

Innovative Analysis Methods of Mobile Phone Data in the Best Travel Demand Modeling Practice in Kentucky

Yang Han, PhD, PE (The Corradino Group)

Ken Kaltenbach, PE (The Corradino Group)

Scott Thomson, PE (KYTC)

Jayalakshmi Balaji, PE (KYTC)

Daniel Hulker (KYTC)

15th National Tools of the Trade Conference

Charleston, South Carolina

September 12, 2016



Outline

- **Background**
- **AirSage Data Pre-processing & Adjustment**
- **AirSage Data Application**
 - ❑ Trip Generation
 - ❑ Trip Distribution
 - ❑ Time-of-Day
 - ❑ Directional Factor
- **Findings**

Background

KYTC's regional TDMs

- Trip-based, 3-step (TG, TD & TA), TOD model
- Standardized organization of files, model stream, script and interface
- Flexibility of using available data sources (e.g., mobile phone data) to enhance model components.

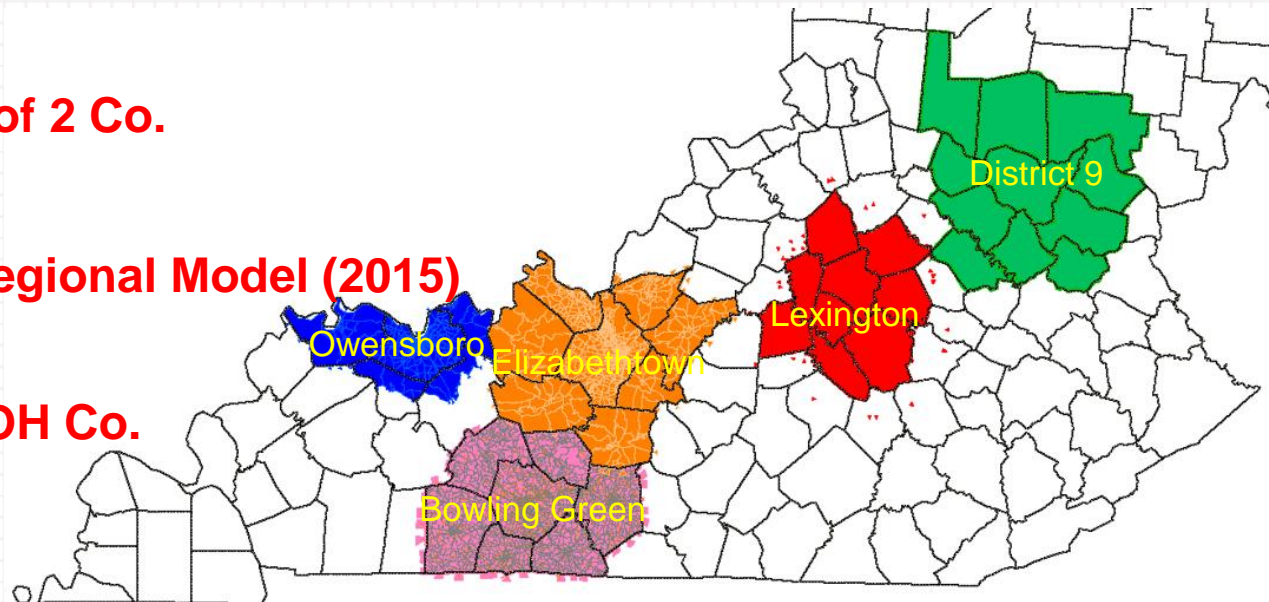
Background – AirSage Data

- No current HH OD survey data in KY. No NHTS add-on.
- Lower cost compared to traditional HH travel surveys.
- Quick turnover of data.
- KYTC first purchased AirSage OD data for Lexington Model (2012). Completed 5 regional models since 2012. All used AirSage data.
- KYTC and its consultants have gained thorough insight in the use of AirSage OD data for model development.
- AirSage OD data provides internal-internal trips within the model area by purpose and time period (more options are available).
- When data collection area is large enough, external trip information was obtained from AirSage. Good data at the edge of study area.

