Experience Using Air Sage Data for Model Development in Kentucky

presented to
The Tennessee Model Users Group

presented by
The Corradino Group

Mt. Juliet, TN
February 5, 2015

THE CORRADINO GROUP
Engineers • Planners • Program Managers • Environmental Scientists

Objectives

- Recount the Kentucky experience with using Air Sage data,
- Describe:
  - apparent advantages,
  - disadvantages, and
  - limitations of the data

Background

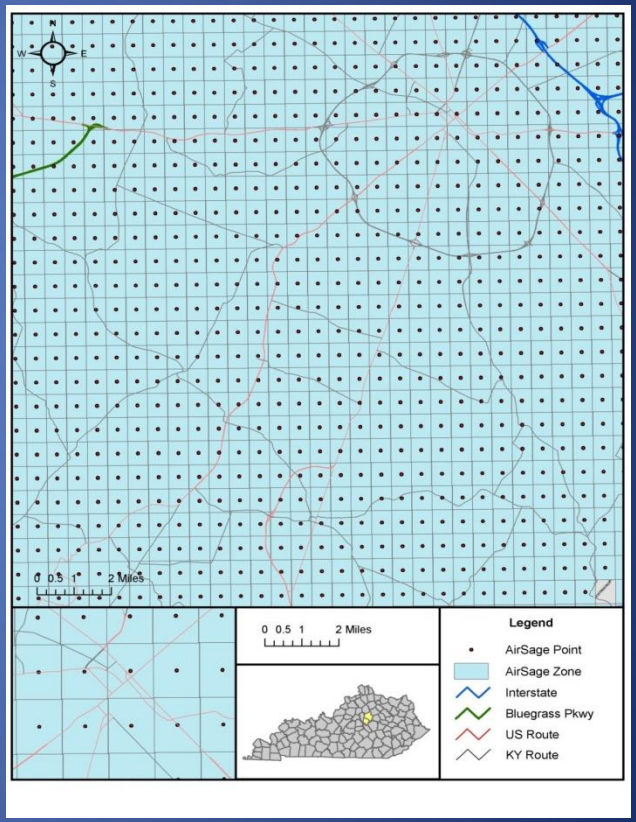
- No current HH OD survey data. No NHTS add-on.
- The Kentucky Transportation Cabinet (KYTC) has purchased Air Sage data for some regions.
- Corradino used Air Sage data (2012) in the development of an eight-county model in the Lexington/Central Kentucky region.
- Corradino is now developing a new model for the Owensboro/Henderson Region (2014), which covers three entire counties and parts of two other counties.

The Air Sage Product

- Created TransCAD matrix files from Air Sage (.CSV).
- Time Periods: (AM peak, midday, PM peak, and night),
- Trip Purposes: (H-home, W=work, O=other)
  - HBW <– HW, WH
  - HBO <– HO, OH, HH
  - NHB <– WO, OW, WW, OO

