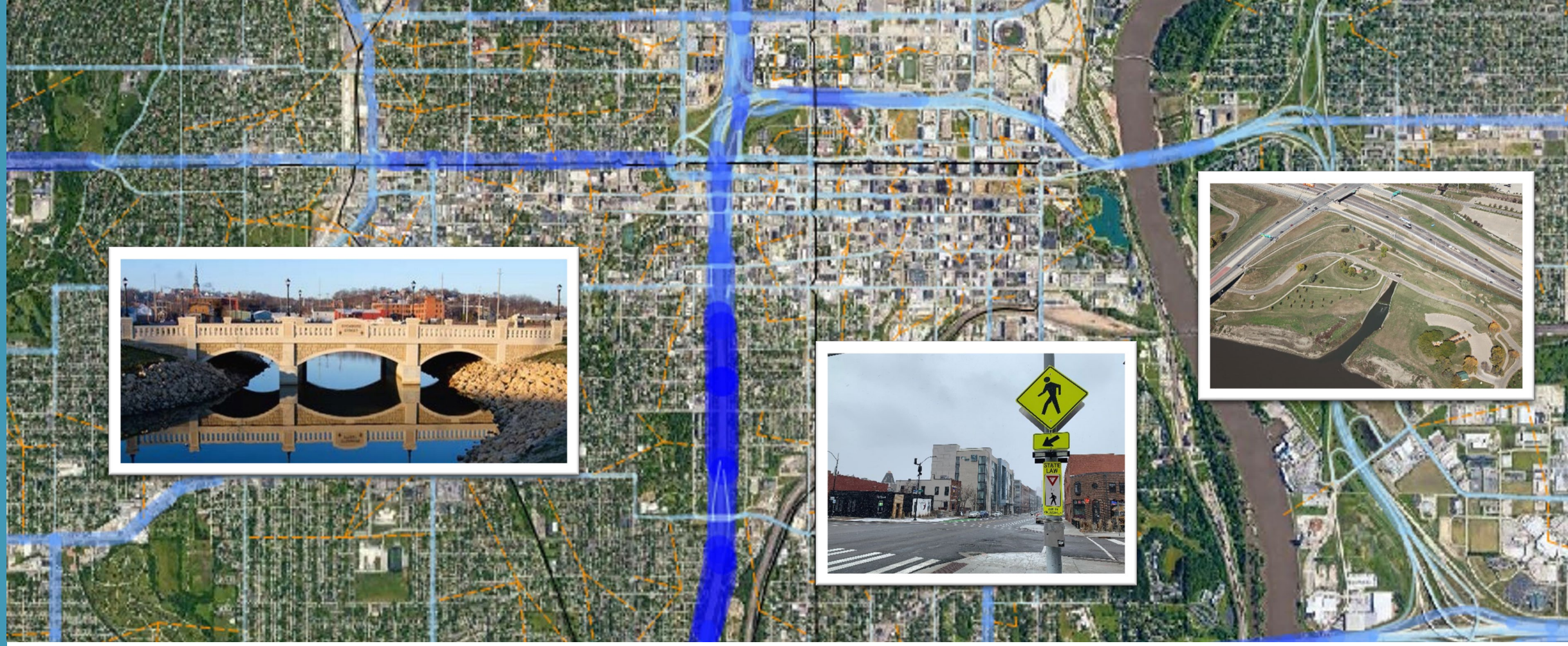


Model Calibration, Validation, and Reasonableness Checking

MTMUG

May 9, 2023

01 Calibration versus Validation versus Reasonableness Checking

02 Calibration Approach & Tips/Suggestions

03 (Some) Calibration/Validation Resources

04 (Some) Validation Data Sources

05 (Some) Typical Validation Checks

06 Other Considerations & Final Thoughts

Calibration versus Validation versus Reasonableness Checking

- **Calibration**: adjustment of model constants to better replicate observed results.

- **Validation**: comparison of a model to observed data not directly used in the model development.

- **Reasonableness Checking**: comparing model outputs to expected results. The expected results may be ranges of outputs from other models across the country or outputs from other mathematical means such as linear regression or professional experience.