Background – AirSage Data

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.
 - 1 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)

Origin_Zone	Destination_Zone	Start_Date	End_Date	Aggregation	Subscriber_Class	Purpose	Time_of_Day	Count
74	286	20150415	20150514	WD	Resident	OH	H1800:H2400	12.47
315	225	20150415	20150514	WD	Visitor	OO	H0900:H1500	5.9
432	433	20150415	20150514	WD	Resident	WH	H0900:H1500	11.84

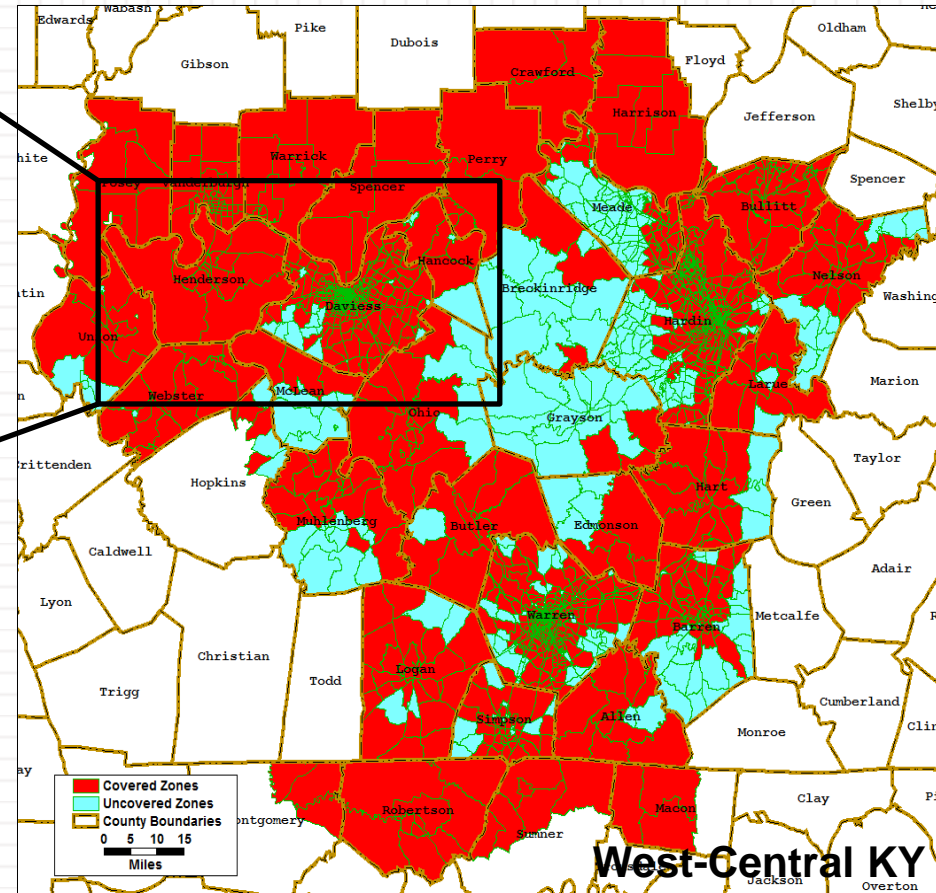
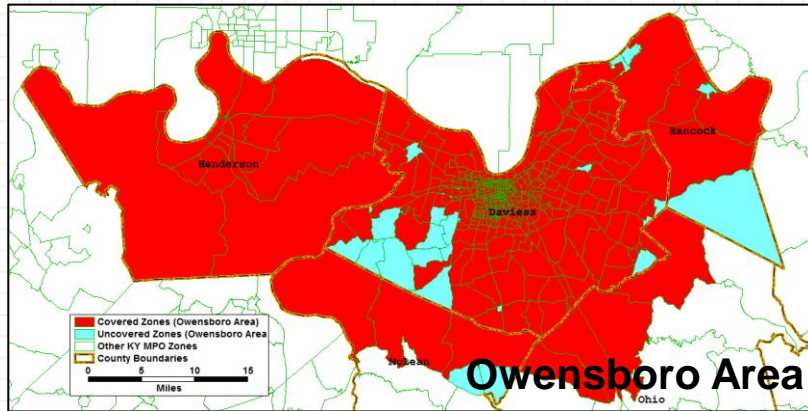
- **Must provide pre-defined study area TAZ polygons to AirSage. Consider AirSage resolution limit (0.25 square mile grid).**
- **AirSage provided “expanded” data which matches census population by carriers.**
- **When multiple carriers exist (e.g., Owensboro/Daviess Co.), averaging data sets minimizes 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)**