The Air Sage Spatial Coding

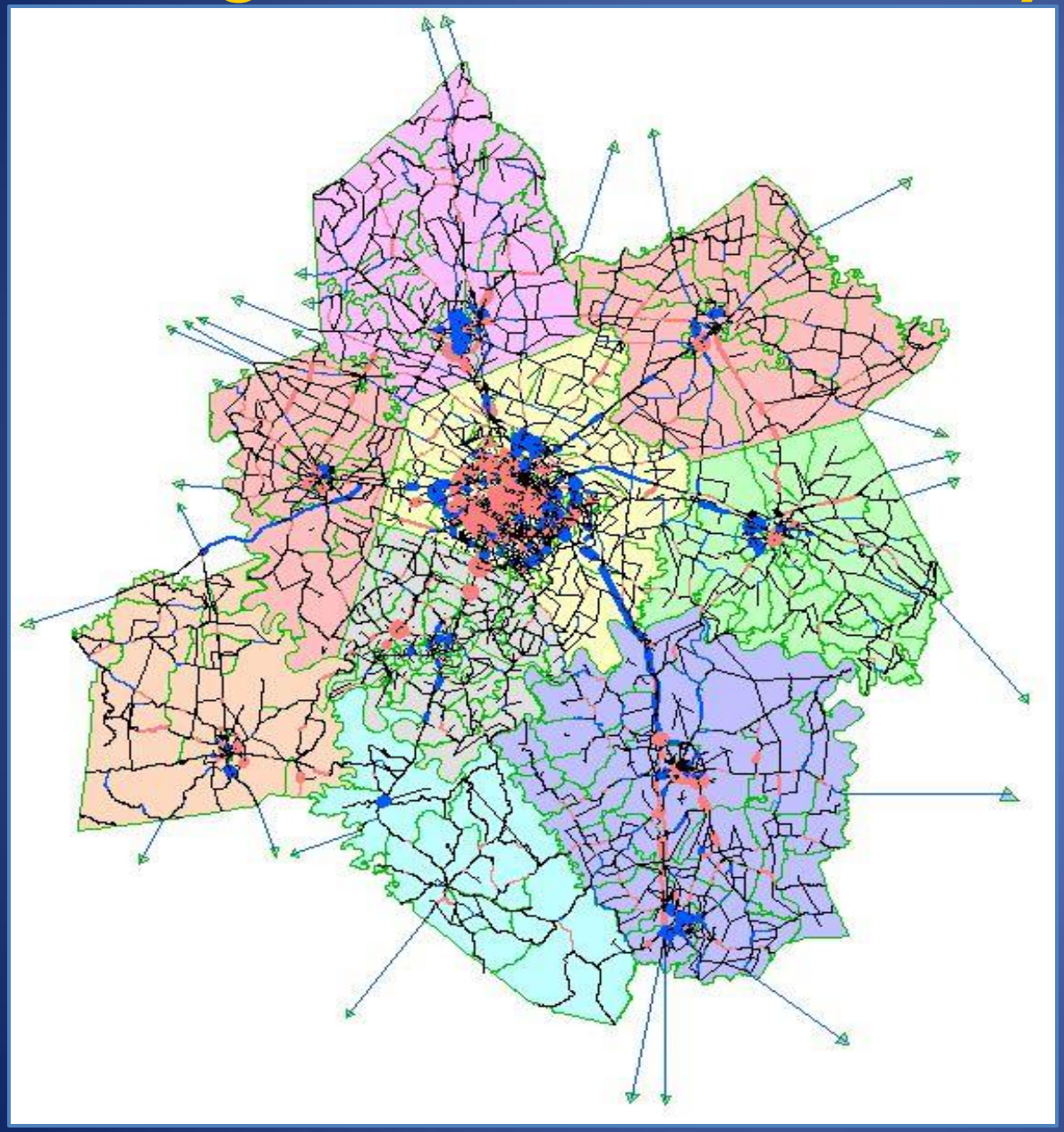
- Air Sage determined locations from cell phone towers, using triangulation and the directional antennas. Not GPS-based.
- For Lexington, Air Sage provided data located on a 1,000 x 1,000 meter grid, which were associated with TAZ polygons.
- For Owensboro, Air Sage determined the best TAZ from the model's TAZ shape files.
- Because of privacy issues, addresses of cell phone owners (home-address) could not be provided.



The Air Sage Product

- Air Sage uses proprietary methods for determining trip purposes.
- Methods depend on an assessment of the land use where the cell phones are detected, and for how long.
- For a cell phone to be detected, data must flow by phone call, text message, web access or other data transmission. Locations are not recorded from the cell phone “pinging” a cell phone tower.
- Air Sage data generally cannot be used to estimate travel time because the time of arrival and departure at a location is somewhat uncertain.
- Air Sage data does not contain any disaggregate user information.

Lexington/Central Kentucky Study Area



Lexington/Central Kentucky Zonal Data

Category	Source	Fayette County	Jessamine County	Bourbon County	Clark County	Garrard County	Madison County	Mercer County	Scott County	Woodford County
Population	Census Bureau[1]	295,803	48,586	19,985	35,613	16,912	82,916	21,331	47,173	24,939
Occupied Housing Units	Census Bureau[1]	123,043	17,642	7,976	14,267	6,668	31,973	8,682	17,408	9,806
Retail Employment	ES-202[2] /KYSTM	84,578	8,275	3,250	5,629	1,146	18,127	2,110	7,273	4,902
NonRetail Employment	ES-202[2] /KYSTM	20,976	4,139	2,351	2,869	265	5,833	4,270	9,454	2,925
Service Employment	ES-202[2] /KYSTM	55,107	3,384	1,304	3,853	529	7,131	3,927	5,006	2,179
Total Employment	ES-202[2] /KYSTM	160,661	15,804	6,905	12,351	1,940	31,091	10,307	21,733	10,006

[1] 2010 US Census Bureau Data

[2] 2012 ES-202 data was original source of employment data and was processed by KYTC staff.

