Development of Standardized Models in Kentucky Using AirSage Data

CATMUG Meeting

Chicago, IL
December 7, 2016
Outline

- Background
- AirSage Pre-processing & Adjustment
- AirSage Application
  - Trip Generation
  - Trip Distribution
  - Time-of-Day
  - Directional Factor
- Findings
KYTC’s regional TDMs

- Currently using TransCAD 5. In transition to TransCAD 6.
- Standardized organization of files, model stream, script and interface
- Trip-based, 3-step (TG, TD & TA).
- 5 auto trip purposes (HBW, HBO, NHB, HBSchool, HBUniversity).
- 3 truck trip purposes (light, medium, heavy)
- 4 time periods (AM, PM, MD, NT)
- Flexibility of using Big Data to enhance model components.
AirSage in Kentucky

- No recent HH OD survey data in KY. No NHTS add-on.
- Cell phone data (not GPS data).
- Lower cost compared to traditional HH travel surveys.
- Quick turnover.
- AirSage provides ODs by purpose and time period.
- When data collection area is large enough, external trip info. can be obtained. Good data at the edge of study area.
- KYTC first purchased AirSage data for Lexington Model (2012). Complete 5 regional models so far. All used AirSage data.
- KYTC and its consultants have gained thorough insight in using AirSage data for model development.
Background

KYTC’s Completed Regional Models

- Lexington/Central KY Region (2012)
- Bowling Green/Warren Co. (2014)
- E-town/Hardin Co./Meade Co. (2014)
- Owensboro/Daviess Co. Region (2014)
  - 160K pop
  - 3 Co. + parts of 2 Co.
  - Small MPO
- KYTC District 9 Regional Model (2015)
  - 300K pop
  - 8 KY Co. + 3 OH Co.
  - Rural
AirSage Data Pre-processing

Example of AirSage Data (.CSV)

<table>
<thead>
<tr>
<th>Origin_Zone</th>
<th>Destination_Zone</th>
<th>Start_Date</th>
<th>End_Date</th>
<th>Aggregation</th>
<th>Subscriber_Class</th>
<th>Purpose</th>
<th>Time_of_Day</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>74</td>
<td>286</td>
<td>20150415</td>
<td>20150514</td>
<td>WD</td>
<td>Resident</td>
<td>OH</td>
<td>H1800:H2400</td>
<td>12.47</td>
</tr>
<tr>
<td>315</td>
<td>225</td>
<td>20150415</td>
<td>20150514</td>
<td>WD</td>
<td>Visitor</td>
<td>OO</td>
<td>H0900:H1500</td>
<td>5.9</td>
</tr>
<tr>
<td>432</td>
<td>433</td>
<td>20150415</td>
<td>20150514</td>
<td>WD</td>
<td>Resident</td>
<td>WH</td>
<td>H0900:H1500</td>
<td>11.84</td>
</tr>
</tbody>
</table>

- Must provide pre-defined TAZ polygons to AirSage. AirSage has resolution limit (0.25 square mile grid).
- AirSage provided “expanded” data which matches census population by carriers.
- When multiple carriers exist (e.g., Owensboro Model), averaging data sets minimizes data bias.
AirSage Data Pre-processing

- **Day Parts (can be customized to match model)**
  - AM peak = 6:00-9:00
  - Mid-day = 9:00-15:00
  - PM peak = 15:00-18:00
  - Night = 18:00-24:00 & 0:00-6:00

- **Trip Purpose: (H-home, W-work, O=other)**
  - HBW = HW, WH
  - HBO = HO, OH, HH
  - NHB = WO, OW, WW, OO

- Convert .CSV file to OD matrix (3 purposes x 4 periods)
Data collected for west-central KY
- 95% of 566 Owensboro zones had data
- 94% Pop & 97% Emp are covered
- Lack of data in adjacent rural counties with poor cellar service. No impact on external trips (0.2% of Owensboro work flows - CTPP).
Data collected only for D9 area

A 500-zone structure was used for AirSage data acquisition.

85% of 500 zones had data.