AirSage Data Coverage Assessment

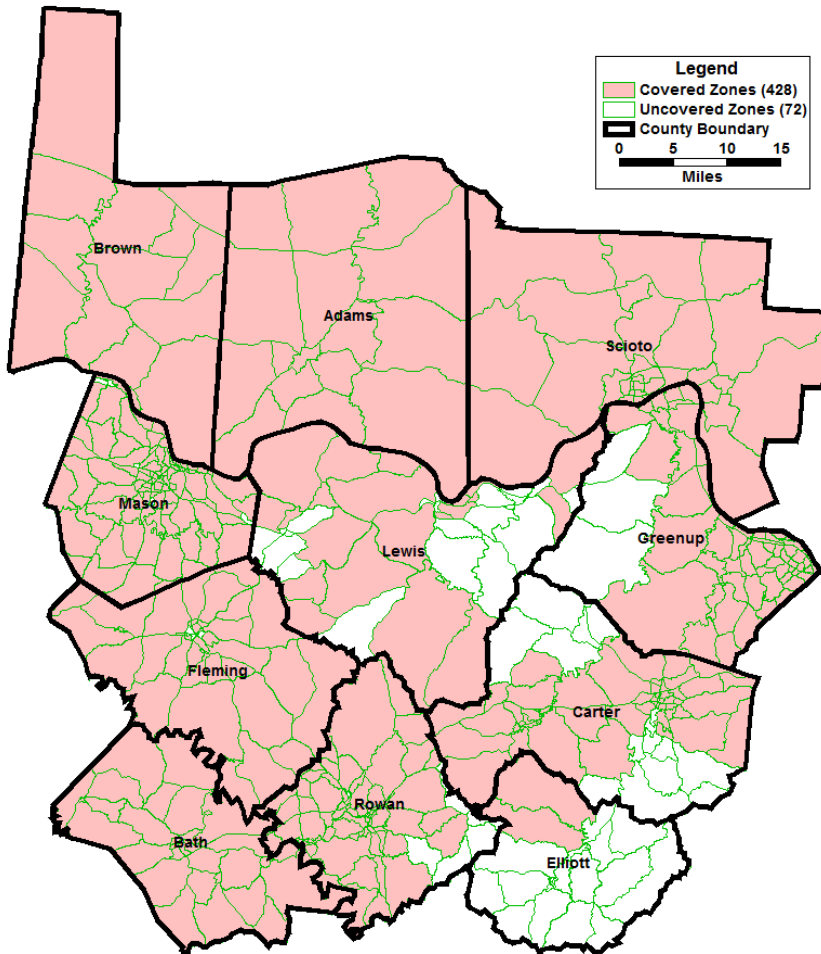
Owensboro/Daviess Co.



- Data collected for west-central KY
- 95% of 566 Owensboro zones had data
- 94% Pop & 97% Emp are covered
- Lack of data in Breckinridge, Grayson & Meade Counties (rural area with poor cellular service). 0.2% of Owensboro worker flow (CTPP). No impact on external trips.

AirSage Data Coverage Assessment

KYTC District 9 Area



- 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.

