a microsimulation, SILO needs microscopic data.
• As these data are not publicly available, a synthetic population has been created.
• Microscopic PUMS is used to create micro data.
• Micro data cover households, person, dwellings, employment and non-residential floorspace
Microdata: Households

- Household size
- Number of autos

- Link to every person in household
- Link to dwelling
Microdata: Persons

- Age
- Gender
- Education
- Occupation
- Drivers license
- Workplace zone (given by PUMA)
- Income
- Relationship (single, married, child)
Microdata: Dwellings

• Zone
• Type
• Area in ft$^2$
• Quality (complete plumbing/complete kitchen)
• Monthly cost (rent or mortgage)
Public Use Micro Sample (PUMS)

- Use PUMS 2000 data
- Use 5% sample to increase spatial resolution
20 PUMA Regions
Methodology

• Every PUMS record provides information on household, the dwelling, and all persons in this household.

• Every PUMS record has a weight that is used to expand the PUMS data to full population.
PUMS Sample Size

- PUMS 5% sample file for Minnesota has 359,332 records
- 37,341 thereof are in Minneapolis/Saint Paul study area
- Expanding these data generates
  - 1,017,153 Households
  - 2,405,601 Persons
Households by Size, Age and PUMA
Dwellings by type and PUMA
Validation of Synthetic Population

<table>
<thead>
<tr>
<th>Population</th>
<th>Observed</th>
<th>Modeled</th>
<th>Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Households</td>
<td>1,024,263</td>
<td>1,017,153</td>
<td>-0.69%</td>
</tr>
<tr>
<td>Persons</td>
<td>2,415,649</td>
<td>2,405,601</td>
<td>-0.42%</td>
</tr>
</tbody>
</table>
Issues encountered

• PUMS defines age as 0 to 90+, need to generate age distribution for 91, 92, 93, ...

• Ignore records with dwelling type 10 (Boat, RV, Van → irrelevant for land-use model) and -999 (unknown).

• PUMS provides
  • 1,017,682 Dwellings for
  • 1,017,153 Households
Vacant dwellings by zone (raw)
Vacant dwellings by zone (adjusted)
Validation of Synthetic Population

<table>
<thead>
<tr>
<th>Dwellings</th>
<th>Observed</th>
<th>Modeled</th>
<th>Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>SFD</td>
<td>633,811</td>
<td>630,599</td>
<td>-0.5%</td>
</tr>
<tr>
<td>SFA</td>
<td>87,348</td>
<td>89,072</td>
<td>2.0%</td>
</tr>
<tr>
<td>MF234</td>
<td>62,790</td>
<td>60,670</td>
<td>-3.4%</td>
</tr>
<tr>
<td>MF5plus</td>
<td>246,555</td>
<td>250,482</td>
<td>1.6%</td>
</tr>
<tr>
<td>MH</td>
<td>16,718</td>
<td>16,178</td>
<td>-3.2%</td>
</tr>
<tr>
<td>Total</td>
<td>1,047,222</td>
<td>1,047,001</td>
<td>-0.02%</td>
</tr>
</tbody>
</table>
CONCLUSIONS
Next steps

• SILO will be calibrated for Minneapolis/Saint Paul.
• ACS data for 2010 will be used to validate the land-use model for 10 years of simulation
Conclusions

• Complexity must be well-balanced: As complex as needed yet as simple as possible (but not simpler).

• Independent data (i.e. data not used in model development) are crucial for model validation

• No need to over-engineer micro data