92% Pop & 93% Emp are covered

Uncovered zones are very rural, sparsely populated areas

A refined 620 zones (final model TAZ) was used for friction factor development.
Power of Big Data

- HH survey: 11% of total zones (an Indiana example)
- AirSage: 85 ~ 95% of total zones
  (Owensboro – 28,085 OD pairs with data)
  (District 9 – 22,111 OD pairs with data)
Power of Big Data

What can we see based on 11%?
Power of Big Data

How about based on 90%?
Power of Big Data

Big Data = Insight
### AirSage Data Adjustments (Intrazonals)

**Owensboro**

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Total</th>
<th>Intrazonal</th>
<th>Intrazonal %</th>
</tr>
</thead>
<tbody>
<tr>
<td>HBW</td>
<td>150,012</td>
<td>1,951</td>
<td>1.3%</td>
</tr>
<tr>
<td>HBO</td>
<td>330,101</td>
<td>89,822</td>
<td>27.2%</td>
</tr>
<tr>
<td>NHB</td>
<td>95,250</td>
<td>29,640</td>
<td>31.1%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>575,363</strong></td>
<td><strong>121,413</strong></td>
<td><strong>21.1%</strong></td>
</tr>
</tbody>
</table>

**District 9**

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Total</th>
<th>Intrazonal</th>
<th>Intrazonal %</th>
</tr>
</thead>
<tbody>
<tr>
<td>HBW</td>
<td>118,429</td>
<td>5,339</td>
<td>4.5%</td>
</tr>
<tr>
<td>HBO</td>
<td>654,781</td>
<td>324,717</td>
<td>49.6%</td>
</tr>
<tr>
<td>NHB</td>
<td>355,881</td>
<td>109,520</td>
<td>30.8%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1,129,090</strong></td>
<td><strong>439,576</strong></td>
<td><strong>38.9%</strong></td>
</tr>
</tbody>
</table>

Note: Data is from raw AirSage (all trips)

- Except for HBW trips, intrazonal trip percentages seem too high
- HBO & NHB introzonals were reduced to more conventional levels before developing friction factors.
Trip Generation - Trip Rates

- Previous experience indicates AirSage may not be suitable for directly determining trip generation rates. Analysis is limited due to aggregated data & resolution limits.
- Owensboro (small MPO) has too few NHB trips. Home-based additional trips or short-distance trips may not be captured.
- D9 region (rural area) generally has more trips.
- Not much improvement by ODME

### Owensboro

<table>
<thead>
<tr>
<th>Purpose</th>
<th>AirSage Internal P's</th>
<th>AirSage %</th>
<th>Calibrated Model Internal P's</th>
<th>Calibrated Model %</th>
</tr>
</thead>
<tbody>
<tr>
<td>HBW</td>
<td>104,628</td>
<td>24%</td>
<td>100,543</td>
<td>20%</td>
</tr>
<tr>
<td>HBO</td>
<td>262,339</td>
<td>60%</td>
<td>262,892</td>
<td>52%</td>
</tr>
<tr>
<td>NHB</td>
<td>72,025</td>
<td>16%</td>
<td>140,013</td>
<td>28%</td>
</tr>
<tr>
<td>Total</td>
<td>438,992</td>
<td>100%</td>
<td>503,448</td>
<td>100%</td>
</tr>
</tbody>
</table>

| Trip Rates | 6.8 | 7.7 |

### District 9

<table>
<thead>
<tr>
<th>Purpose</th>
<th>AirSage Internal P's</th>
<th>AirSage %</th>
<th>Calibrated Model Internal P's</th>
<th>Calibrated Model %</th>
</tr>
</thead>
<tbody>
<tr>
<td>HBW</td>
<td>118,429</td>
<td>10%</td>
<td>160,704</td>
<td>18%</td>
</tr>
<tr>
<td>HBO</td>
<td>654,781</td>
<td>58%</td>
<td>466,274</td>
<td>53%</td>
</tr>
<tr>
<td>NHB</td>
<td>355,881</td>
<td>32%</td>
<td>258,846</td>
<td>29%</td>
</tr>
<tr>
<td>Total</td>
<td>1,129,090</td>
<td>100%</td>
<td>885,824</td>
<td>100%</td>
</tr>
</tbody>
</table>

| Trip Rates | 9.6 | 7.6 |
Trip Generation - Area Type Factors

Assumption: Data bias is diminished, if not completely offset between area type samples, by a factoring process.

- Rural, Town & 2nd City
- Convert OD to PA
  - HBW = HW + Transpose(WH)
  - HBO = HO + Transpose(OH) + HH
- NHB data not used (not real P’s by HH)
- AirSage P’s rates (by purpose & area type)
  - Remove outliers (95% CI)
Trip Generation - Area Type Factors

- Area Type Factors (t-test of P rates between area types)
  - If significantly different, factor = ratio of P rates
  - Otherwise, factor = 1.0
- Factors of Rural vs. Towns are reasonable (HBW=1.0, HBO=0.9)
- Factors for Rural vs. 2nd City, Towns vs. 2nd City are too large.
- Available data (KYSTM) suggests same rates for Towns & 2nd City (consolidated as Non-Rural).
- For NHB, factor (=1.0) is derived from NCHRP 716 and KYSTM.