Calibration Approach

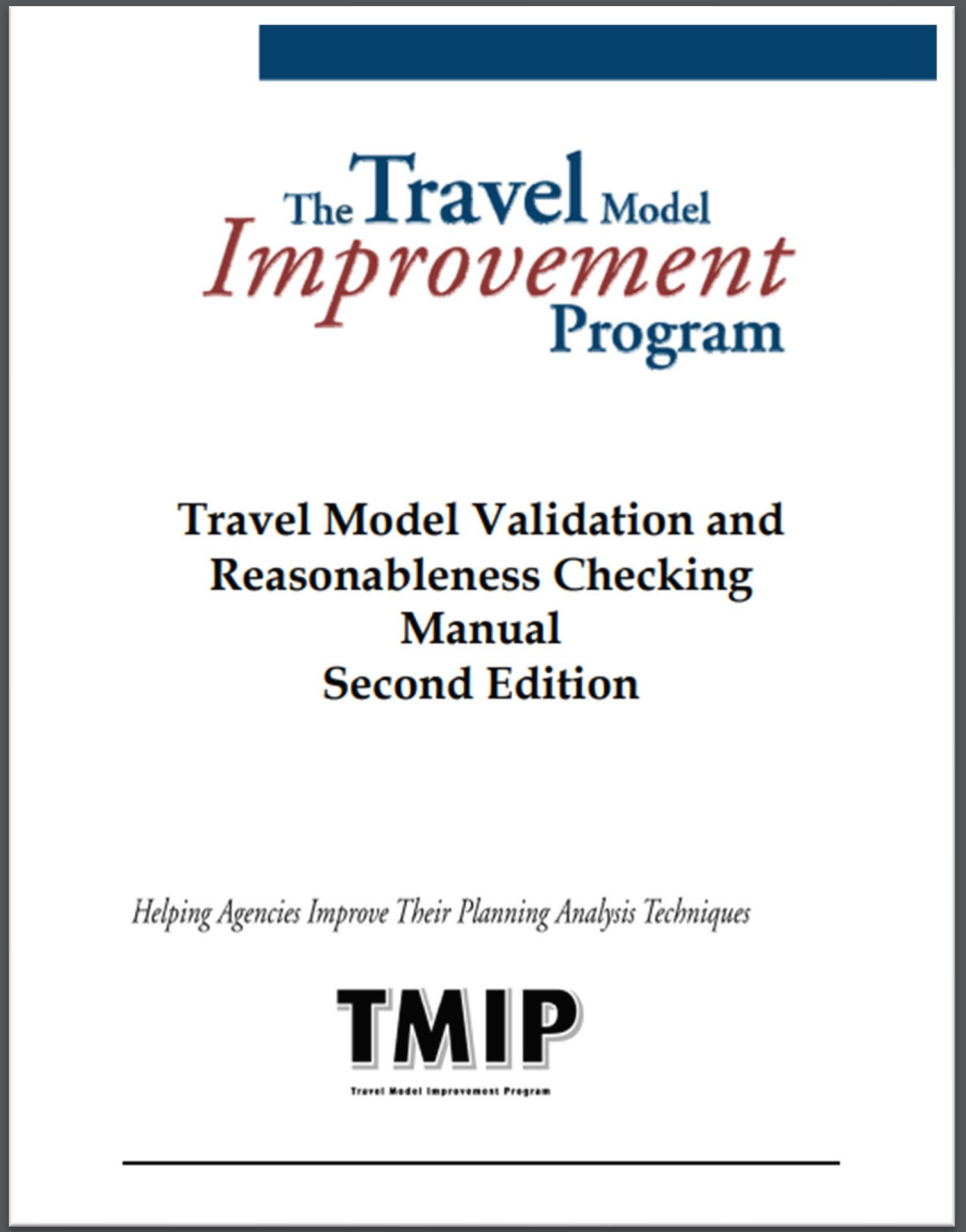
Tips/Suggestions:

- Check all major steps
- Make “apples-to-apples” comparisons
- Look for errors before making calibration adjustments
- Focus on systematic adjustments
- Don’t over-calibrate
- Know when to stop

Source: DALLI-E OpenAI
“Jigsaw puzzle of traffic congestion”

