RSG
the science of insight

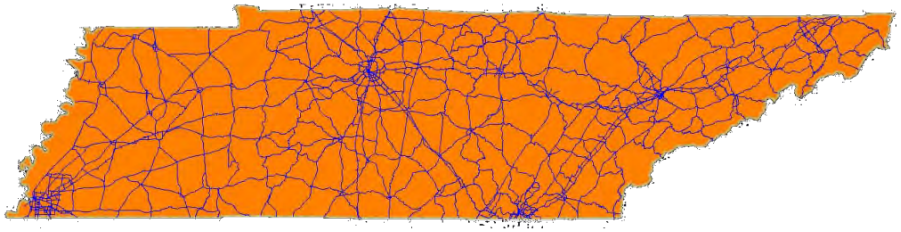
Tennessee Statewide Travel Forecasting Model Update

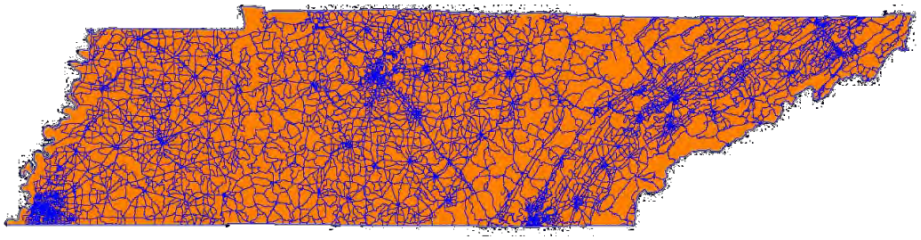
Vince Bernardin, RSG

May 15, 2014

Phase 1 of the Statewide Model Update: Data Development

Statewide Model Update





Existing Statewide Model

- 2003 Base – 2030 Horizon
- Only Total Daily Traffic
- Limited Network Coverage
- Limited Sensitivity
  - Re-routing Only

New Statewide Model

- 2010 Base Year – 2040 Horizon Year
- Peak Hour and Daily Traffic
- Expanded Network Coverage
- New Sensitivity to:
  - Network changes
  - Induced demand
  - Alternative future land use scenarios
  - Population changes (aging, etc.)
- Version 3 – Commodity Flow Modeling

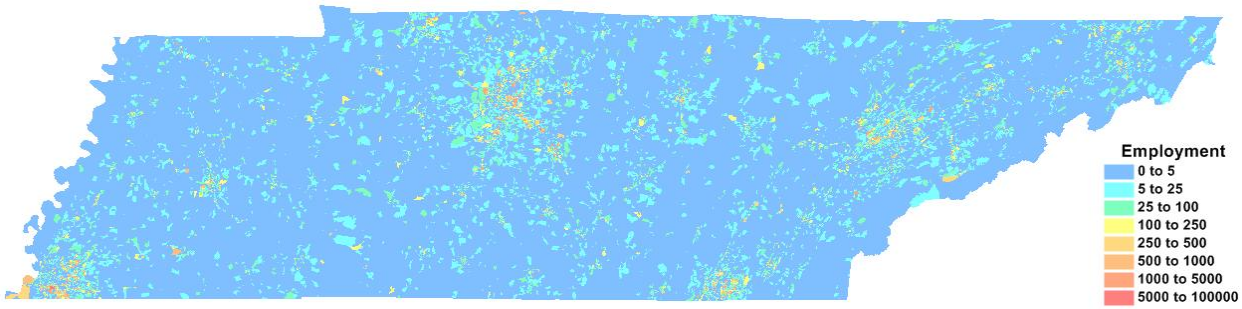
Statewide Model Update: Phases 1 & 2

Phase 1: Data Development (completed in April)

- New, Expanded Network
- New, More Detailed Zone System
- Obtain & Process Socioeconomic Data
- New Socioeconomic Forecasts
- Obtain & Process ATRI Truck GPS Data
- Combine NHTS & MPO Household Travel Survey Data

Phase 2: Model Development (begun in February)

- New Trip-based Model
- Time-of-Day Modeling (peak hour volumes)
- Destination Choice Models (greater accuracy)
- Possible Pivot-Point Structure (greater accuracy)
- Truck/Freight Modeling still being scoped
- Post-processing for Performance Measures (access to jobs, hospitals, etc.)



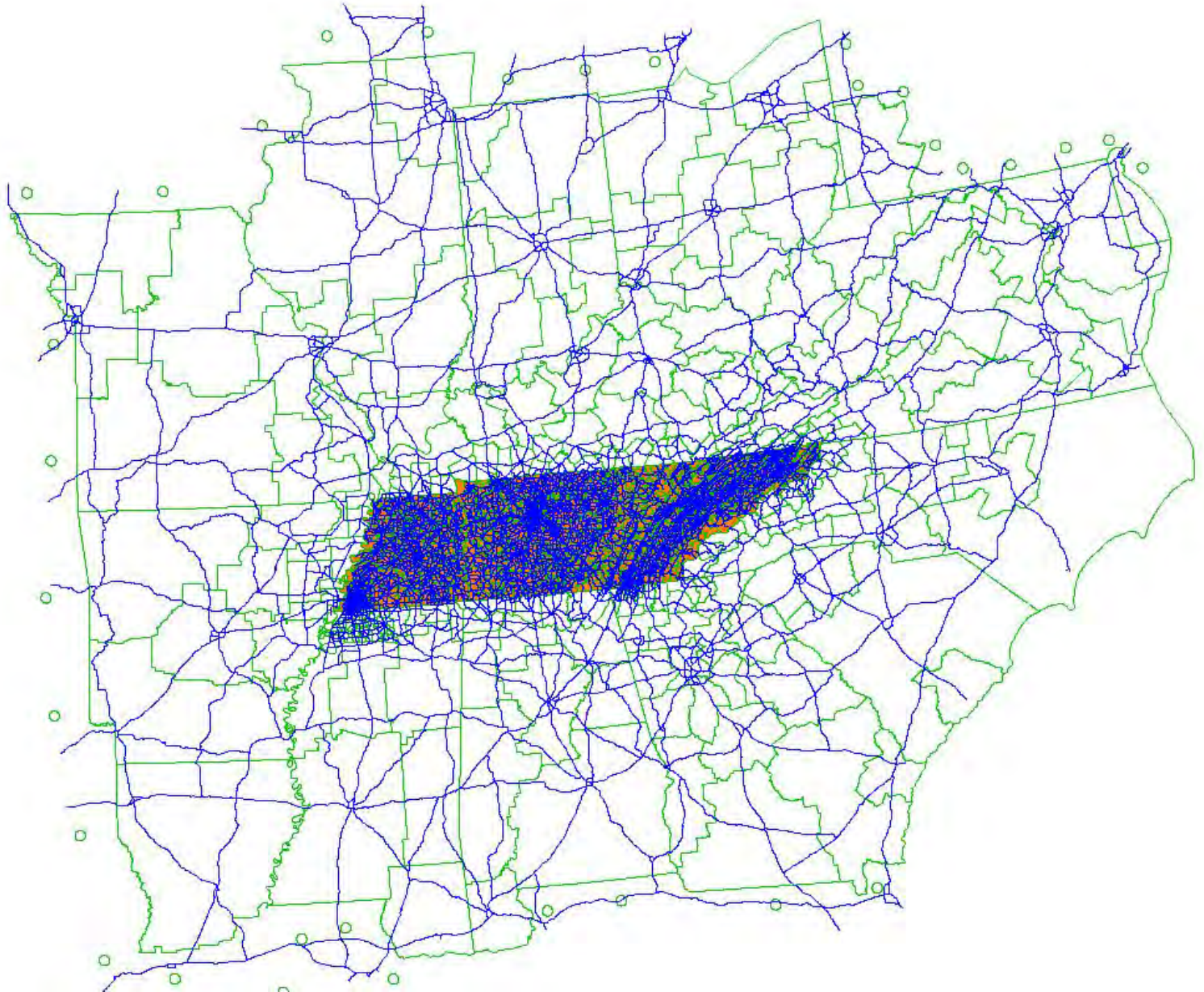
Zone Size and Network Coverage

	Ohio	Iowa	Indiana	Tennesse v1	Tennessee v2
Population	11,500,000	3,100,000	6,500,000	6,500,000	6,500,000
Road Miles*	42,000	45,000	19,000	9,421	32,546
TAZ in state	3,660	1,866	4,690	1,222	3,293
Total TAZ	5,116	3,314	4,831	1,397	3,684
Pop / TAZ*	3,200	1,600	1,400	5,300	2,000
Acres / TAZ*	12.2	30.2	7.8	34.5	12.8
Miles / Acre	0.9	0.8	0.5	0.2	0.8
Pop / Miles	270	70	340	690	200
Miles / TAZ	11.5	24.1	4.1	7.7	9.9

