The FHWA
Travel Model Improvement Program
Travel Model Validation Manual Update

(or Why Your Model is Wrong and What to Do About it)

Part I

presented by
Thomas Rossi
David Kurth
Cambridge Systematics, Inc.

October 27, 2010

1

Acknowledgments

- Webinar and Validation Manual prepared by Cambridge Systematics, Inc.
  - Thomas Rossi
  - David Kurth
- Prepared for Federal Highway Administration
  - H. Sarah Sun
- Arranged by Texas Transportation Institute
  - Gary Thomas

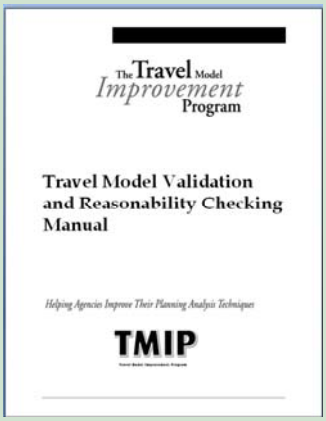
2

Acknowledgments

- Review panel of experts
  - Paul Agnello, Virginia DOT
  - Ken Cervenka, FTA
  - Alan Horowitz, Univ. of Wisconsin, Milwaukee
  - Ken Kaltenbach, The Corradino Group
  - Guy Rousseau, Atlanta Regional Commission
  - David Schmitt, AECOM
  - Ed Weiner, consultant

3

Travel Model Validation and Reasonability Checking Manual



4

Manual Chapters

1: Introduction – today

2: Model Validation Plan Specification – today

3: Validating Model Inputs – today

4: Socioeconomic Models – today

5: Amount of Travel/Activity– today

5

Manual Chapters (continued)

6: Trip Distribution/Destination Choice/ Location Choice – today

7: Mode Choice – tomorrow

8: Time of Day – tomorrow

9: Assignment Procedures

- Highway – today
- Transit – tomorrow

6

Manual Chapters (continued)

10: Temporal Validation and Sensitivity Testing – tomorrow

11: Validation Documentation – today

7

Webinar Focus

- Why update the 1997 manual?
  - Validation practices peer exchange questions
- New information in the manual
- Examples of procedures
  - Not a detailed review (you can read the manual)
- A “living” document

8

Peer Exchange Questions

- How important is it to match base year observations?
- How close is “close enough?”
- What steps needed to ensure accurate validation data?
- How should model sensitivity be tested?

9

Peer Exchange Questions

- Other tests to assess ability to forecast future travel?
- Should risk analysis of future forecasts be performed?
- What is the role of validation documentation in raising model credibility?

10

The New Manual Discusses, for Each Model Process…

- Sources of Data
- Aggregate Checks
- Disaggregate Checks
- Criteria Guidelines
- Reasonability and Sensitivity Testing
- Troubleshooting Strategies
- Forecasting Checks

11

Validation Is…

- "*Essentially, all models are wrong, but some are useful…the practical question is how wrong do they have to be to not be useful.*" – George Box
- So validation may be defined as "*steps to verify the ability of the model system to make reasonable predictions over a range of development patterns, transportation operations, and external factors.*"

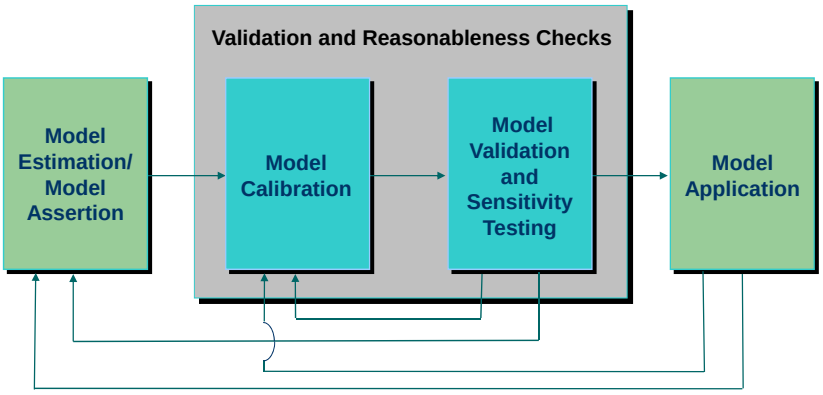
12

Validation Is Not…

- Standards a model must meet
- "Proof" of a model's
  - Appropriateness
  - Accuracy
- A "check-off" report card for the model