(Some) Calibration/Validation Resources

- TMIP Validation Manual
- NCHRP 716 (NCHRP 365)
- ISMS Manual



Published in 2010

(Some) Calibration/Validation Resources

- TMIP Validation Manual
- NCHRP 716 (NCHRP 365)
- ISMS Manual

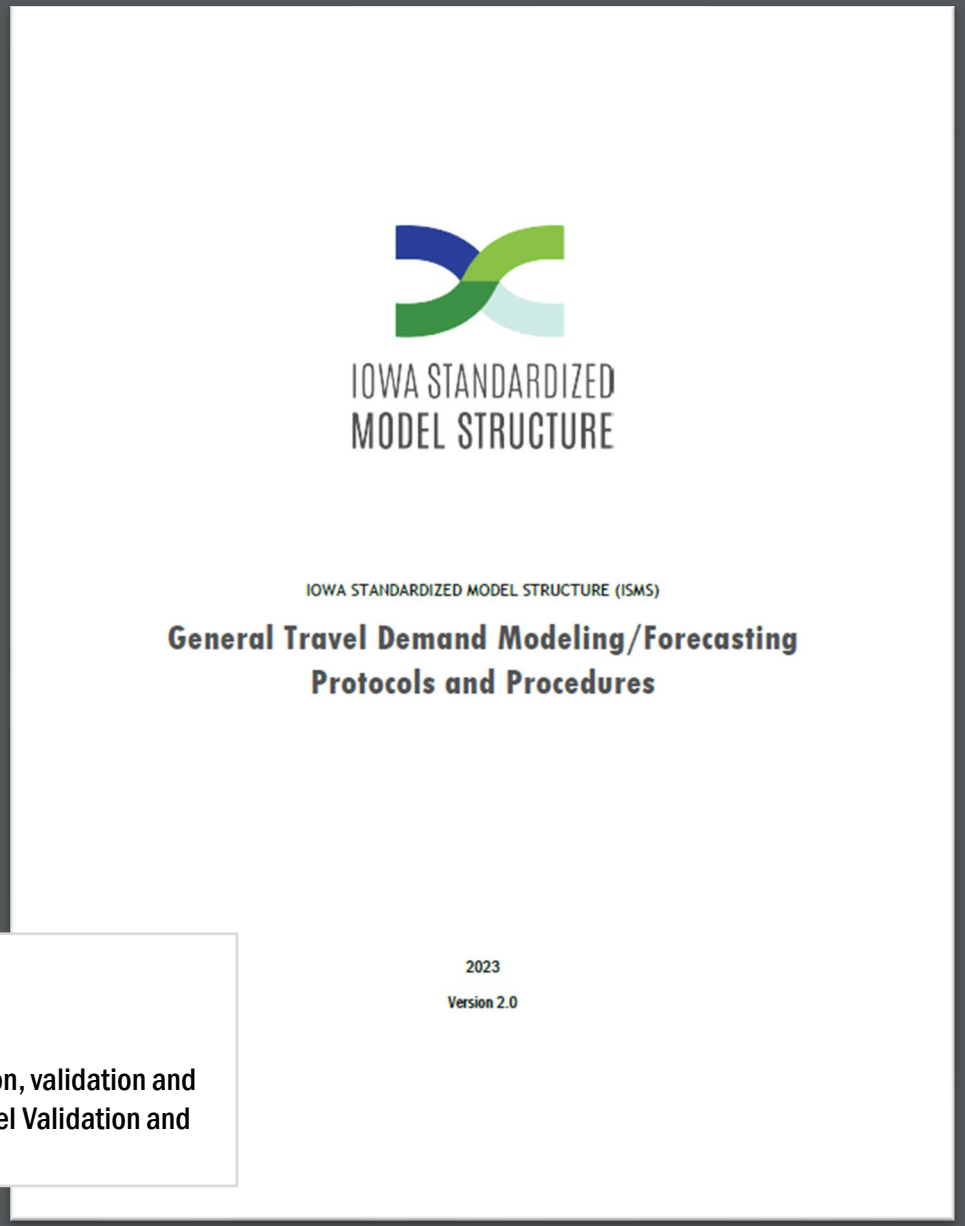
Table 3.5. Example primary and secondary model validation tests.

