Quality Assurance/Quality Control for the Atlanta Activity-Based Model: A “Look Under the Hood”

Kyeil Kim, Guy Rousseau, Kyung-Hwa Kim & David D’Onofrio

4th International Conference on Innovations in Travel Modeling
April 30-May 2, 2012  Tampa, Florida
“Anti-intellectualism ... nurtured by the false notion that democracy means that ‘My ignorance is just as good as your knowledge.’”

– Isaac Asimov
Background

- ARC travel forecasting for the metro Atlanta 20-county nonattainment area
  - 4-Step Trip-Based Model (TBM) & Activity-Based Model (ABM)
  - Transportation plan development
  - Air quality conformity determinations
  - Sales tax referendum

- Motivation
  - Thorough understanding of ABM (capabilities and limits)
  - Ensure model quality through critical review
Objective

- Quality Assurance (QA)
  - A systematic review process by personnel not directly involved in model development

- Quality Control (QC)
  - A technical routine to control the quality of the model performed in model development

Objective

- Present the QA/QC project implemented for the Atlanta Activity-Based Model by ARC staff
- Focus on the QA/QC “process”
Structure of ARC ABM (Overall)
Model Structure (CT-RAMP)

- **CT-RAMP**: Coordinated Travel Regional Activity-Based Modeling Platform

![Diagram of Model Structure (CT-RAMP)]
Model Run Setup

- **System setup**
  - Three 64-bit Windows Server machines
  - 16 cores and 32 GB of RAM for each server
  - Cube v5.1.3 (clusters)
  - 64-bit Java Development Kit 1.6
  - 32-bit Java Runtime Environment

- **Model Runtime**
QA/QC Procedures

- **Guidelines**
  - ARC, *Quality Assurance/Quality Control of Travel Demand Models*, April 2011
  - **Inputs**
    - Inputs from ARC modelers and non-modeling planners
  - **Contents**
    - Reasonableness checking on all modeling components
    - Comparability between TBM and ABM
    - Temporal validation between base and forecast years
QA/QC Procedures (cont’d)

- **Limitations**
  - Vast output database
    - Year 2040 - approx. 8.0 million persons; 9.5 million tour records, 25.2 million trip records
    - Each model scenario ranging 15GB ~ 20GB
  - Software/Programs
    - MS Access, Excel spreadsheet, Programming language

- **Tools**
  - Requirements: Interactive, Flexible, Responsive
  - SQL Server Database Management Studio
  - Visualization system (ABMVIZ)
SQL Server DB Management Studio

- **Structured Query Language**
- **Transact-SQL**
- **SQL Server Database**
  - A collection of data tables
  - Relational database
  - Schema: organizational description of database, including the data tables

- **ABM Run Outputs**
  - SQL Server Database (ARCABM)
  - Each model run scenario stored in a separate schema
  - Each schema contains various model outputs for the scenario
SQL Server DB Management (cont’d)
Visualization

- **ABMVIZ**
  - Model visualization dashboard application in Adobe AIR (Adobe Integrated Runtime) Flash
  - Linked to the SQL database through Data Access Layer
  - Interactive, dynamic and visual
  - Tables, charts, time-use diagrams, tour tracing, tree map and radar charts
ABMVIZ

- Bar Chart/Map
- Tree Map
- Time-use Diagram
- Radar Chart
Review of Model Components

- Identification of model building blocks
- Understand interactions (model flow)
- Focus on each modeling step
Population Synthesizer (PopSyn)

- Creates synthetic populations based on household control variables, Census data and PUMS

- Data segmentation by key attribute
  - Single and cross tabulations across the attributes (e.g., distributions by person type, employment category, student category, and their combinations)

- Temporal changes of the distributions (aging factor)
Population Synthesizer (cont’d)

- Reasonableness in income distributions by person type
- Comparison with trip-based model

Zonal workers (TBM)  
- Wrkrs_@home TBM
  - 0 to 271
  - 272 to 554
  - 555 to 799
  - 800 to 1099
  - 1100 to 1459
  - 1460 to 1934
  - 1935 to 2629
  - 2630 to 10000

Zonal workers (ABM)  
- Wrkrs_@home ABM
  - 271 and below (248)
  - 272 to 554 (243)
  - 555 to 799 (246)
  - 800 to 1099 (248)
  - 1100 to 1459 (244)
  - 1460 to 1934 (247)
  - 1935 to 2629 (245)
  - 2630 and above (246)
  - Other (57)
Mandatory Activity Location Choice

- Work/K-12/university locations for synthesized population
- \( f(\text{size variables, income, area type, mode choice logsum}) \)
- County-by-county comparisons with ACS for work destination choice model
- Comparison with real-world mandatory activity location data
Mandatory Activity Location Choice

School Location

University Location
Vehicle Ownership Model

- Number of autos owned by household
  - \( f(\text{household size, household income, drivers/workers in household, parking costs at residential locations, and mode choice logsums}) \)

- Reasonableness of single/crosstab relationships between dependent and independent variables

- Model’s sensitivity to increasing congestion in the future

- Discrepancies between ABM and TBM (average car ownership, geographical distributions)
Vehicle Ownership Model (cont’d)