AirSage Data Adjustments (Intrazonals)

Owensboro

Purpose	Total	Intrazonal	Intrazonal %
HBW	150,012	1,951	1.3%
HBO	330,101	89,822	27.2%
NHB	95,250	29,640	31.1%
Total	575,363	121,413	21.1%

Note: Data is from raw AirSage (all trips)

District 9

Purpose	Total	Intrazonal	Intrazonal %
HBW	118,429	5,339	4.5%
HBO	654,781	324,717	49.6%
NHB	355,881	109,520	30.8%
Total	1,129,090	439,576	38.9%

Note: Data is from raw AirSage (all trips)

- **Except for HBW trips, intrazonal trip percentages seem too high**
- **HBO & NHB intrazonals 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 region (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

Purpose	AirSage		Calibrated Model	
	Internal P's	%	Internal P's	%
HBW	104,628	24%	100,543	20%
HBO	262,339	60%	262,892	52%
NHB	72,025	16%	140,013	28%
Total	438,992	100%	503,448	100%
Trip Rates	6.8		7.7	

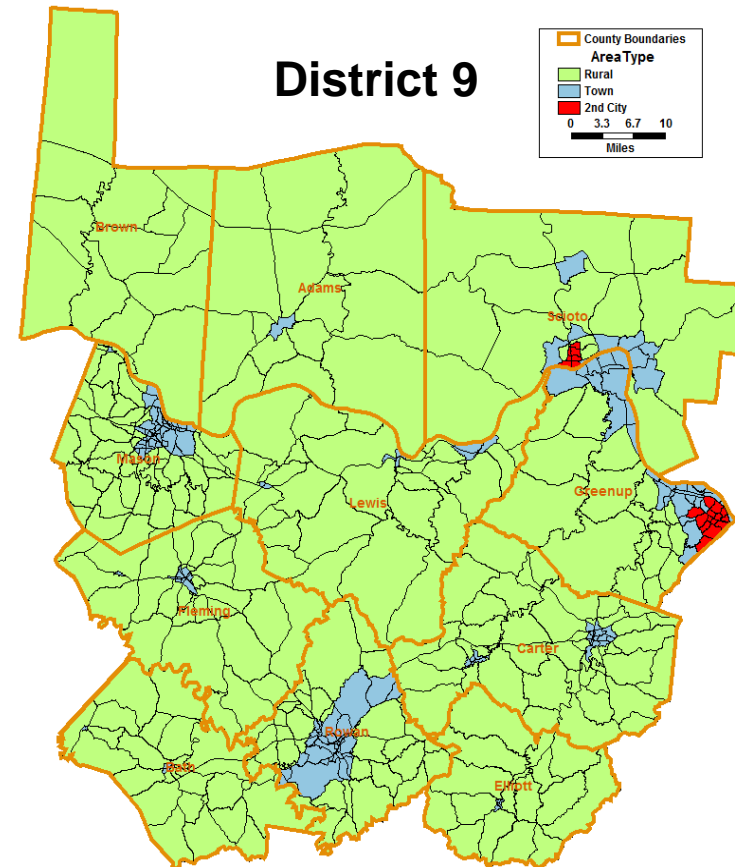
District 9

Purpose	AirSage		Calibrated Model	
	Internal P's	%	Internal P's	%
HBW	118,429	10%	160,704	18%
HBO	654,781	58%	466,274	53%
NHB	355,881	32%	258,846	29%
Total	1,129,090	100%	885,824	100%
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.

- Convert AirSage OD tables to PA tables
 - ❑ $HBW = HW + \text{Transpose}(WH)$
 - ❑ $HBO = HO + \text{Transpose}(OH) + HH$
- NHB data not used (not real P's by zonal HH)
- AirSage P's rates (by purpose)
 - ❑ zonal P's ÷ zonal HH
 - ❑ Remove abnormal values (95% CI)
- Rural, Town & 2nd City