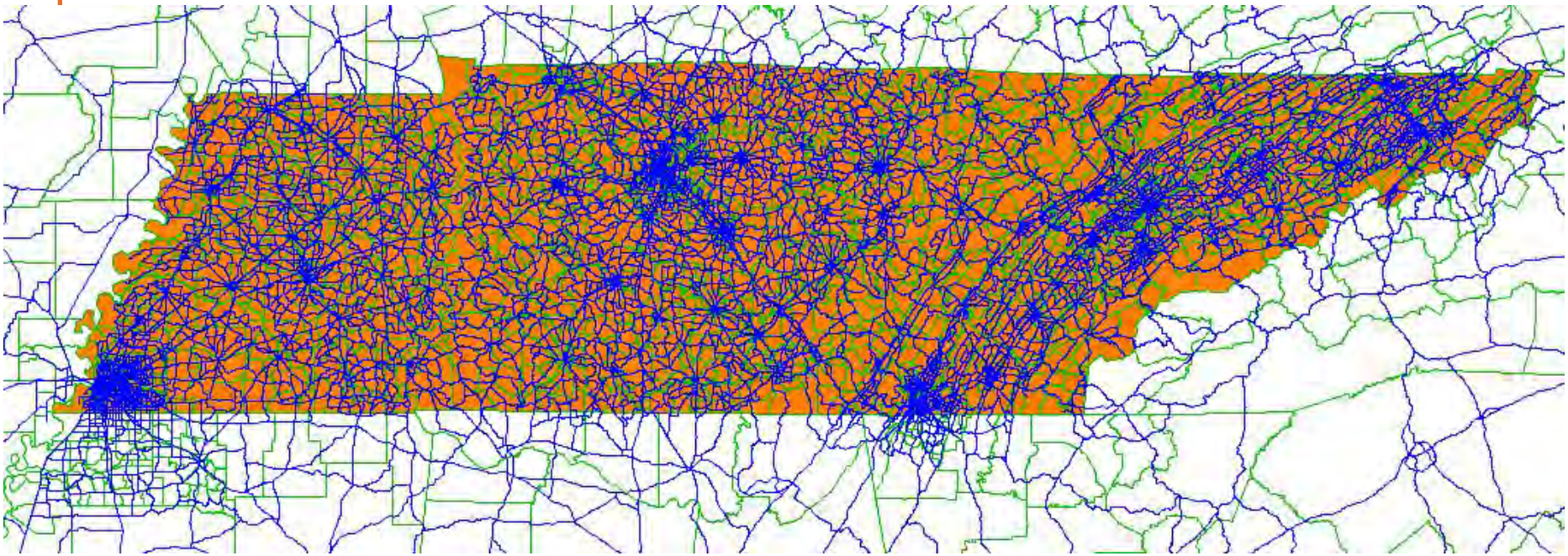
*in state

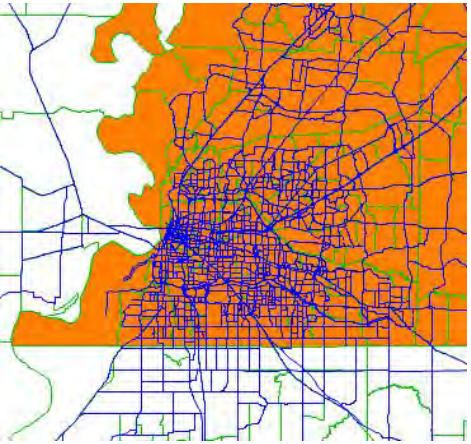
- New model has **triple** the network and zones

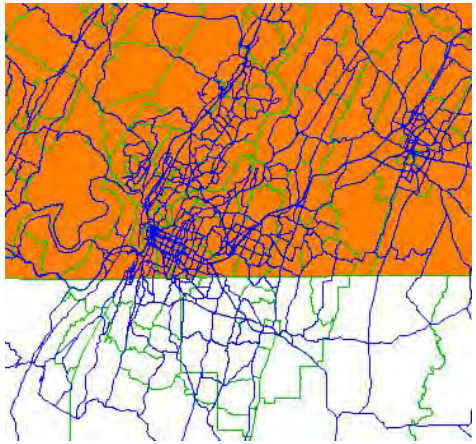
Version 2 Network and Zones

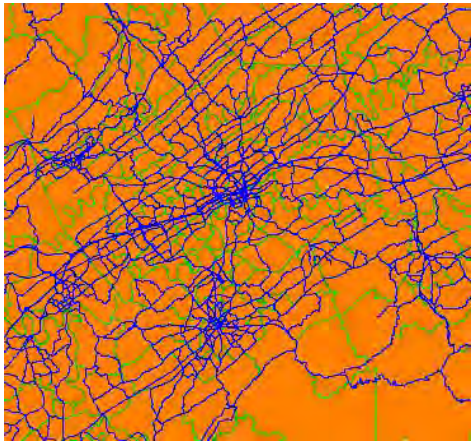


Version 2 Network and Zones









Network Development

Defining the Network

How do we decide what to include in the network?

Old Model

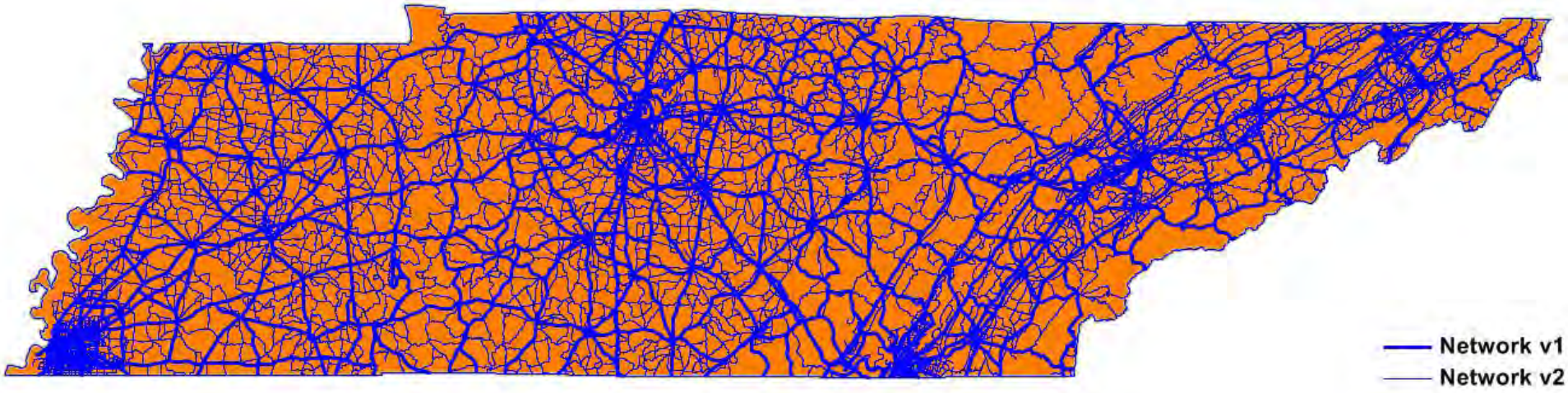
- Interstates & Principal Arterials

New Model

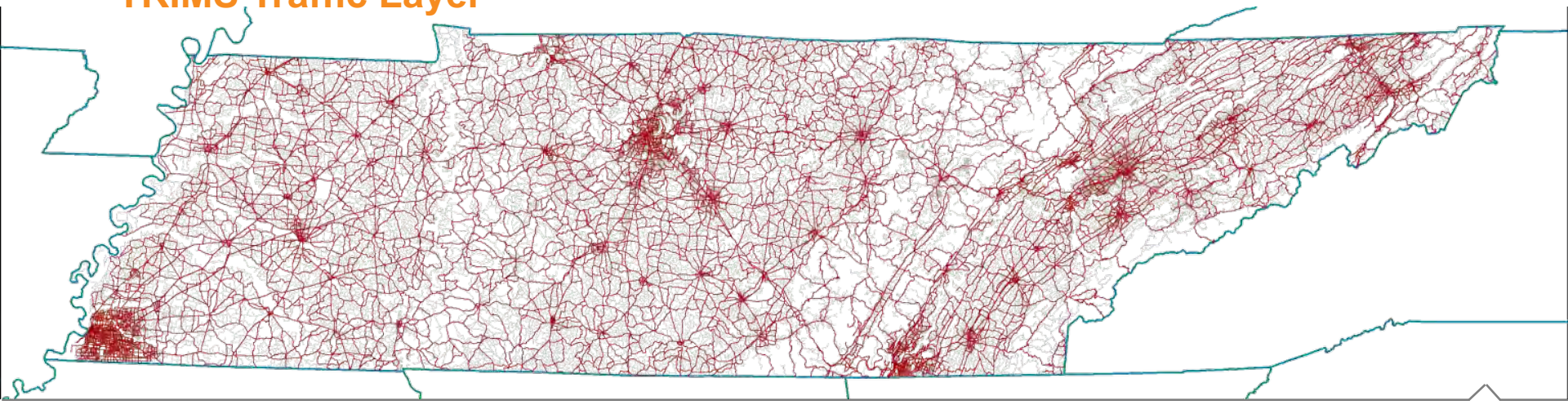
- Started with minimum criteria
  - Anything in the old model
  - Anything in the National Highway Planning Network (NHS, etc.)
  - All minor arterials
- Want network coverage one class lower than desired forecasts
- Began to look at TRIMS data, to consider volume thresholds, etc.
- Found the TRIMS Traffic layer (e.g., roads with TDOT traffic counts) has roughly the right level of network coverage
- Had to add ~100 links to minimum criteria
- May still add/remove a small number of roads in Phase 2 to ensure good loadings / balance with TAZ layer