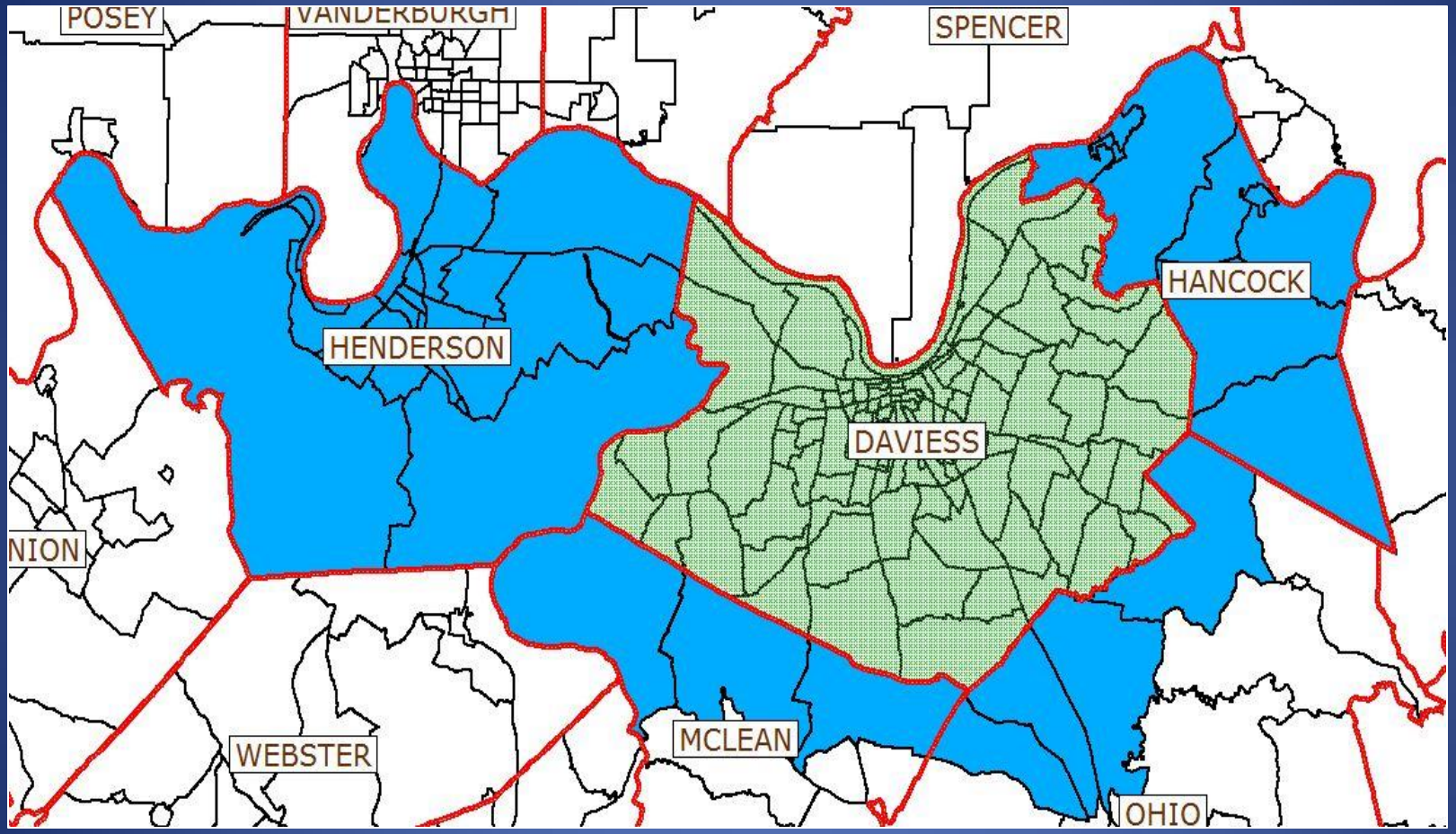
Lexington Method

- Trip generation: linear regression. Analysis limited because no disaggregate user data are available – only aggregate TAZ data. Zonal average income, households, and total employment were used.
- Trip distribution: used network times/skims, and Air Sage trip tables to estimate friction factors.
- Time-of-day/diurnal factors were developed from the Air Sage trip tables.