13

The Model Development Process



Model Calibration

- More than just regionwide validation – "adjusting constants"
- Segmentation variables – revising, adding, deleting
- Adjusting network parameters and settings
- Often "points back" to issues with other model steps

15

Overall Model Validation

- The overall results are the results of the final step (assignment)
- But results may indicate things to check in earlier model steps:
  - Screenline issues → check trip distribution
  - VMT too high or low → check trip rates
  - Link volume issues → check networks

16

Aggregate and Disaggregate Validation

- Disaggregate validation – required for disaggregate models
  - Explores how well model fits observed data at the household or individual level
  - Involves defining subgroups of observations
  - Compares model results with observed data to reveal systematic biases
- Aggregate validation – required for aggregate and disaggregate models
  - Provides a general overview of model performance through regional travel characteristics
  - Applies model at the regional, district, and zonal level

17

What's New
(Compared to 1997 Validation Manual)

- Taking advantage of new experience, data, and insights
- Validating emerging models (e.g. ABM)
- Validation plans
- Documentation of validation
- Web-based document

18

Model Validation Plan Specification (Chapter 2)

"You've got to be very careful if you don't know where you are going, because you might not get there." – Yogi Berra

19

Planning for Validation Tests

- Reason for test – what it is to demonstrate
- Detail required
- Goals for accuracy and sensitivity
- Observed data required / data sources
- Guide allocation of available funds for:
  - Model development
  - Model validation
  - Validation data collection

20

Planned Tests

- Define validation context – how/why models will be used
- Planned model uses help define validation tests and tolerance

21

Example
Daily Activity Pattern Model Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOMES	PRIORITY

Comparison of model parameters to other regions	• Comparison of model coefficients to:
	  • Sacramento
	  • San Francisco
	  • Columbus
		• No expectations; comparison only.	Level 1

Disaggregate	• Prediction success of modeled daily activity pattern choices against observed survey estimation data	• Prediction success likely to be very low	Level 3

Aggregate	• Numbers or percents of residents making tours and intermediate stops by activity type:
	  • For the region
	  • By county
	  • By household size and income group
	  • By gender and age group
	• Percent of "immobiles" by:
	  • By household size and income group
	  • By gender and age group
		• Compare modeled to expanded observed numbers or percents
		• Review for reasonable patterns
		• Compare to results summarized in external report
			Level 2

Source: DRCOG

22

Example
Volume-Over-Count Ratios and Percent Error

Standards	Statistic Acceptable	Preferable
Freeway Volume-over-Count (FT1x, FT8x, FT9x)	+/- 7%	+/- 6%
Divided Arterial Volume-over-Count (FT2x)	+/- 15%	+/- 10%
Undivided Arterial Volume-over-Count (FT3x)	+/- 15%	+/- 10%
Collector Volume-over-Count (FT4x)	+/- 25%	+/- 20%
One way/Frontage Road Volume-over-Count (FT6x)	+/- 25%	+/- 20%
Freeway Peak Volume-over-Count	75% of links @ +/-20%	50% of links @ +/-10%
Major Arterial Peak Volume-over-Count	75% of links @ +/-30%	50% of links @ +/-15%
Assigned VMT-over-Count Areawide	+/-5%	+/-2%
Assigned VHT-over-Count Areawide	+/-5%	+/-2%
Assigned VMT-over-Count by FT/AT/NL	+/- 25%	+/- 15%
Assigned VHT-over-Count by FT/AT/NL	+/- 25%	+/- 15%

Source: Florida DOT

23

Identifying Required Validation Data

- For each data item
  - Content
  - How used in validation
  - Geographic scale and extent
  - Time when data were collected
  - Consistency with previous data

24

Identifying Required Validation Data

- Validation data are subject to sampling error
  - Fallacy of "tight validation" to something with large confidence interval
- Assessment can focus/provide priorities for validation data collection
  - Assembly/mining of existing data
  - Collection of new data

25

Example
Allocation of Modeling Resources

Model Development Process Component	Desired Resource Allocation (Percent)
Data collection	39
Estimation	16
Calibration	17
Validation	17
Documentation	9

Source: FHWA Travel Model Validation Practices Peer Exchange

26

Validation of Model Inputs (Ch. 3)
Why Do It?