- Zonal average car ownership

<table>
<thead>
<tr>
<th>Miles</th>
<th>TBM vehown</th>
<th>ABM vehown</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-3</td>
<td>1.00 to 1.33</td>
<td>1.00 to 1.33</td>
</tr>
<tr>
<td>3-10</td>
<td>1.33 to 1.57</td>
<td>1.33 to 1.57</td>
</tr>
<tr>
<td>10-15</td>
<td>1.57 to 1.70</td>
<td>1.57 to 1.70</td>
</tr>
<tr>
<td>15-20</td>
<td>1.70 to 1.82</td>
<td>1.70 to 1.82</td>
</tr>
<tr>
<td>20-25</td>
<td>1.82 to 1.90</td>
<td>1.82 to 1.90</td>
</tr>
<tr>
<td>25-30</td>
<td>1.90 to 1.97</td>
<td>1.90 to 1.97</td>
</tr>
<tr>
<td>30-35</td>
<td>1.97 to 2.03</td>
<td>1.97 to 2.03</td>
</tr>
<tr>
<td>35-40</td>
<td>2.03 to 2.09</td>
<td>2.03 to 2.09</td>
</tr>
<tr>
<td>40+</td>
<td>2.09 to 10.00</td>
<td>2.09 to 2.16</td>
</tr>
</tbody>
</table>

OTHE
Coordinated Daily Activity Pattern

- Daily activity patterns by household member
  - Mandatory, Non-mandatory or At-home for each person
  - Coordinated patterns among household members
  - $f$ (Person/HH attributes, car ownership, income, accessibility measures, intra-household interaction terms)

- Relationship between person type and DAP
- Changes in distribution of activity patterns from base year to future years (~aging factor)
Tour Models

- Predict the number and purpose of tours for each person, destinations, and time-of-day choices
- Four different tours

Diagram:

- Individual Mandatory
  - Residual Time
  - Joint Non-Mandatory
  - Individual Non-Mandatory
  - At-Work Sub-Tours
## Individual Tour Models

<table>
<thead>
<tr>
<th>Major Explanatory Variables</th>
<th>Individual Mandatory Tour</th>
<th>Individual Non-Mandatory Tour</th>
<th>At-Work Sub-Tour</th>
</tr>
</thead>
<tbody>
<tr>
<td>Household Characteristics</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Person Characteristics</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Income</td>
<td>✓</td>
<td>❌</td>
<td>✓</td>
</tr>
<tr>
<td>Car Ownership</td>
<td>✓</td>
<td>❌</td>
<td>✓</td>
</tr>
<tr>
<td>Location of Tour Activities</td>
<td>✓</td>
<td>❌</td>
<td>✓</td>
</tr>
<tr>
<td>Size Variables</td>
<td>✓</td>
<td>❌</td>
<td>✓</td>
</tr>
<tr>
<td>Area Type</td>
<td>✓</td>
<td>❌</td>
<td>✓</td>
</tr>
<tr>
<td>Duration of Parent Work Tour</td>
<td>✓</td>
<td>❌</td>
<td>✓</td>
</tr>
<tr>
<td>Mode-Choice Logsums</td>
<td>✓</td>
<td>❌</td>
<td>✓</td>
</tr>
</tbody>
</table>

- **frequency**
- **time-of-day**
- **destination**
- **time-of-day**
Individual Tours

- Aggregation check
  - Average tours/person, temporal variations
  - Reasonability in proportions of tours by person type, tour category, and tour purpose and by their combinations
  - Trip rate comparisons with TBM

- Zonal tour distributions by activity pattern
  - Workplace location choice model vs. tour destination choice model

- Tour time-of-day choice models
  - Reasonability in departure hours, arrival hours and tour durations by tour purpose
Individual Tours

- Tour frequency model (example)
Individual Tours (cont’d)

- Tour time-of-day choice models (examples)
Joint Non-Mandatory Tour

Major Explanatory Variables

- Household Characteristics
- Person Characteristics
- Overlapping Time Window
- Income
- Car Ownership
- Area Type
- Tour Purpose
- Mode-Choice Logsums

Joint Non-Mandatory Tour
Joint Tours

- Tour frequency by tour purpose
- Tour frequency by household size and number of joint trips
  - 1 joint tour: 85% of household making joint tours
- Distribution of tour compositions
  - Adults-accompanied: 94%
- Number of participants by tour purpose
  - 2-person joint tours: 74%
- Variations of tour departure hours, arrival hours and durations by tour purpose
Joint Tours

Distribution of Tours by Participant

Tour Durations by Tour Purpose
Lessons Learned

- Exposure to the SE and travel data unavailable in TBM - beneficial to both modelers and non-modelers
- Understand the working mechanism of ABM
- Evaluate the model’s capabilities and limitations
- A team approach: technical and pragmatic
- SQL/ABMVIZ: efficient tools
- Opportunity to improve both TBM and ABM through cross-referencing
- Resulted in corrective actions for both developers and users
Next Steps

- Continue on the QA/QC project
  - Validation through the Census data and the new household travel survey
- Eventual goal of using ABM for the plan development and update
- Investigate the issues of randomness in the ABM simulation process
- More scenario and sensitivity analyses
- Run the ABM all the way through vehicle emission model (MOVES)
Team

- **ARC**
  - Kyeil Kim, kykim@atlantaregional.com
  - Guy Rousseau, grousseau@atlantaregional.com
  - Kyung-Hwa Kim, kkim@atlantaregional.com
  - David D’Onofrio, dd’onofrio@atlantaregional.com

- **Consultants**
  - Joel Freeman (PB), freeman@pbworld.com
  - Ben Stabler (PB), stabler@pbworld.com
  - Patti Schropp (Atkins), patti.schropp@atkinsglobal.com
  - Jonathan Nicholson (Atkins), jonathan.nicholson@atkinsglobal.com