Trip Generation - Area Type Factors

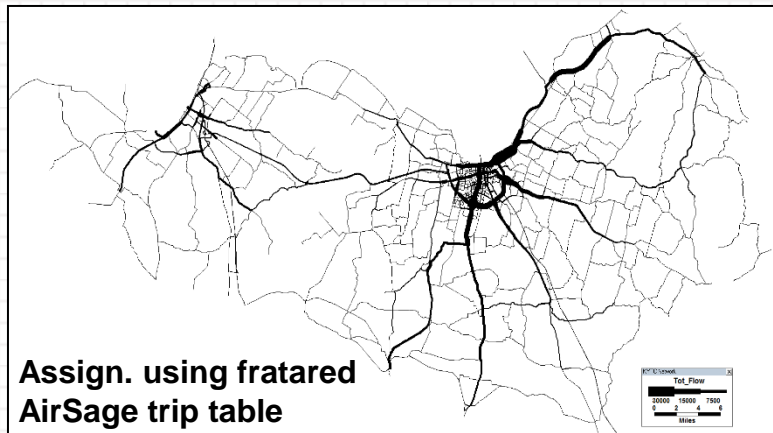
- Area Type Factors (**t-test** of P rates between area types)
 - If significantly different, factor = ratio of P rates
 - Otherwise, factor = 1.0
- Factors of Rural vs. Towns are reasonable (**HBW=1.0, HBO=0.9**)
- Factors for Rural vs. 2nd City, Towns vs. 2nd City are too large.
- Available data (KYSTM)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

Purpose	D9 Model Area Type Factor	
	Rural	Non-Rural
HBW	1.0	1.0
HBO	0.9	1.0
NHB	1.0	1.0

Trip Distribution – Friction Factors

Owensboro Method

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



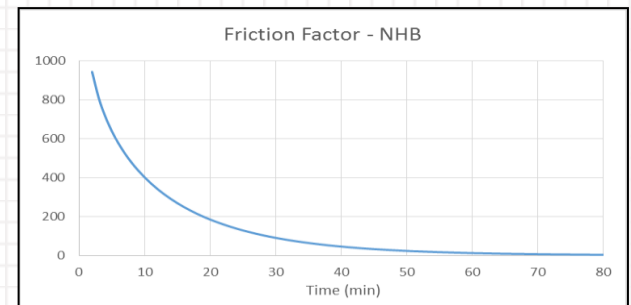
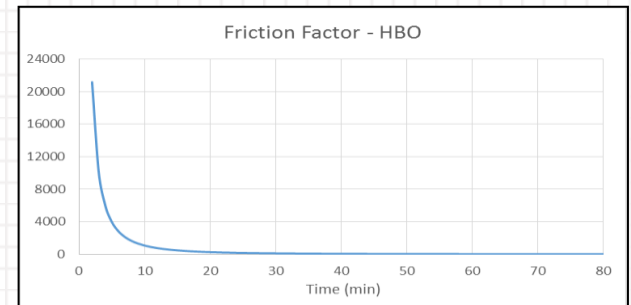
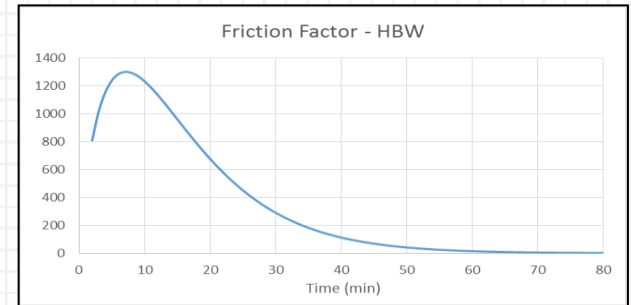
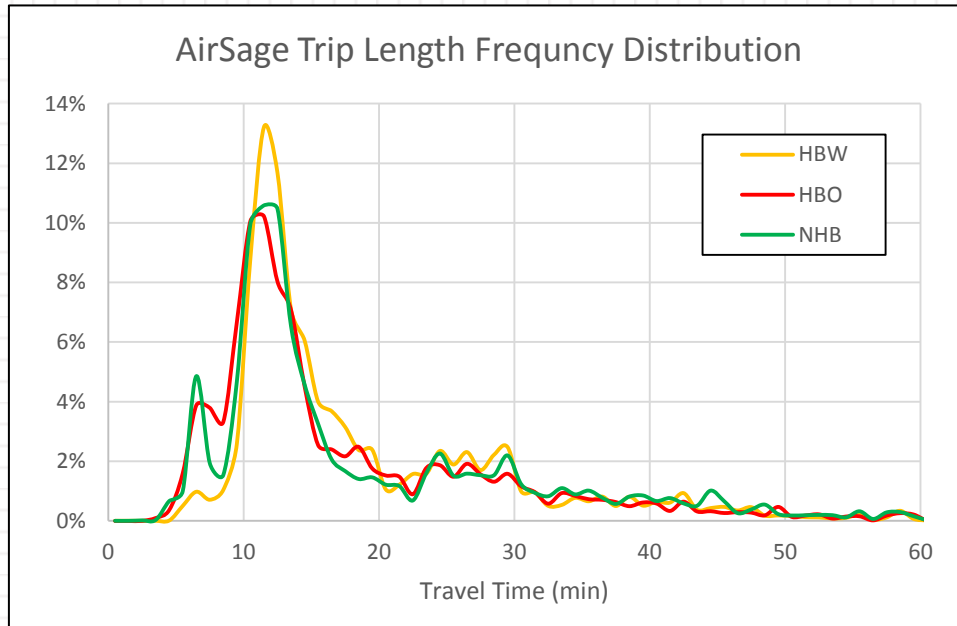
Traffic Assignment	Total Traffic (Count Links)	Avg. Dev % from Counts *
AirSage (raw)	907,259	-59.2%
AirSage (frateded)	1,397,946	-19.8%
Traffic Counts	2,275,196	