27

The Major Input Data Groups

- Network data
  - Highway
  - Transit
- Socioeconomic data (“zone” data)
  - Households/persons by type
  - Employment by type
  - Other

28

Aggregate Input Data Checks

- Socioeconomic data
  - Summation of TAZ data for different segments
  - Comparison to other sources of observed data
- Network data
  - Aggregation of coded network data by segment
  - Comparison to independently summarized data for the same strata

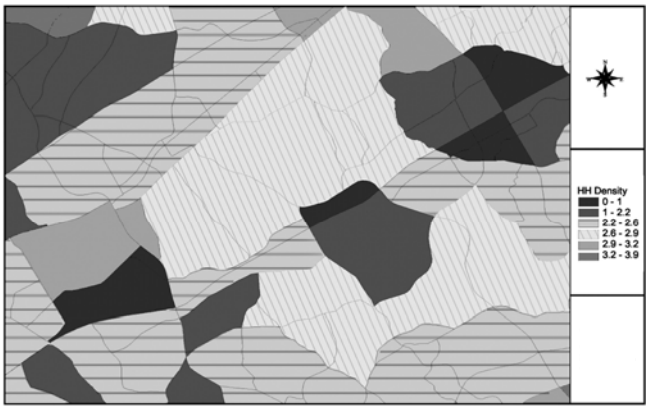
29

Disaggregate Input Data Checks

- Comprehensive disaggregate checks generally not feasible (no independent data source)
- Spot checks
- More ACS data becoming available for checking socioeconomic data
- Check all links in the region for “exceptional” characteristics

30

Reasonability Checks
Example



31

Forecasting Checks
Example

	Population Range				
Region	**> 1,000,000**	**500,000-1,000,000**	**200,000-500,000**	**50,000-200,000**	**All**
Average Growth Rates					
California	1.3%	2.0%	0.8%	1.8%	1.3%
Midwest	0.7%	0.6%	0.6%	0.4%	0.5%
Northeast	0.3%	0.4%	0.7%	0.3%	0.4%
Other West	2.5%	2.1%	2.1%	1.9%	2.0%
South Central	2.1%	1.6%	1.2%	0.5%	1.1%
Southeast	1.5%	2.0%	1.4%	1.3%	1.5%
All	**1.3%**	**1.3%**	**1.1%**	**0.9%**	**1.1%**
Ranges of Growth Rates					
California	0.2%–3.5%	0.8%–2.7%	-0.2%–2.3%	0.9%–2.6%	-0.2%–3.5%
Midwest	-0.3%–1.5%	-0.7%–1.9%	-0.5%–2.9%	-1.0%–1.9%	-1.0%–2.9%
Northeast	-0.5%–1.4%	-0.3%–1.2%	-0.3%–2.4%	-0.8%–2.8%	-0.8%–2.8%
Other West	1.2%–4.5%	0.5%–3.6%	0.3%–4.7%	0.3%–6.7%	0.3%–6.7%
South Central	1.2%–3.8%	0.7%–3.4%	-0.3%–3.5%	-0.8%–1.9%	-0.8%–3.8%
Southeast	0.7%–3.3%	0.8%–4.7%	-0.9%–3.7%	-0.5%–10.7%	-0.9%–10.7%
All	**-0.5%–4.5%**	**-0.7%–4.7%**	**-0.9%–4.7%**	**-1.0%–10.7%**	**-1.0%–10.7%**

32

Network Skims and Path Building

Transit Path-Building Prediction Success Table

PM Period Work Trips Using Walk to Rail		Modeled Boardings			
		No Path	1	2	3+
Reported Boardings	No Path	0	0	0	0
	1	7	3	4	0
	2	1	0	0	0
	3+	0	0	0	0

33

Socioeconomic Models (Ch. 4)

- Aggregate share models
- Discrete choice models
- Population synthesis

34

Example
Vehicle Availability Model

	County							
Variable	Moe	Larry	Curly	Groucho	Chico	Harpo	Zeppo	Region
Percent 0 Vehicles								
Observed	5%	13%	12%	7%	12%	7%	37%	17%
Model	3%	8%	8%	3%	5%	4%	24%	11%
Percent 1 Vehicle								
Observed	28%	35%	35%	31%	34%	33%	42%	35%
Model	27%	38%	38%	30%	34%	33%	49%	37%
Percent 2 Vehicles								
Observed	44%	37%	38%	44%	38%	43%	18%	34%
Model	47%	39%	39%	45%	43%	43%	21%	36%
Percent 3 Vehicles								
Observed	22%	14%	15%	18%	16%	17%	3%	13%
Model	22%	15%	15%	21%	18%	20%	5%	15%
Average Number of Vehicles								
Observed	1.86	1.55	1.57	1.78	1.58	1.74	0.88	1.44
Model	1.95	1.65	1.65	1.90	1.79	1.82	1.11	1.59