Lexington Limitations and Findings

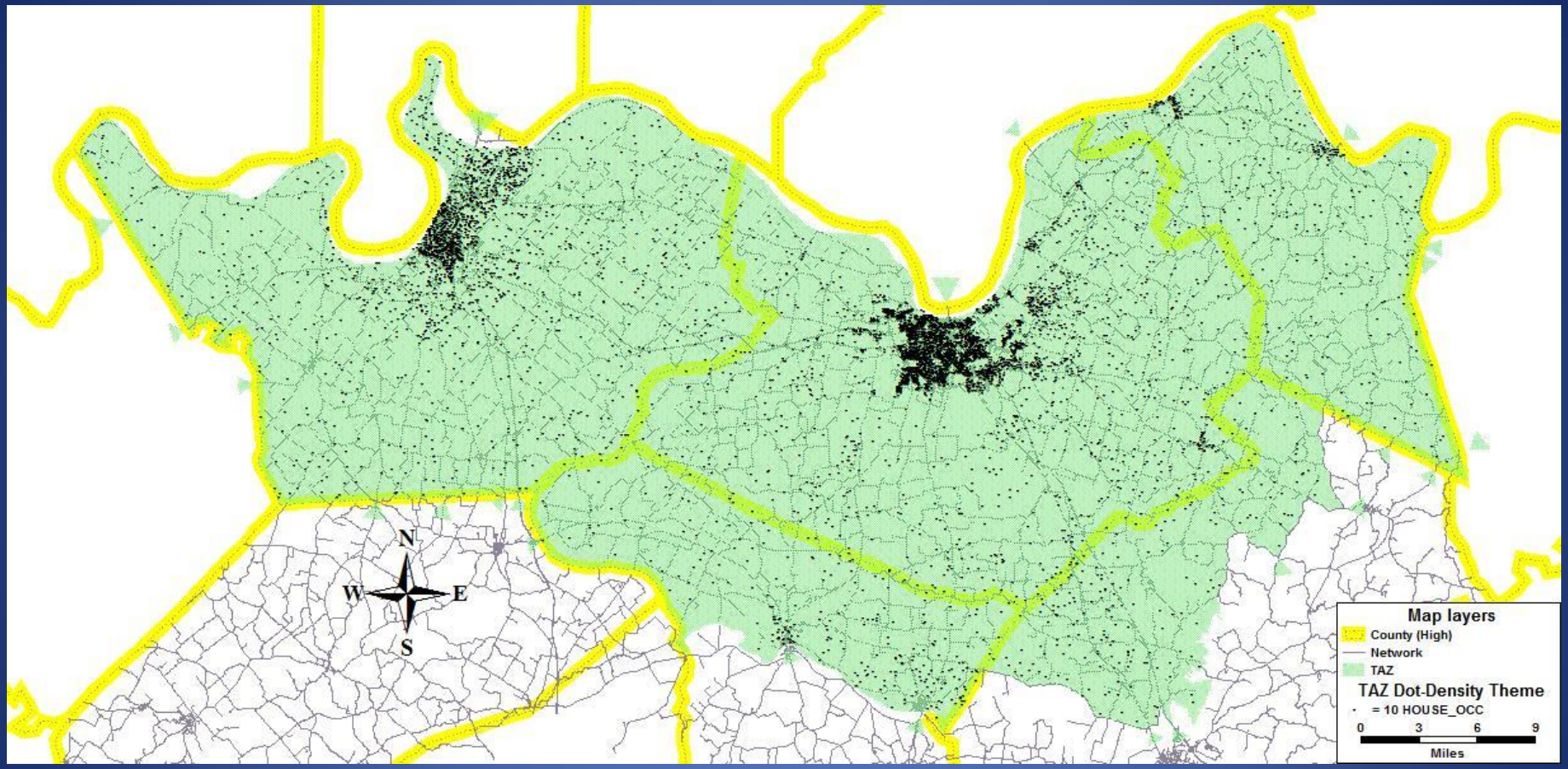
- Results similar to “synthesized” models developed with little data other than traffic counts.
  - 8-county daily %RMSE ~ 49%, similar to statewide models
  - Fayette/Jessamine (urban portion) % RMSE ~ 40%
- Trip generation too simplistic – most zonal data are collinear, and since no disaggregate user data are available, classification analyses could not be done.
- Data shed little light on travel characteristics for major colleges and universities (University of Kentucky, Eastern Kentucky University, Georgetown College, and Asbury College).
- Air Sage contains no data on modes and auto occupancy.
- Some unusual results – HBW was not the purpose with the longest trips (before calibration adjustments).