<table>
<thead>
<tr>
<th>Purpose</th>
<th>D9 Model Area Type Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rural</td>
<td>Non-Rural</td>
</tr>
<tr>
<td>HBW</td>
<td>1.0</td>
</tr>
<tr>
<td>HBO</td>
<td>0.9</td>
</tr>
<tr>
<td>NHB</td>
<td>1.0</td>
</tr>
</tbody>
</table>
Owensboro Method

- AirSage PA tables (internal, by purpose) are fratared to match default target P’s (NCHRP 716) and A’s (NCHRP 365).
- Fratared trip tables work better than raw data

<table>
<thead>
<tr>
<th>Traffic Assignment</th>
<th>Total Traffic (Count Links)</th>
<th>Avg. Dev % from Counts *</th>
</tr>
</thead>
<tbody>
<tr>
<td>AirSage (raw)</td>
<td>907,259</td>
<td>-59.2%</td>
</tr>
<tr>
<td>AirSage (fratared)</td>
<td>1,397,946</td>
<td>-19.8%</td>
</tr>
<tr>
<td>Traffic Counts</td>
<td>2,275,196</td>
<td></td>
</tr>
</tbody>
</table>

* Simple average error, not RMSE

- Friction factors were developed using TransCAD gravity model calibration function. Further adjustment in model calibration.
Trip Distribution – Friction Factors

Owensboro Method

<table>
<thead>
<tr>
<th>Trip Purpose</th>
<th>Mean Trip Length (min)</th>
<th>AirSage</th>
<th>Calibrated Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>HBW</td>
<td>19.1</td>
<td></td>
<td>15.7</td>
</tr>
<tr>
<td>HBO</td>
<td>17.8</td>
<td></td>
<td>14.9</td>
</tr>
<tr>
<td>NHB</td>
<td>19.4</td>
<td></td>
<td>14.4</td>
</tr>
</tbody>
</table>
Trip Distribution – Friction Factors

District 9 Method (revised approach)

- KY local trip rates usually are unknown and differ from national defaults. Calibrate AirSage OD tables to traffic counts.

- ODME

  \[
  \text{Full seed OD table} = \text{I-I} \uparrow + \text{I-E/E-I} \uparrow + \text{E-E} \uparrow
  \]
  
  - AirSage (all purpose)
  - Land Use Data + NCHRP 716
  - KYSTM extraction + Fratar

- Disaggregate ODME trip table by purpose – based on original AirSage data

- Friction factors (by purpose) were developed using TransCAD gravity model calibration function.
### AirSage – Trip Length Distribution

#### District 9 Method (revised approach)

<table>
<thead>
<tr>
<th>Trip Purpose</th>
<th>Mean Trip Length (min)</th>
<th>AirSage</th>
<th>Calibrated Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>HBW</td>
<td>16.8</td>
<td>17.3</td>
<td></td>
</tr>
<tr>
<td>HBO</td>
<td>13.3</td>
<td>12.5</td>
<td></td>
</tr>
<tr>
<td>NHB</td>
<td>15.3</td>
<td>13.7</td>
<td></td>
</tr>
</tbody>
</table>

![AirSage Trip Length Frequency Distribution](image)

![Friction Factor - HBW](image)

![Friction Factor - HBO](image)

![Friction Factor - NHB](image)
### AirSage TOD & Directional Factors

#### Owensboro (TOD)

<table>
<thead>
<tr>
<th>Purpose</th>
<th>AM (6-9am)</th>
<th>Mid-Day (9am-3pm)</th>
<th>PM (3-6pm)</th>
<th>Night (6pm-6am)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HBW</td>
<td>0.26</td>
<td>0.25</td>
<td>0.18</td>
<td>0.31</td>
</tr>
<tr>
<td>HBO</td>
<td>0.17</td>
<td>0.30</td>
<td>0.19</td>
<td>0.34</td>
</tr>
<tr>
<td>NHB</td>
<td>0.12</td>
<td>0.47</td>
<td>0.23</td>
<td>0.18</td>
</tr>
<tr>
<td>E-I Auto</td>
<td>0.20</td>
<td>0.30</td>
<td>0.20</td>
<td>0.30</td>
</tr>
<tr>
<td>E-E *</td>
<td>0.19</td>
<td>0.31</td>
<td>0.20</td>
<td>0.30</td>
</tr>
</tbody>
</table>