* 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



Trip Purpose	Mean Trip Length (min)	
	AirSage	Calibrated Model
HBW	19.1	15.7
HBO	17.8	14.9
NHB	19.4	14.4

Trip Distribution – Friction Factors

District 9 Method (revised approach)

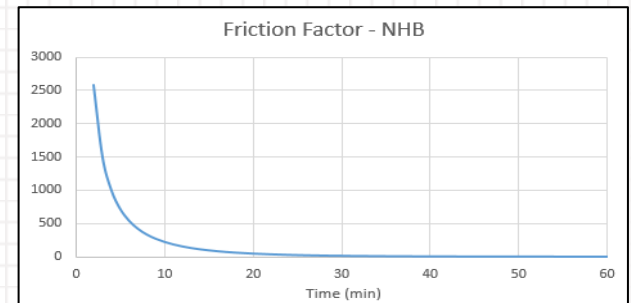
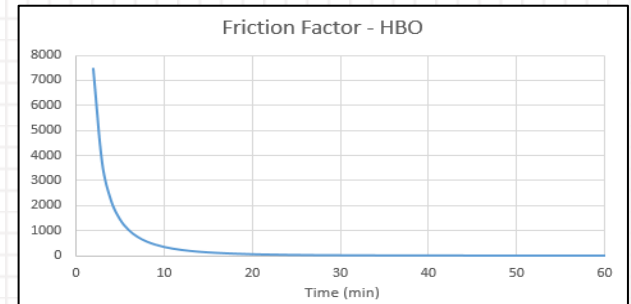
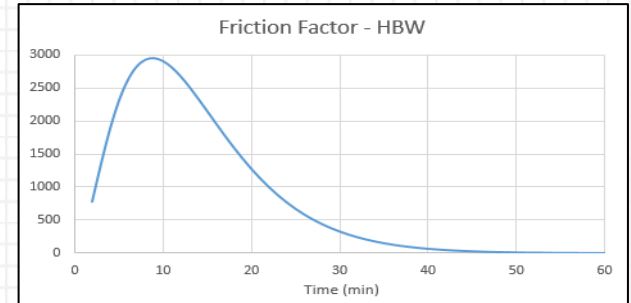
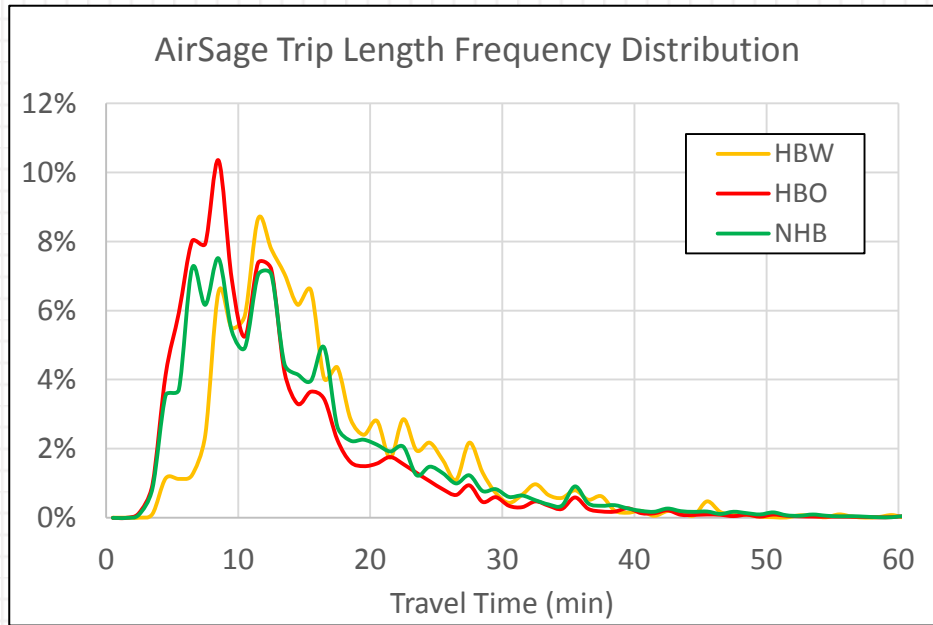
- KY local trip rates usually are unknown and differ from national defaults. Calibrate AirSage OD tables to traffic counts.
- ODME