Owensboro Study Area



- All of Daviess, Henderson and Hancock Counties
- Parts of McLean and Ohio Counties

Owensboro Region Households



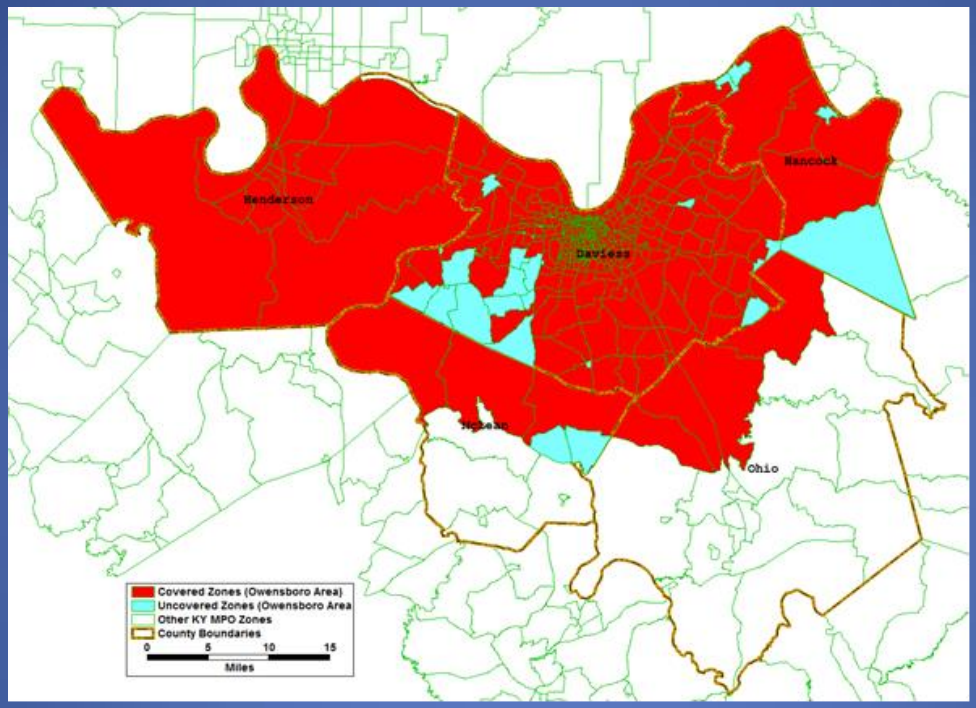
Summary of 2010-2040 Population

Population Estimates			
	Census 2010	2040	Growth
Kentucky	4,339,367	5,162,292	822,925
Green River ADD	213,472	223,771	10,299
Study Area Counties			
Daviess	96,656	108,317	11,661
Hancock	8,565	8,980	415
Henderson	46,250	47,576	1,326
McLean*	9,531	8,332	-1,199
Ohio*	23,842	25,611	1,769
*=partial			

Revised Approach

- No compelling reason to believe that trip generation rates should vary widely between regions. They are generally transferrable.
- Use NCHRP 716 and 365 to estimate trip generation.
- But, time-of-day factors and trip distribution models are much less transferrable.
- Adjust Air Sage trip tables to match NCHRP 365-based P's and A's before estimating friction factors.
- Use Air Sage trip tables to estimate trip distribution and time-of-day models.

Owensboro Air Sage Data Adjustments (two carriers)



- Two carriers covered most of the region. Only very rural, sparsely populated areas were not covered.
- The records for each carrier were expanded independently.
- The values were averaged when coverage was duplicated.

Owensboro Air Sage Data Adjustments (intrazonals)

Purpose	Total Trips	Intrazonal Trips	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%

- Except for HBW trips, intrazonal trip percentages seemed too big.
- HBO and NHB intrazonals were reduced to more conventional levels before developing friction factors.