Defining the Network

New vs. Old Network



TRIMS Traffic Layer



Network Topology: Connectivity & Routing

Model Requirements

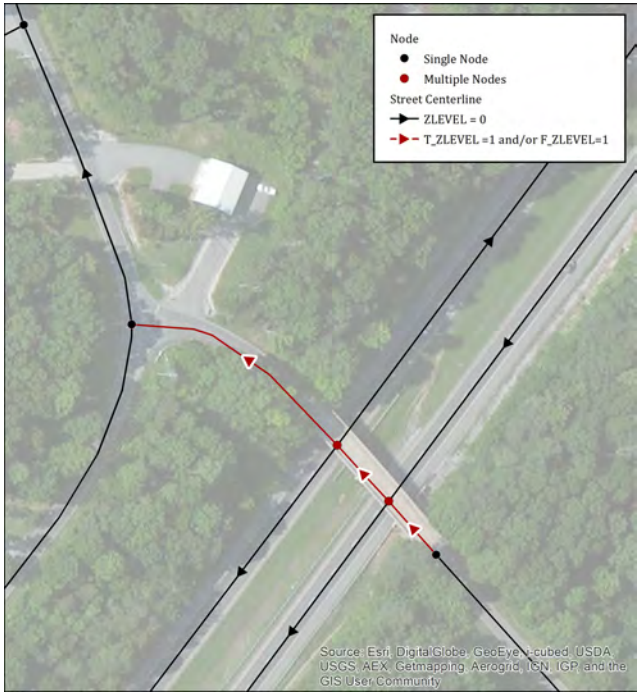
- Connected and routable network

Options

- Connect TRIMS GIS layers – Not enough time & budget
- TN OIR E99 layer – Not ready until 2014
- TeleAtlas network – Chosen

Issues

- How to connect with TRIMS
- TRIMS missing ramps
- Elevation (Z) data



Network Attributes

Chosen Attributes

- Design
  - Direction
  - Divided
  - Access Control
  - Ramp
  - Lanes
  - TWTL
  - Reversible lane
  - Lane width
  - Shoulder width
  - Terrain
  - Water Xing
  - RR Xing
  - Speed Limit
- Administrative
  - Name
  - Functional Class
  - Ownership
  - County
  - State
- Intersection
  - Control Type
- Traffic
  - AADT
  - Peak hour %
  - MU Truck %
  - SU Truck %

Attribute	Missing
Speed Limit	15.94%
Divided	0.01%
Lanes	0.03%
AADT	0.05%

Missing Attributes

- Most attributes substantially complete from TRIMS
- 90% of roads missing speeds were rural minor arterials
- Speed was missing on roughly half of this class
- No volume / geographic pattern – ok to impute

Network Development Process

Getting the TRIMS attributes on the routable TeleAtlas network…

- Develop a Least-Common-Denominator (LCD) TRIMS line layer network with a nodes anywhere a chosen attribute changes
- Get all the TRIMS attributes onto the single LCD layer
- Simplify and reduce the LCD representation above if possible
- Develop a common segmentation between LCD TRIMS and TeleAtlas
- Pass the TRIMS attributes over onto the newly segmented TeleAtlas layer
- Simplify the newly segmented TeleAtlas layer (remove unnecessary nodes)

TAZ Development

A Good Lookin' TAZ

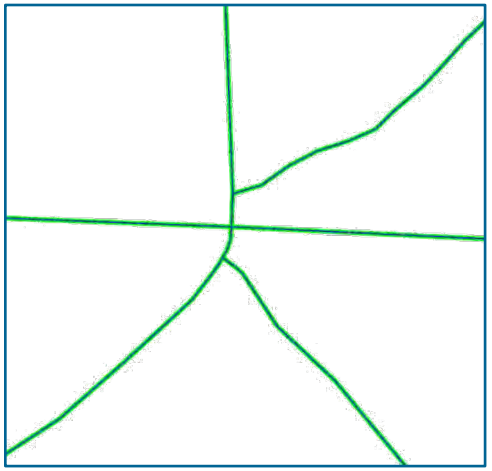
What should a TAZ look like? What makes for a good TAZ?

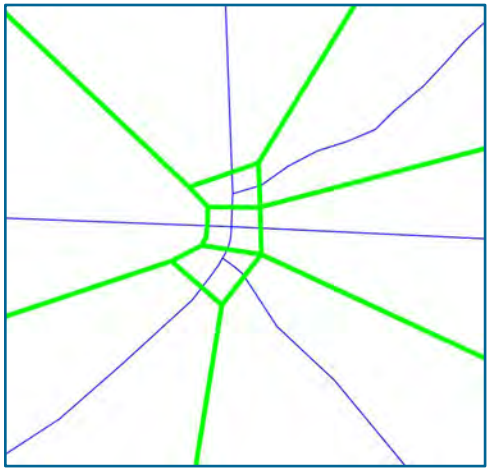
Traditionally

- Zone boundaries conform to the network
- And other boundaries, maybe
- And homogenous land use, maybe

Travel Sheds

- Zones as catchment areas around network
- Borrowed from hydrology
- First used for TAZ in NW 20+ years ago
- Increasingly common in statewide models
- Clearer relationship to the network, less ambiguity about loading points / centroid connectors
- Better able to represent distinct rural and small urban zones
- Take other boundaries more seriously





Building Blocks

What are the statewide TAZ made of?

Urban

- Aggregations of MPO zones (except conform to 2010 Census geography)
- Mix of traditional and travel sheds

Rural / Small Town

- Aggregations of Census blocks
- Less traditional, mostly travel sheds

Putting the Puzzle Pieces Together

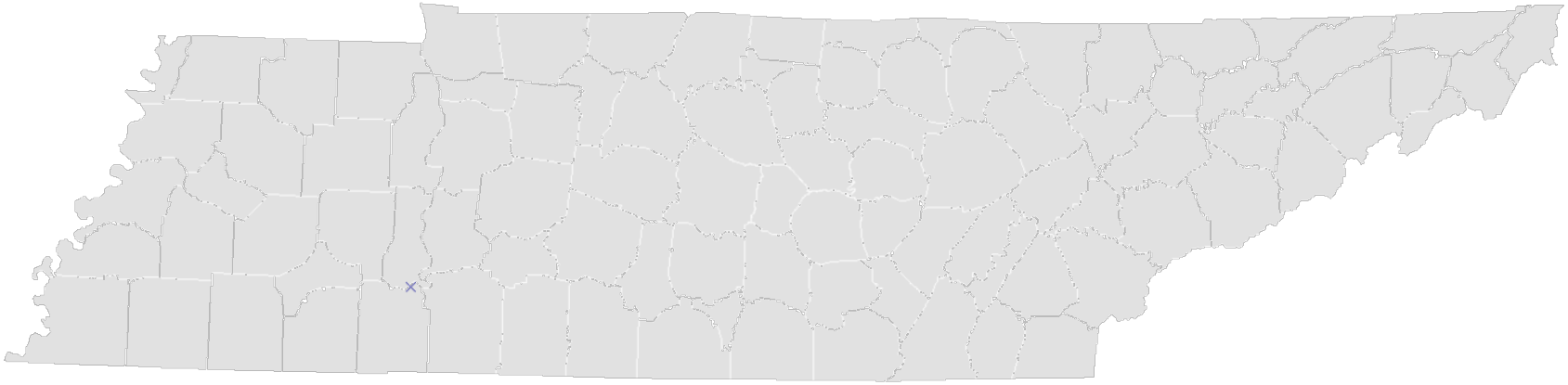
How do you group MPO zones / Census blocks into SWM TAZ?