Model Component	Primary Tests	Secondary Tests	Potential Validation Data Sources
Networks/Zones	• Correct distances on links • Network topology, including balance between roadway network detail and zone detail • Appropriateness of zone size given spatial distribution of population and employment • Network attributes (managed lanes, area types, speeds, capacities) • Network connectivity • Transit run times	• Intrazonal travel distances (model design issue) • Zone structure compatibility with transit analysis needs (model design issue) • Final quality control checks based on review by end users • Transit paths by mode on selected interchanges	• GIS center line files • Transit on-board or household survey data
Socioeconomic Data/Models	• Households by income or auto ownership • Jobs by employment sector by geographic location • Locations of special generators • Qualitative logic test on growth • Population by geographic area • Types and locations of group quarters • Frequency distribution of households and jobs (or household and job densities) by TAZ	• Dwelling units by geographic location or jurisdiction • Households and population by land use type and land use density categories • Historical zonal data trends and projections to identify "large" changes (e.g., in autos/ household from 1995 to 2005)	• Census SF-3 data • QCEW • Private sources, such as Dun & Bradstreet
Trip Generation	• Reasonableness check of trip rates versus other areas • Logic check of trip rate relationships	• Checks on proportions or rates of nonmotorized trips • Reasonableness check of tour rates • Cordon lines by homogeneous land use type	• Chapter 4 of this report • Traffic counts (or intercept survey data) for cordon lines • Historic household survey data for region • NHTS (2001 or 2009)
Trip Distribution	• Trip length frequency distributions (time and distance) by market segments • Worker flows by district • District-to-district flows/desire lines • Intrazonal trips • External station volumes by vehicle class	• Area biases (psychological barrier—e.g., river) • Use of k-factors (Design Issue) • Comparison to roadside intercept origin-destination surveys • Small market movements • Special groups/markets • Balancing methods	• ACS/CTPP data • Chapter 4 of this report • Traffic counts (or intercept survey data) for screenlines • Historic household survey data for region • NHTS (2001 or 2009)

Published in 2012

(Some) Calibration/Validation Resources

- TMIP Validation Manual
- NCHRP 716 (NCHRP 365)
- ISMS Manual



3.3.8 Validation, Calibration, and Reasonableness Checking

...

Each step in the ISMS recommended model architecture includes guidance on model calibration, validation and reasonableness checks. Many reference industry standards from publications such as the Model Validation and Reasonableness Checking Manual.

(Some) Validation Data Sources

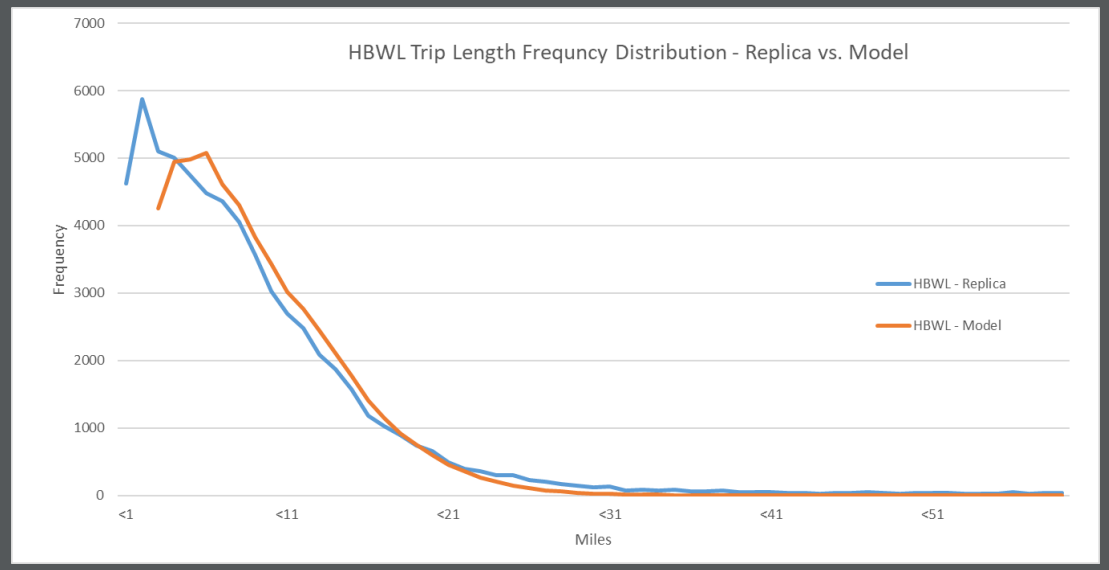
- Census
- CTPP
- LEHD OnTheMap
- Wood & Poole / BLS
- NHTS Add-on
- Other local data or studies (e.g., transit ridership data, local truck study, university employee travel survey, etc.)
- On-Board Travel Survey
- NTD
- Big Data (StreetLight, AirSage, ATRI)
- Replica
- INRIX
- Google Maps / Aerials
- Previous model
- Historic data
- Other models
- “Industry Standards”
- Traffic Counts
- Your brain!