Owensboro Air Sage Data Adjustments (trip rates)

Purpose	NCHRP Equation (balanced)		AirSage (expanded)	
	I-I Prod.	I-I Attr.	I-I Prod.	I-I Attr.
HBW	79,177	79,177	104,628	104,628
HBO	294,198	294,198	262,339	262,339
NHB	178,959	178,959	72,025	72,025
Total	552,333	552,333	438,992	438,992
Daily I-I Person Trip/HH	8.5		6.8	

- Trip rates seem too low – not unexpected.
- Not used for trip generation.
- Trip tables adjusted before using them for analysis of trip distribution.

Owensboro Air Sage Data Adjustment (trip assignment test)

Traffic Assignment	Total Volume on Counted Links	Average Dev % from Counts
Air Sage Expanded	907,259	-59.2%
Fratared AirSage Expanded	1,397,946	-19.8%
Fratared AirSage Unexpanded	1,374,049	-21.5%
Traffic Counts	2,275,196	

- Used in the assessment of possible methods to adjust Air Sage before use in trip distribution analysis.
- Data as delivered from Air Sage had the biggest error (simple average error, not RMSE).
- Smallest error was to adjust Air Sage expanded trip tables to match NCHRP 365 P's and A's.

Time-of-Day Factors

Purpose	AM Peak	Midday	PM Peak	Night	Daily
HBW	0.2572	0.2493	0.1821	0.3114	1.0000
HBO	0.1736	0.3000	0.1893	0.3371	1.0000
NHB	0.1233	0.4681	0.2268	0.1817	1.0000
E-I Auto	0.1956	0.3006	0.2019	0.3018	1.0000
E-E	0.1883	0.3101	0.1951	0.3064	1.0000

Purpose	AM Peak P to A	Midday P to A	PM Peak P to A	Night P to A
HBW	0.9082	0.5088	0.1579	0.4348
HBO	0.8533	0.5262	0.4337	0.3315
E-I Auto	0.8533	0.5262	0.4337	0.3315

Air Sage time of day and directional factors seem reasonable. These data are based on adjusted Air Sage trip tables.

Owensboro Model Results

SUMMARY METRICS Daily	
Name	Value
COUNT VMT =	1,475,746
FLOW VMT =	1,600,220
FLOW VMT/COUNT VMT =	1.084
%RSME =	35.58
FLOW VMT (all links) =	4,507,018

RMSE BY VOLUME GROUP		
Count Range	% RMSE	Desired Range
0-2000	86.52	55 plus
2000-5000	34.52	45 - 55
5000-10000	29	35 - 45
10000-20000	22.85	27 - 35
20000-30000	12.65	24 - 27
30000-40000	3.26	22 - 24

RMSE BY FACILITY TYPE					
Facility Type	% RMSE	VMT Ratio	Count	Count Ratio	#Links
Other fwy xway (2)	20.72	1.155	251,397	1.164	23
Other Principal arterial (3)	22.98	1.101	632,151	1.012	68
Minor arterial (4)	29.52	1.009	721,888	0.96	97
Major collector (5)	43.98	1.089	480,076	0.946	130
Minor collector (6)	82.29	1.11	80,616	1.207	96
Local (7)	106.02	0.587	31,466	0.639	59

RMSE BY County		
County	% RMSE	Count Ratio
DAVIESS	34.72	1.026
HANCOCK	44.61	1.067
HENDERSON	33.27	0.932
MCLEAN	25.11	1.008
OHIO	29.63	1.125

Air Sage Advantages/Disadvantages

- Advantages
  - Very large dataset.
  - Low cost when compared to surveys.
  - Large coverages 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 trip times and travel times may be limited.
  - Unusual things in the trip tables (IZ, trip lengths, trip rates)

Summary of Findings

- Findings are based on our understanding of the Air Sage product. It is clear that Air Sage is refining their product, so future data may address these issues.
- Air Sage data may not be suitable for determining trip generation rates. Analyses are very limited because of the aggregate data.
- Air Sage data should be adjusted before use in trip distribution analysis.
- Air Sage seems to provide reasonable estimates of existing temporal distribution.
- Air Sage provides no information on the user, or any other disaggregate observations because of privacy restrictions.

Acknowledgements

- Kentucky Transportation Cabinet (KYTC)
  - Scott Thomson, PE
- Consultant partners
  - The Corradino Group
  - Parson Brinckerhoff
  - Stantec

THE CORRADINO GROUP
Engineers • Planners • Program Managers • Environmental Scientists