Two step process

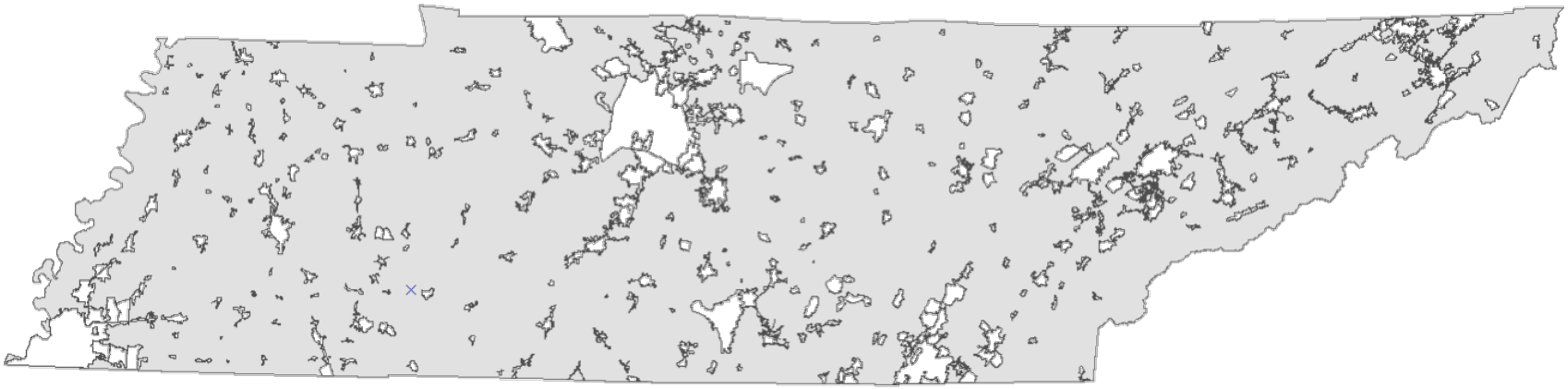
- First identify boundaries TAZ should not cross
  - County boundaries
  - Place boundaries (loose)
  - Major Rivers
  - Freeways
  - Railroads
  - Major ridgelines / slopes
- Then within the areas defined by these boundaries, group building blocks (blocks/MPO TAZ) into travel sheds around largest (non-freeway) facilities
  - Estimate number of desired zones and choose corresponding number of corridors, starting with highest AADT
  - Successively buffer around each corridor in increments (0.5 mi) until all blocks/MPO TAZ are assigned to a travel shed
  - Manually review and clean

Easy Criteria

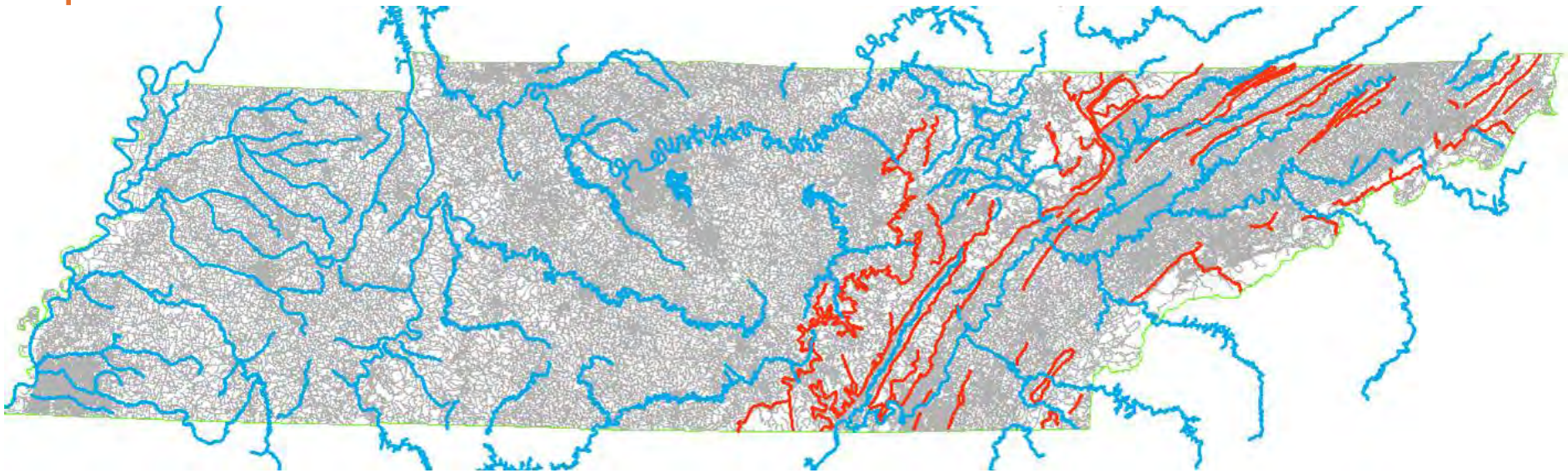
Counties



Places



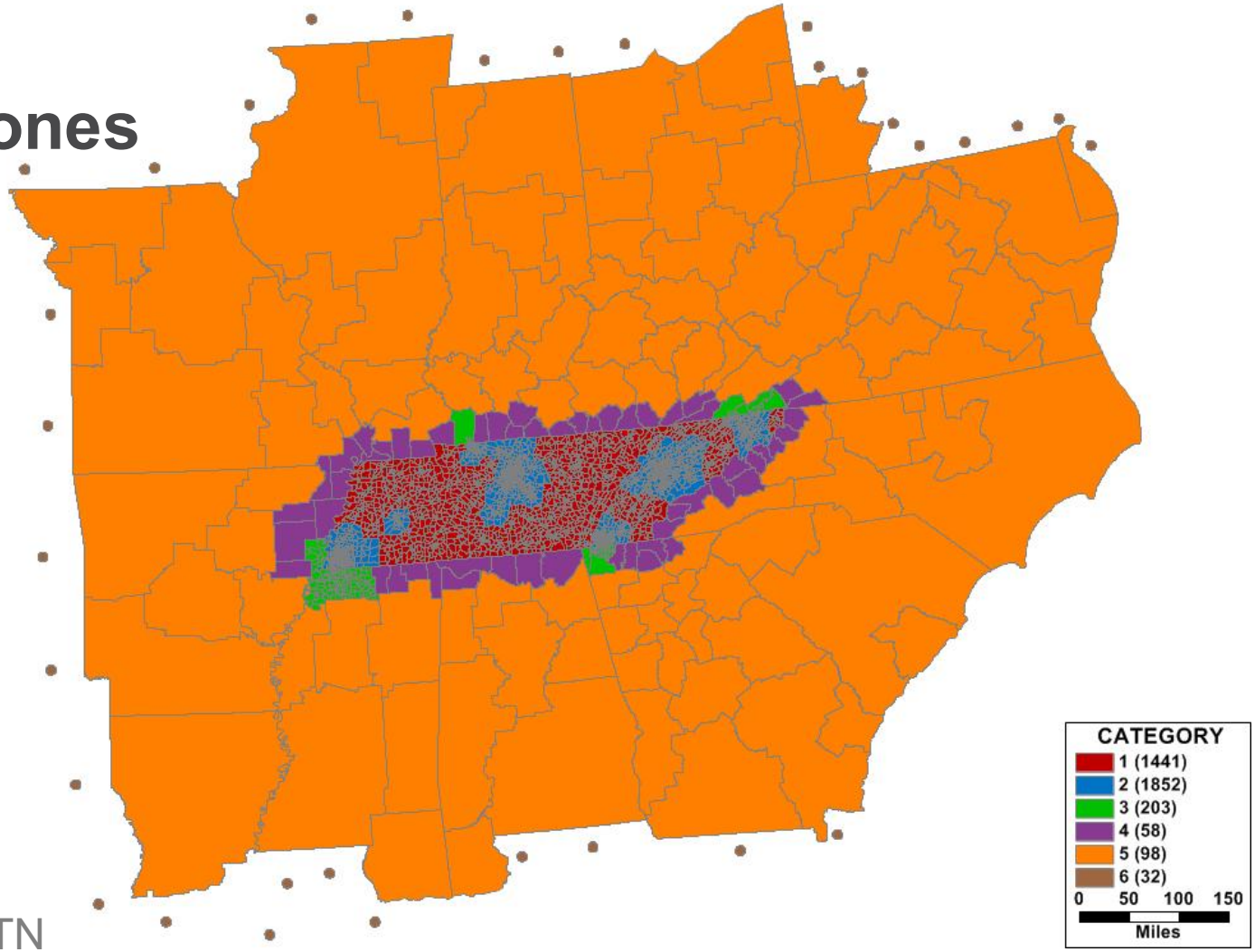
Harder Criteria



Slopes, Ridgelines and Water Features

- Plenty of water layers, but how to define “major”
- No canned “ridgeline” layers
- Created ridgelines by processing DEMs
- Tried to define “internal” criteria
- Instead, used visual inspection against TeleAtlas, looking for network gaps
- Found slopes more a barrier than ridges in some areas