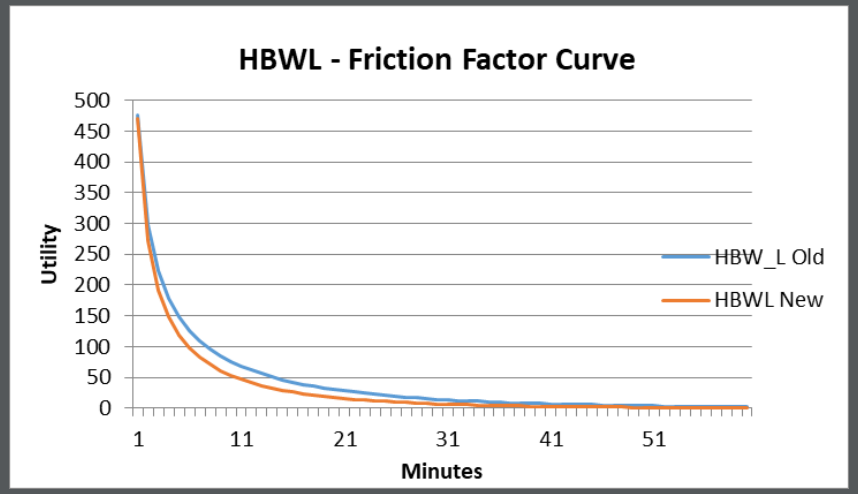
(Some) Typical Validation Checks

- **Trip Generation:**
  - **Households**: Parcel household data checks (vs. Census data)
  - **Non-Residential Land Use**: General parcel land use data checks for reasonableness (vs. other models)
  - **Unbalanced Trips**: Unbalanced trip productions versus trip attractions from model (10% rule)
  - **Trip Purpose Percentages**: Percentage of each balanced trips by trip purpose (vs. NHTS/Other Models/"industry standards")
  - **Trips per Household**: Total trips divided by households (vs. NHTS/Other Models/"industry standards")