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

- 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)



Trip Purpose	Mean Trip Length (min)	
	AirSage	Calibrated Model
HBW	16.8	17.3
HBO	13.3	12.5
NHB	15.3	13.7

AirSage TOD & Directional Factors

Owensboro (TOD)

Purpose	AM (6-9am)	Mid-Day (9am-3pm)	PM (3-6pm)	Night (6pm-6am)
HBW	0.26	0.25	0.18	0.31
HBO	0.17	0.30	0.19	0.34
NHB	0.12	0.47	0.23	0.18
E-I Auto	0.20	0.30	0.20	0.30
E-E *	0.19	0.31	0.20	0.30

* Assumed by average of all purposes

District 9 (TOD)

Purpose	AM (6-9am)	Mid-Day (9am-3pm)	PM (3-6pm)	Night (6pm-6am)
HBW	0.28	0.27	0.19	0.26
HBO	0.17	0.30	0.19	0.34
NHB	0.19	0.37	0.22	0.22
E-E *	0.19	0.32	0.20	0.29

* Assumed by average of all purposes

Owensboro (direction – P to A)

Purpose	AM (6-9am)	Mid-day (9am-3pm)	PM (3-6pm)	Night (6pm-6am)
HBW	0.91	0.51	0.16	0.43
HBO	0.85	0.53	0.43	0.33
E-I Auto *	0.85	0.53	0.43	0.33

* Assume equal to HBO

District 9 (direction – P to A)

Purpose	AM (6-9am)	Mid-day (9am-3pm)	PM (3-6pm)	Night (6pm-6am)
HBW	0.94	0.49	0.16	0.39
HBO	0.85	0.54	0.42	0.34
E-I Auto *	0.85	0.54	0.42	0.34

* 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.

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.

▪ Disadvantages – **adjustments required**

- Everything is aggregate.
- Characteristics of the traveler are not available.
- Trip purpose is based on apparent land use and times of data transmissions.
- Data transmission times may not accurately reflect when travel occurs. Estimates of travel patterns may be limited.
- 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