Types of Zones



Six Categories

1. TN Rural
2. TN MPO
3. MPO outside TN
4. Single county outside TN
5. Multi-county outside TN
6. External stations

Socioeconomic Data Development

Demographics

Decennial Census

- Population, Households, Children, Seniors
- Block level data

American Communities Survey

- Workers, Vehicles, Income
- Block group level available
- Disaggregate to blocks proportionally to households

Employment Categories

Industry Categories

- Using standard 20 two digit NAICS categories for data development to support commodity flow modeling in Version 3

NAICS Code	Description
11	Agriculture, Forestry, Fishing and Hunting
21	Mining, Quarrying, and Oil and Gas Extraction
22	Utilities
23	Construction
31-33	Manufacturing (31, 32, 33)
42	Wholesale Trade
44-45	Retail Trade (44 & 45)
48-49	Transportation and Warehousing (48 & 49)
51	Information
52	Finance and Insurance
53	Real Estate and Rental and Leasing
54	Professional, Scientific and Technical Services
55	Management of Companies and Enterprises
56	Administrative and Support and Waste Management and Remediation Services
61	Educational Services
62	Health Care and Social Assistance
71	Arts, Entertainment, and Recreation
72	Accommodation and Food Services
81	Other Services, except Public Administration
92	Public Administration

Employment Data Sources

MPOs

- Total employment estimate at MPO TAZ level
- Not full two-digit NAICS breakouts

InfoGroup

- Purchased data for all Tennessee – available to MPOs
- Individual business with lat, long locations
- Based on phone surveys, aggregated data

LEHD

- Freely available federal data
- Employment by NAICS category by Census block
- Based on administrative (tax) records, with some 'fuzziness' for privacy

BEA

- Freely available federal data
- Total employment by NAICS category at County level

Woods & Poole

- Purchased data including employment forecasts consistent with BEA

Urban Employment Allocation Process

Iterative Proportional Fitting (IPF)

- For each county:
- Scale MPO total employment by TAZ to 2010 BEA (if necessary)
- Apply average of InfoGroup and LEHD 2-digit NAICS breakout within each zone to create seed distribution of employment by TAZ by industry
- Apply IPF to seed distribution
- Resulting employment must
  - Respect MPO TAZ total employment (scaled to BEA)
  - Match county level BEA totals by industry
- Some limited manual cleaning/adjustment was necessary

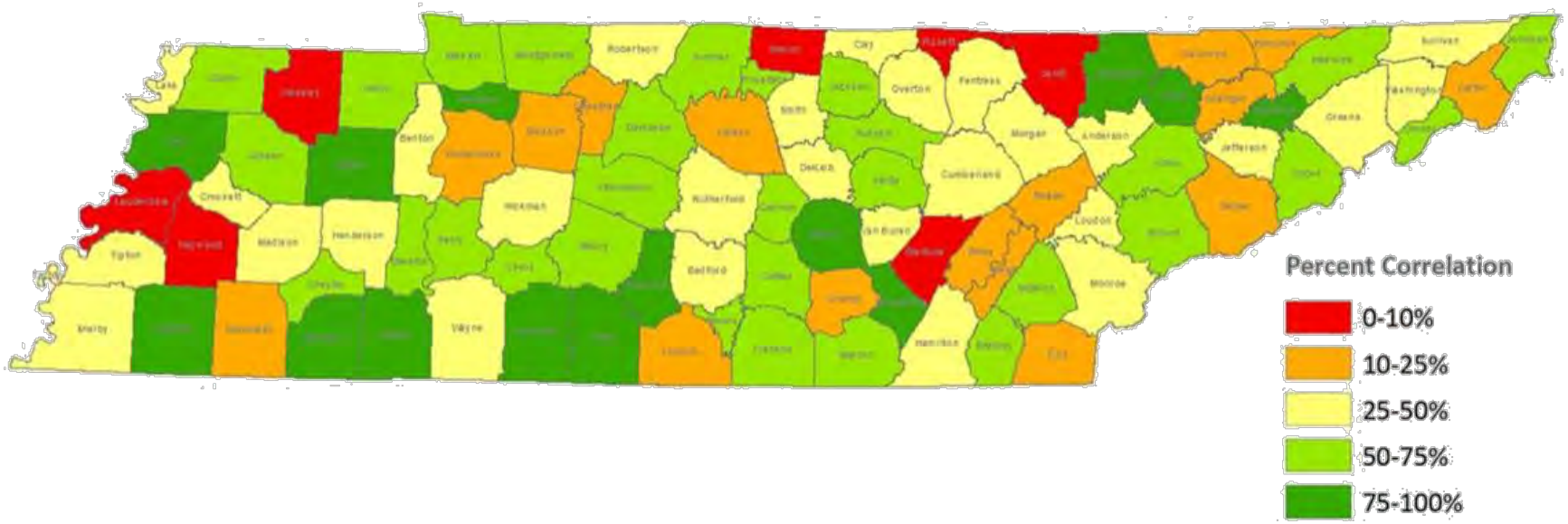
Using InfoGroup and LEHD Together

Cleaning

- Compare differences and correlations look for outliers

Combining

- Both InfoGroup and LEHD account for roughly 85% of BEA
- If they are independent, together they would account for 98% of BEA
- Research in Ohio suggests they are close to independent



Rural Employment Allocation Process

Using InfoGroup and LEHD together

- First at the TAZ level, identify any cases where InfoGroup (IG) and LEHD differ by > 200 employees for any industry
  - Determine whether to use IG, LEHD or split the difference based on IG and LEHD coverage of that industry in that county or manual investigation for very large discrepancies
  - Replace IG and/or LEHD estimate with chosen value to create 'cleaned' versions
- Second, choose how to use the two datasets together and scale them to BEA
  - Using 'cleaned' IG and LEHD calculate the ratio of their sum to the BEA total for each county and industry

    If < 1.5, scale maximum of 'cleaned' IG, LEHD to BEA

    If between 1.5 and 2.5, scale average of 'cleaned' IG, LEHD to BEA

    If > 2.5, possibly scale minimum of 'cleaned' IG, LEHD to BEA

Socioeconomic Forecasts

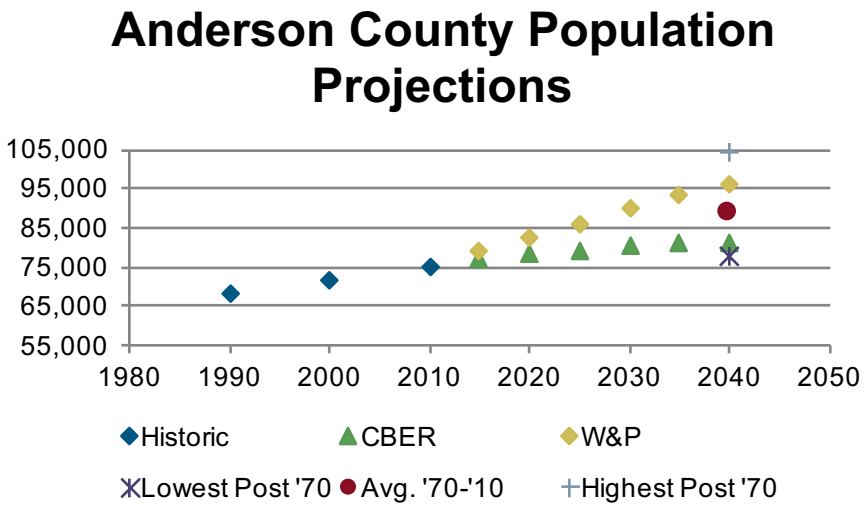
County Control Totals

Sources of Forecasts

- UT's Center for Business and Economic Research (CBER) – population only
- Woods & Poole
- MPO forecasts
- Historic growth rates and trends

Recommended Control Totals

- If 2 sources (rural employment)
  - If W&P within 10% of historic, use it, otherwise average with historic
- If 3 sources
  - use middle estimate
- If 4 sources (urban population)
  - If MPO forecast is not highest or lowest, use it
  - If MPO forecast is highest, use second highest
  - If MPO forecast is lowest, use second lowest