(Some) Typical Validation Checks

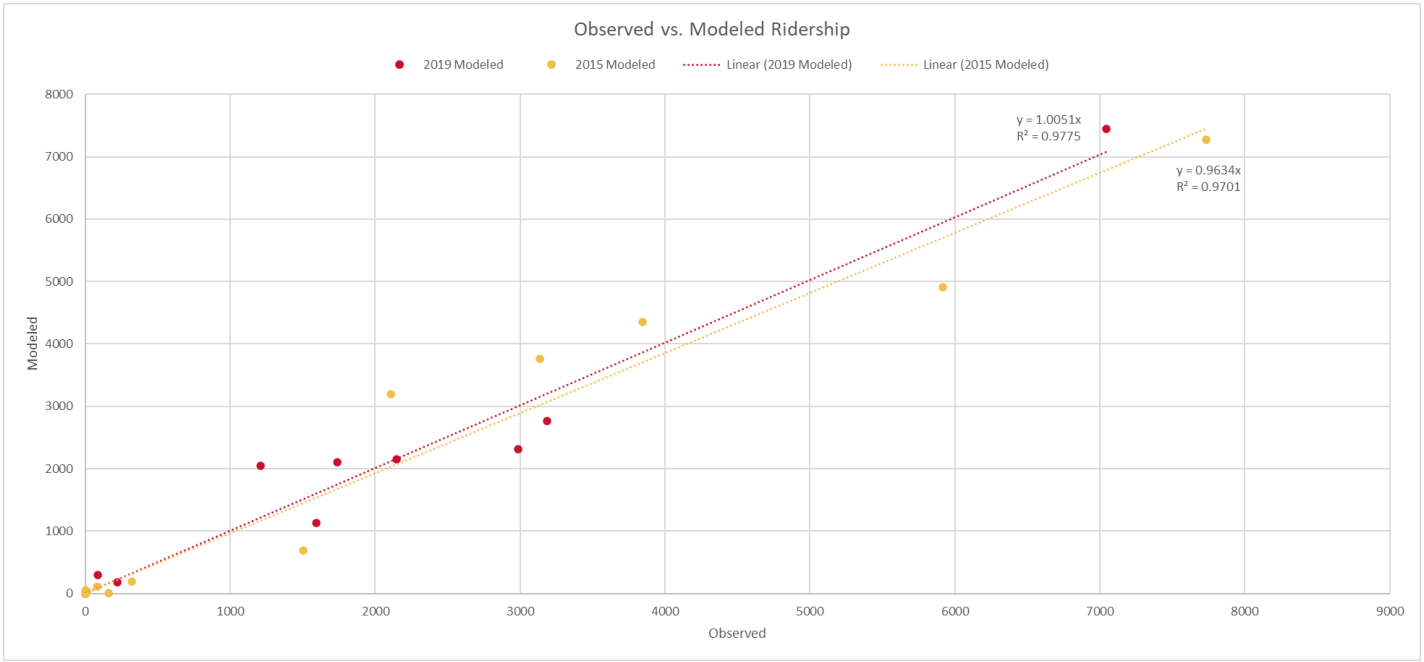
- **Trip Distribution:**
  - **Average Trip Time or Length**: Average trip time or distance by trip purpose (vs. NHTS/CTPP/Replica/StreetLight)
  - **Trip Length Frequency Distribution Curves (TLFD)**: Histogram of trips by time or distance bins (vs. NHTS/CTPP/Replica/StreetLight)
    - Coincidence Ratio (> 0.70)
  - **Friction Factor Curves**: Graphical display of the utility (or disutility) of making trips of various lengths of time (vs. Old model/"industry standards")





(Some) Typical Validation Checks

- **Mode Choice/Transit Assignment**
  - **Non-Motorized Trips**: Number and percentage of walk and bike trips (vs. NHTS/Replica)
  - **System Ridership:** Total number of “unlinked” transit trips (vs. NTD/local data)
  - **Ridership by Route:** Total ridership by route (vs. local data)