* Assumed by average of all purposes

#### District 9 (TOD)

<table>
<thead>
<tr>
<th>Purpose</th>
<th>AM (6-9am)</th>
<th>Mid-Day (9am-3pm)</th>
<th>PM (3-6pm)</th>
<th>Night (6pm-6am)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HBW</td>
<td>0.28</td>
<td>0.27</td>
<td>0.19</td>
<td>0.26</td>
</tr>
<tr>
<td>HBO</td>
<td>0.17</td>
<td>0.30</td>
<td>0.19</td>
<td>0.34</td>
</tr>
<tr>
<td>NHB</td>
<td>0.19</td>
<td>0.37</td>
<td>0.22</td>
<td>0.22</td>
</tr>
<tr>
<td>E-E *</td>
<td>0.19</td>
<td>0.32</td>
<td>0.20</td>
<td>0.29</td>
</tr>
</tbody>
</table>

* Assumed by average of all purposes

#### Owensboro (direction – P to A)

<table>
<thead>
<tr>
<th>Purpose</th>
<th>AM (6-9am)</th>
<th>Mid-day (9am-3pm)</th>
<th>PM (3-6pm)</th>
<th>Night (6pm-6am)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HBW</td>
<td>0.91</td>
<td>0.51</td>
<td>0.16</td>
<td>0.43</td>
</tr>
<tr>
<td>HBO</td>
<td>0.85</td>
<td>0.53</td>
<td>0.43</td>
<td>0.33</td>
</tr>
<tr>
<td>E-I Auto</td>
<td>0.85</td>
<td>0.53</td>
<td>0.43</td>
<td>0.33</td>
</tr>
</tbody>
</table>

* Assume equal to HBO

#### District 9 (direction – P to A)

<table>
<thead>
<tr>
<th>Purpose</th>
<th>AM (6-9am)</th>
<th>Mid-day (9am-3pm)</th>
<th>PM (3-6pm)</th>
<th>Night (6pm-6am)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HBW</td>
<td>0.94</td>
<td>0.49</td>
<td>0.16</td>
<td>0.39</td>
</tr>
<tr>
<td>HBO</td>
<td>0.85</td>
<td>0.54</td>
<td>0.42</td>
<td>0.34</td>
</tr>
<tr>
<td>E-I Auto</td>
<td>0.85</td>
<td>0.54</td>
<td>0.42</td>
<td>0.34</td>
</tr>
</tbody>
</table>

* Assume equal to HBO

- **AirSage TOD / Directional factors seem reasonable.**
- **TOD factors can be adjusted further to match TOD counts in model calibration.**
- **Large data collection area = good data for external trips & at study area edges.**
Traffic Assignment

D9 Model RMSE%:

- Overall = 42.6%  (KYTC target = 55%)
- Rowan County = 24.0%  (KYTC target = 35%)
AirSage Advantages/Disadvantages

**Advantages**

- Very large dataset.
- Low cost when compared to surveys. Quick turnover.
- Trip purpose, TOD & direction are available.
- Large coverage are readily available.
- Good data penetration / saturation

**Disadvantages** – *adjustments required*

- Everything is aggregate. Characteristics of travelers are not available.
- Trip purpose is based on activity clusters and times of data transmissions.
- Data transmission times may not accurately reflect when travel occurs.
- Resolution limits & thresholds of determining device’s location (5 min / 300 meters) – may miss short- dist./duration trips, i.e., small zones or urban areas.
- Unusual results in the trip tables (intrazonal, trip rates)
Summary of Findings

- Findings are based on our understanding of the AirSage product. AirSage product is evolving, so future data may address the issues.

- AirSage data may not be suitable for directly determining trip generation rates. Analysis are limited due to resolution limit and aggregate data.

- AirSage seems to offer a way of estimating the area type impact on trip generation. Validation with other data sources are required.

- Refinement & adjustment of AirSage trip tables are essential before trip distribution analysis. An ODME process is recommended.

- AirSage seems to provide reasonable estimates of existing temporal distribution and directional patterns.

- AirSage data provides a good-starting point for estimating model parameters in small MPO and rural areas.
Questions?

Johnny Han, PhD, PE
Corradino
jhan@Corradino.com, 317.744.9858

Ken Kaltenbach, PE
Corradino
kkaltenbach@Corradino.com, 502.587.7221

Scott Thomson, PE
KYTC
Scott.Thomson@ky.gov, 502.782.5086

Jayalakshmi Balaji, PE
KYTC
Jayalakshmi.balaji@ky.gov, 502.782.5045