Allocation to TAZ

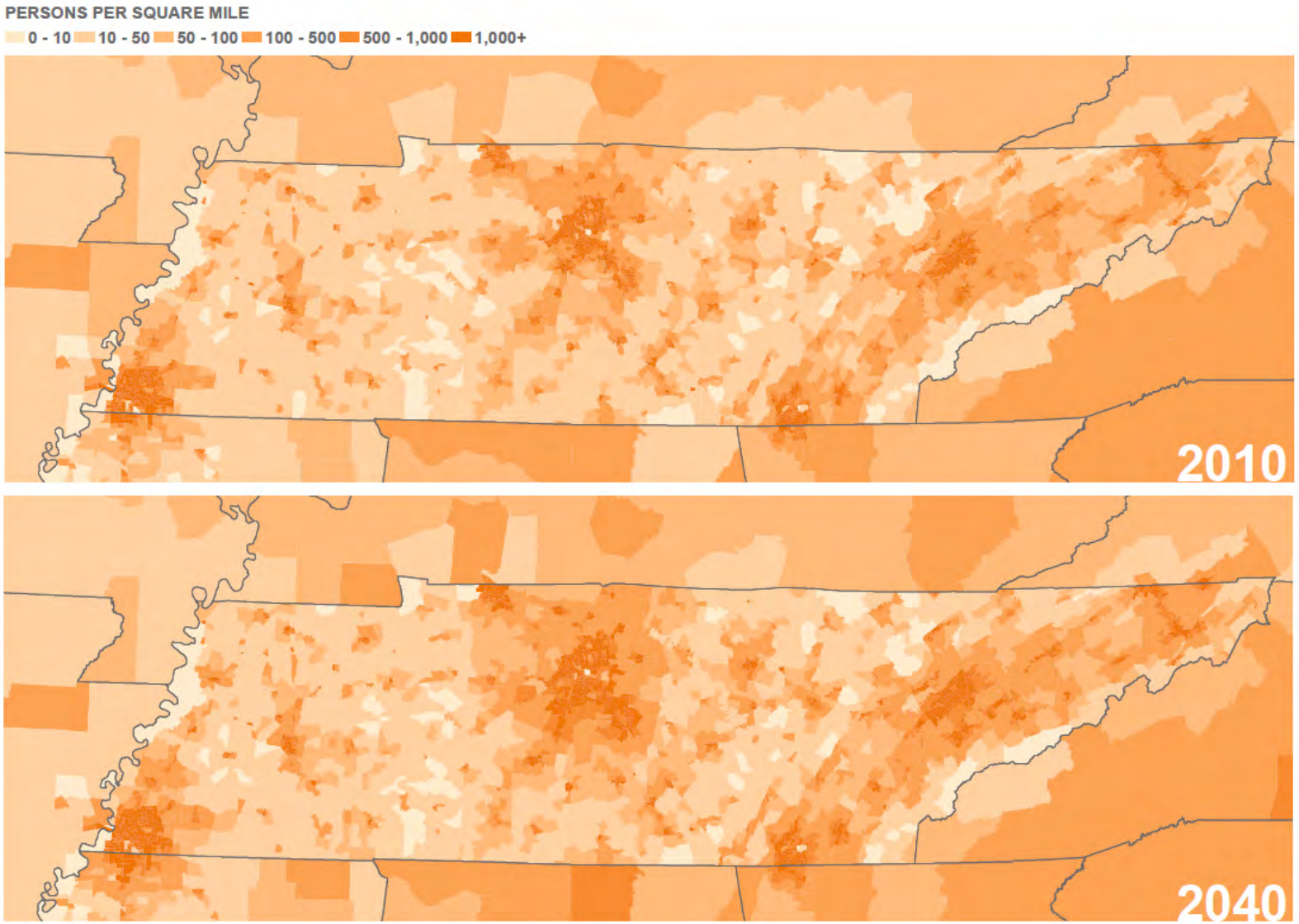
MPO areas

- Use MPO growth allocations
- Absolute growth may not match exactly if control totals differ, but same pattern will be assumed

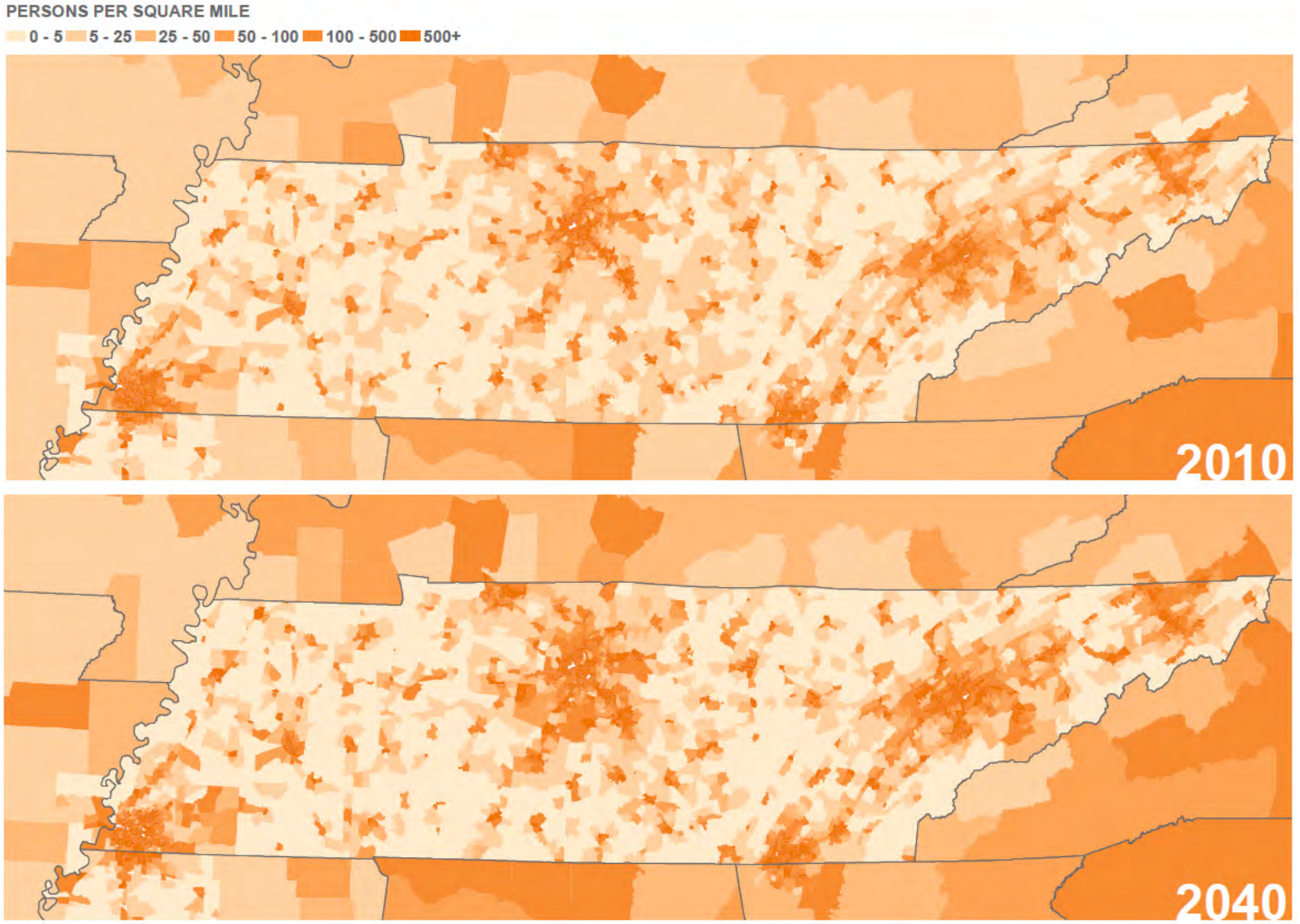
Non-MPO areas

- Population allocation where growth occurred between 2000 and 2010
  - Except imposing a floor of 30% of 2010 population in declining areas
- Employment allocation an average of two allocation processes:
  - Scaling 2010 employment to 2040 control total
  - Allocate future growth where growth has occurred between 2002 and 2012 (according to LEHD)

Population Density



Employment Density



ATRI Truck GPS Data

ATRI Truck GPS Data

What's ATRI?