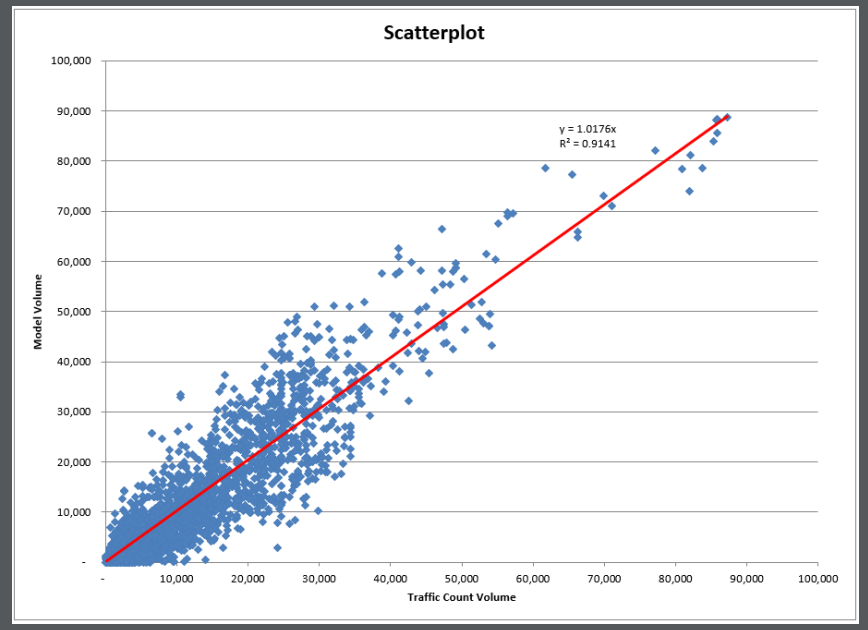


(Some) Typical Validation Checks

- **Traffic Assignment**
  - **VMT Error:** VMT by functional class and volume groups (vs. count data)
  - **Percent RMSE:** Average volume error (vs. count data)
  - **Scatterplot:** Graphical display of volume error (vs. count data)
  - **Level-of-Service:** Areas of congestion (vs. local knowledge)
  - **Screenlines/Cutlines/Cordons:** Regional or local volume error checks (vs. count data)

Table 4-64: Assignment Validation Standards by Facility Type

Facility Type	Total %RMSE - Acceptable	Total %RMSE - Preferred
Total System	35%	30%
Interstate/Freeway/Ramps	30%	25%
Expressway	35%	30%
Principal arterial	35%	30%
Minor arterial	45%	40%
Collectors/Local	65%	50%
Centroid connectors	N/A	N/A



(Some) Typical Validation Checks

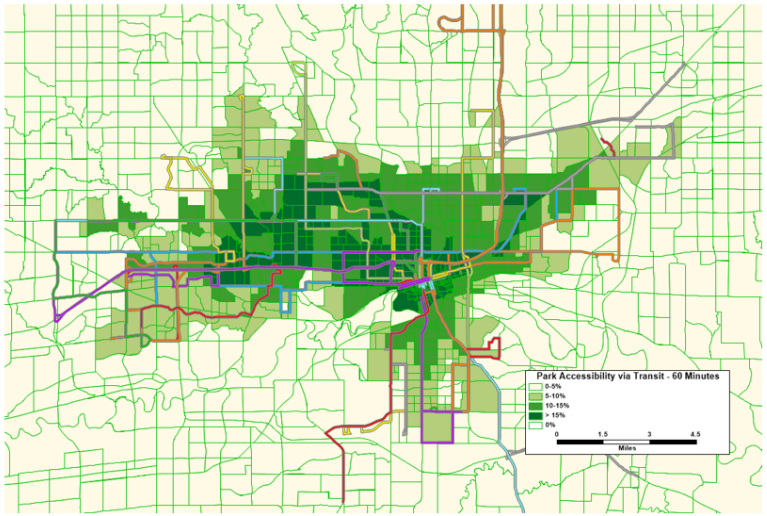
- Sensitivity Testing
  - Future Growth Summary
  - Future Trips in No-growth TAZs
  - Future Level-of-Service
  - New Road Scenario
  - Accessibility Scenario

Table 4-68: Growth Metric Reasonableness Checks

METRIC (BASE AND FORECAST)	SOURCE/FORMULA	DESCRIPTION
Households	Validation Report	Growth in households
Balanced Trips (Weekday)	Validation Report	Growth in trips. Percentage should be similar to growth in households
Vehicle Miles Traveled (Weekday)*	Validation Report	Most communities can expected higher VMT growth than balanced trips or household growth because development often occurs farther away from the city center.
Vehicle Hours Traveled (Weekday)*	Validation Report	Most communities can expect higher VHT growth than VMT growth because congestion usually is expected to increase in the future.
Average Trip Speed	VMT / VHT	Most communities can expected a lower average trip speed in the future because of more congestion
Average Trip Length	VMT / Balanced Trips	Most communities can expect a higher average trip lengths in the future because development occurs farther away from the city center
Average Trip Time	VHT / Balanced Trips	Most communities can expect higher average trip times in the future because congestion increases.

** Exclude connectors*

Figure 4-37: Map of Accessibility by TAZ Output



Other Considerations & Final Thoughts

- Travel time checks (vs. INRIX, Google Maps)
- Truck trips (vs. truck counts, ATRI, StreetLight, Replica)
- External station input checks (vs. StreetLight, Replica)
- Time-of-day checks (vs. hourly counts, StreetLight, Replica)
- Auto occupancies (vs. NHTS, other models, “industry standards”)
- Capacities (vs. other models, “industry standards”)

- **COMMUNICATION & DOCUMENTATION!!!**

Questions?

Thank You!