American Transportation Research Institute (ATRI)

- non-profit funded by the trucking industry
- Receives over 4 Billion GPS truck positions annually from member organizations
- Cannot disclose the individual raw truck traces, but can provided processed data products which avoid disclosure
- Basis of FHWA's Freight Performance Measures Webtool
- Used for major corridor studies, I-95, I-70
- Incorporated in Indiana & Iowa's statewide models
- Will be primary basis of truck model in v2 Tennessee model

Indiana Experience

Data

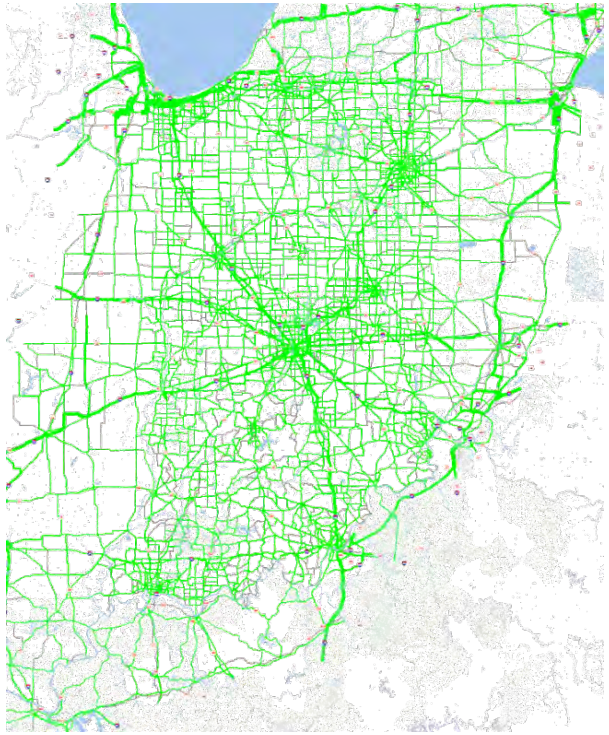
- Eight week sample
- 16 million records
- 305,000 trucks
- 2 million truck trips

Method

- Used existing commodity-flow based model to pivot off of expanded ATRI data

Results

Model	2006 Model	2010 Model
Observations	6,689	5,898
Avg. Count	1,379	1,264
RMSE	69.3%	60.6%
Avg Error	5.4%	-0.1%
MAPE	74%	42%



Data Processing

What constitutes a stop?

Anonymized GPS records converted to ODs

- Criteria based on speed and time
- Duration of a stop necessary to avoid counting traffic stops as destinations

from TAZ	to TAZ	distance	time	elapsed time	speed	status1	status2
10	101032	66.0	57.7	57.7	68.6	moving	moving
101032	101033	16.3	14.3	72.0	68.6	moving	moving
101033	**101015**	26.8	27.9	99.9	57.5	moving	moving
101015	101015	0.0	5.0	5.0	0.0	stopped	stopped
101015	101015	0.2	2.7	7.7	5.2	stopped	stopped
101015	101015	0.3	9.8	17.5	2.0	stopped	stopped
101015	101015	0.1	0.3	0.3	28.2	moving	*stopped?*
101015	2035	37.1	60.0	60.3	37.1	moving	moving
2035	18099	67.8	65.4	125.7	62.2	moving	moving
18099	27006	5.9	5.4	131.1	65.3	moving	moving
27006	**18023**	10.0	15.9	147.0	37.8	moving	moving
18023	18023	0.0	5.0	5.0	0.0	stopped	stopped

Trip	O	D
1	**10**	**101015**
2	**101015**	**18023**

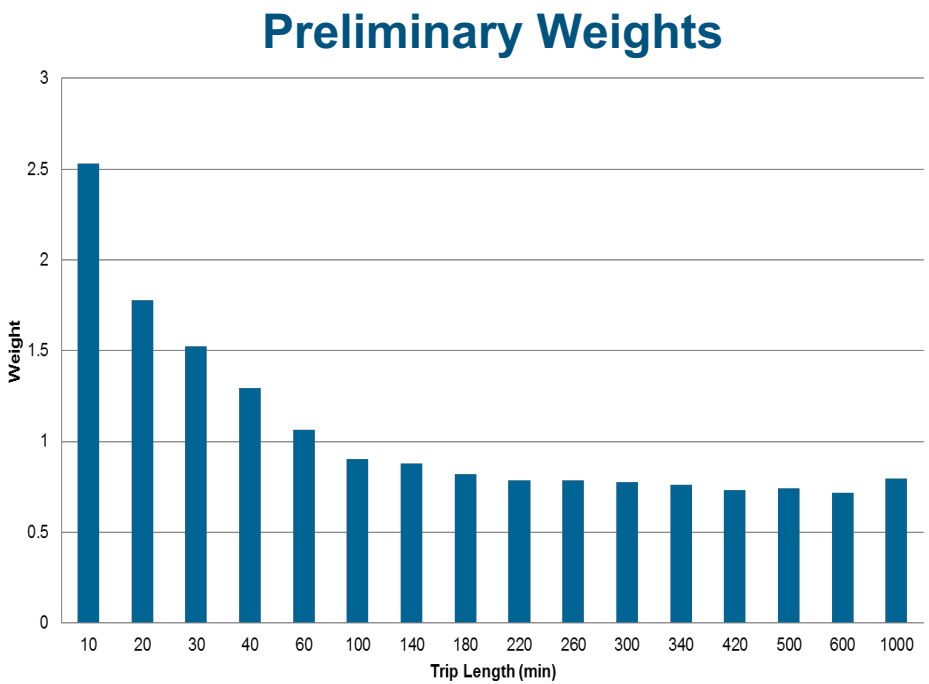
Data Cleaning

Need to clean/filter data for several reasons

- GPS blips
  - GPS location jumps from one place to another in a way that it could not be travel
- Start/End time
  - Trip fragments / partial trips in progress at beginning/end of sample periods
- Internal circuity
  - Correct for undetected stops
- Missing location data
- Long intrazonal trips (undetected stops)

Data Expansion

It's a big sample, but it's still a sample.



Simple Scaling

- Single uniform expansion factor
  - sample truck VMT to HPMS truck VMT

Preliminary Weighting

- Varying weights by trip length
- Weights developed from another study

Final Weighting

- Varying weights by
  - Region
  - Trip length
- Weights developed by analyzing results of ODME

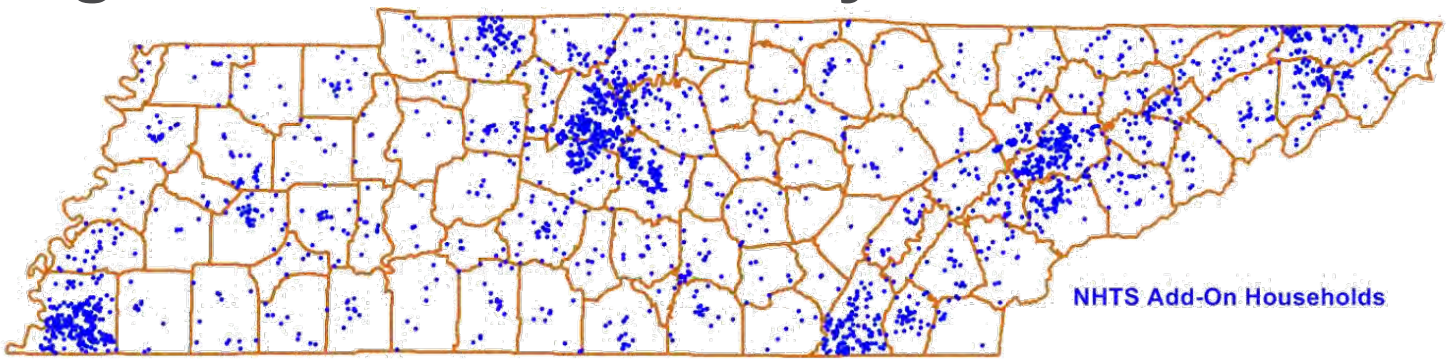
Tennessee ATRI Data

ATRI data includes ~11% of the multi-unit trucks on the road for 56 days

- Observations over 8 weeks in 2013
- Over 234,000 individual trucks
- Over 6.5 million truck trips (5.7 million after cleaning)
- DOT/RSG estimate 14 million daily multi-unit truck VMT in TN
- ATRI data (cleaned) contained 84.1 million truck VMT in TN over 56 days for a daily average of 1.5 million
- This represents 10.7% of the estimated total.
  - To produce the simple scaled OD table, this meant factoring down the raw ATRI data by 0.1672 to represent a daily number.

Household Survey Data

Combining NHTS & MPO Surveys



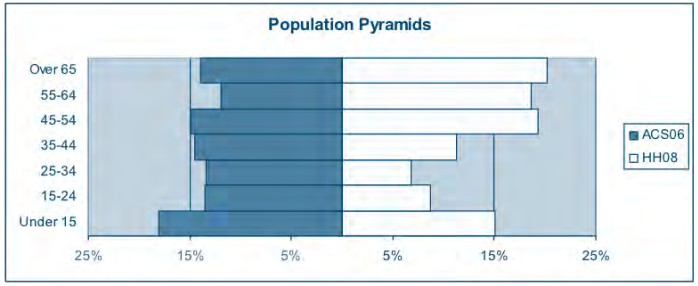
Datasets

- NHTS Add-On for Tennessee
  - Oversampled rural areas
- MPO surveys
  - Complete/complement NHTS
  - Used Middle TN & Knoxville

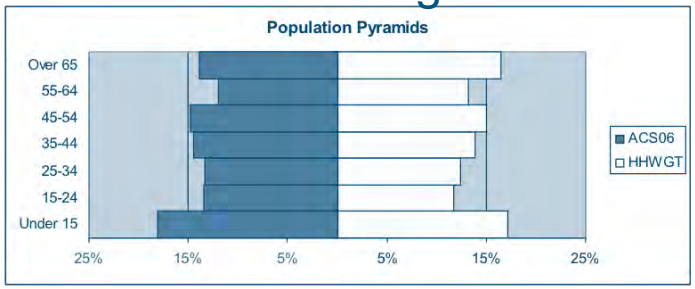
Re-weighting and combining

- Controls
  - Region
  - Household size by vehicles
  - Person age
- Iterative Proportional Fitting

ACS vs Unweighted



ACS vs Weighted



Phase 2 of the Statewide Model Update

Phase 2: Model Development

Advanced Trip-based Passenger Model

- Advanced trip generation
- Destination choice models
- Peak hour models

Truck / Freight Model

- Still being scoped

Validation

Post-processing

- Traffic statistics
- Post-processing

Advanced Trip Generation

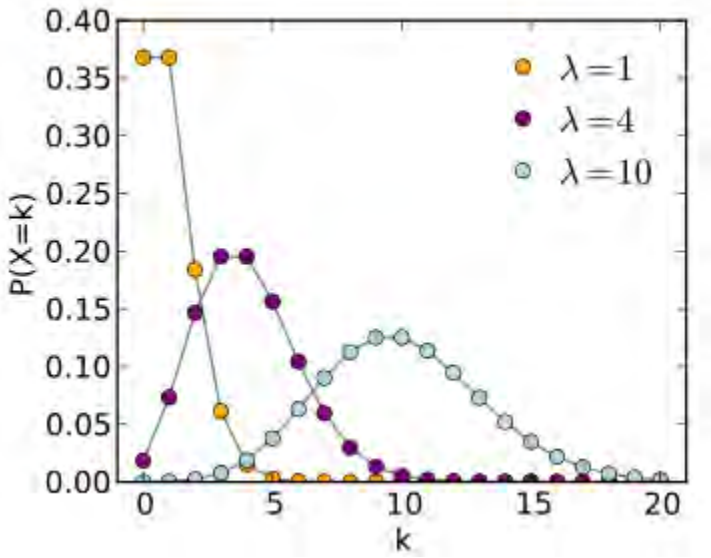
Non-linear Regression Models

- Allow multiple explanatory variables
  - Effect of area type /accessibility
  - Effect of seniors / children
- Capture both rational non-linearities
  - Diminishing returns to scale
  - Interaction effects

Poisson Distributed Household Variables

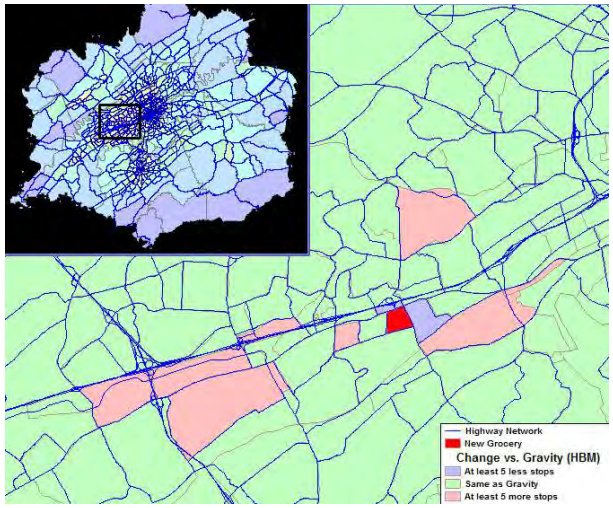
- Reduces aggregation as in cross-class
- But don't require stratification curves, etc.

Poisson Distributions



Source: Wikimedia Commons

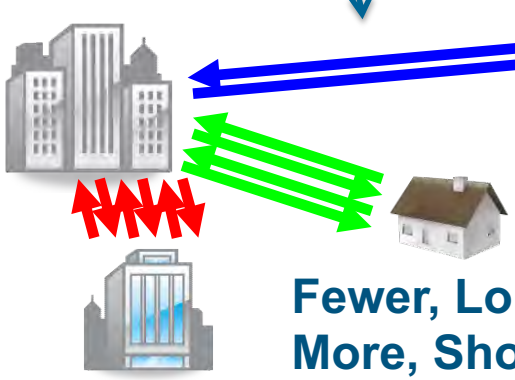
Destination Choice Models



Trip Chaining in Knoxville

Account for More Factors

- Number of Attractions
- Travel Time / Impedance
- Effect of Residence Location on Willingness to Travel
- Psychological Boundaries
  - River Crossings
  - Ridgeline Crossings
  - Major Highway Crossings
  - State / County Line Crossings
- Walkability of Destination
- Mixture of Land Uses at Destination
- Convenience for Trip-Chaining
- Spatial Auto-correlation Effects



Fewer, Longer Rural Trips
More, Shorter Urban Trips

Peak Hour Models

Nested Logit Models

- Must account for duration of long trips
- Upper nest determines at least some portion of the trip occurs in the AM peak hour, PM peak hour or both
- Lower nest determines how much of the trip occurs in the peak hour
- Will consider differences in peak hour factors related to
  - area type / accessibility of origin and destination
  - trip length
  - region

Truck Model

Model Structure

- Three step, like traditional passenger models
- Segmentation
  - by vehicle type
    - *light commercial vehicles*
    - *single unit trucks*
    - *multi-unit trucks*
  - by trip type
    - *commercial passenger trip*
    - *service delivery trip*
    - *freight delivery trip*

Developed based on ATRI data, pivot off of ATRI data

Validation

Demand Validation

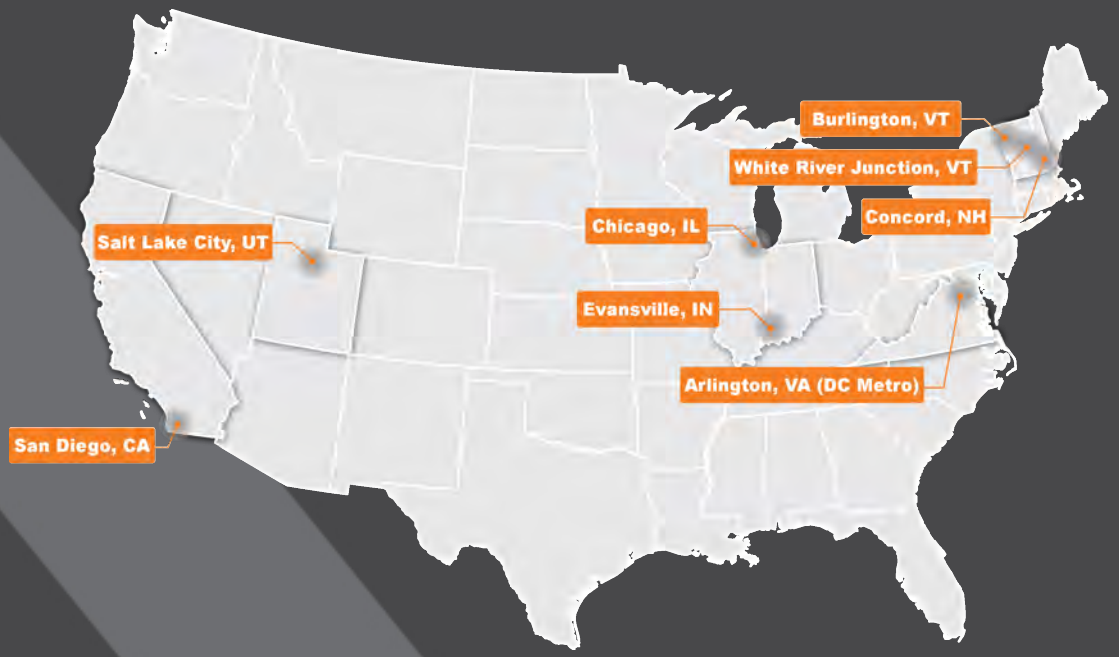
- Generation Rates
- Trip Lengths
- JTW Patterns

Assignment Validation

- Will produce similar statistics as for MPO models
- Different criteria for statewide models

%RMSE for Statewide Models from NCHRP 08-36-91

Volume Range		Average	Avg + 10%
1	5000	92.2	101.4
5000	10000	51.2	56.3
10000	20000	46.7	51.4
20000	30000	32.4	35.7
30000	40000	29.1	32.0
40000	50000	18.0	19.8
50000	60000	18.6	20.5
60000	+	22.2	24.4
Total		54.5	60.0



Contact

Vince Bernardin, PhD, RSG
Statewide Model Update Project Manager

Vince.Bernardin@RSGinc.com
812.200.2351

www.rsginc.com