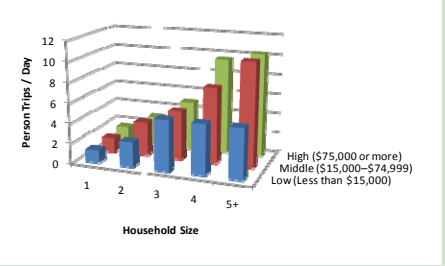
35

Example
Aggregate Share Vehicle Availability Model

	0 Vehicles		1 Vehicle		2 Vehicles		3+ Vehicles		
Range	Obs.	Model	Obs.	Model	Obs.	Model	Obs.	Model	Total
0.00–0.05	181	181	0	0	0	0	0	0	**181**
0.25–0.35	308	293	28	33	5	15	8	8	**349**
0.35–0.45	29	35	17	7	0	3	0	2	**46**
0.45–0.55	382	391	115	119	38	34	33	23	**568**
:									
2.85–2.95	28	6	47	22	49	30	94	160	**218**
2.95 or more	143	10	178	31	87	73	113	406	**521**

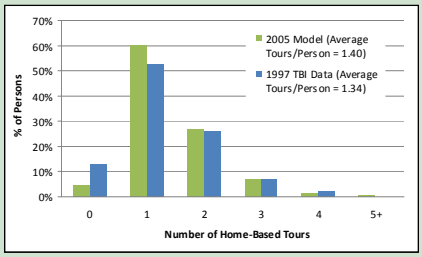
36

Amount of Travel/Activity (Ch. 5)



For each household:
- **Home-based Work Person Trips**
- **Home-based Non-Work Person Trips**
- **Non-home-based Person Trips**

For each person in each household:
- **Number of activities by purpose**
- **Number of tours by purpose**
- **Number of stops on each tour**
- **Number of work-based subtours**



37

Data Sources

- Household travel/activity survey
  - Aggregate validation of trip-based
  - Aggregate & disaggregate validation of activity-based
- National sources
  - Aggregate validation of trip-based
  - Aggregate validation of activity-based

38

National Averages – Person Trips / Day by Household Size & MSA Size

	Person Trips / Day (All Trip Purposes)					
	Number of Persons					All
MSA Size	1	2	3	4	5+	Households
MSA population greater than 3 million	3.4	8.6	12.4	16.8	22.2	10.7
MSA population between 1 and 3 million	3.4	8.9	13.2	17.5	22.1	10.6
MSA population between 500,000 and 1 million	3.3	9.0	12.4	17.4	22.8	10.9
MSA population between 250,000 and 500,000	3.4	8.8	12.8	17.0	24.2	10.8
MSA population less than 250,000	3.8	8.9	13.0	19.4	22.2	10.7
Not in MSA	3.2	8.5	12.8	17.0	21.4	10.2

	Home-Based Work	Home-Based Nonwork	Nonhome-Based
MSA population greater than 3 million	14%	55%	31%
MSA population between 1 and 3 million	15%	53%	32%
MSA population between 500,000 and 1 million	14%	54%	32%
MSA population between 250,000 and 500,000	14%	54%	33%
MSA population less than 250,000	14%	53%	34%
Not in MSA	14%	52%	34%

Source: 2001 NHTS

39

Types of Amount of Travel / Activity Checks

Type of Check	Trip-Based	Activity-Based
Trips per person or household by purpose	✓	✓
Activities per person or household by type		✓
Percent of trips by purpose	✓	✓
Number of activities tours by type		✓
Percent of persons making 0, 1, 2… tours		✓
Percent of "immobile" persons		✓
Production / attraction balance	✓	
Trips per person or household by area type	✓	✓

40

Trip Distribution/Destination Choice/Location Choice (Ch. 6)

Trip-based Models

- Gravity models
- Destination choice models

Activity-based Models

- Regular workplace / school place location
- Destination choice models
- Intermediate stop location

41

Data Sources

- Household travel/activity survey
  - Aggregate trip length frequency distributions
  - Aggregate district interchange (limited)
  - Disaggregate validation of activity-based
- National sources (Census, ACS, NHTS)
  - Aggregate trip length frequency distributions
  - Aggregate district interchange

42

National Averages – Reported Average Trip Durations by MSA Size

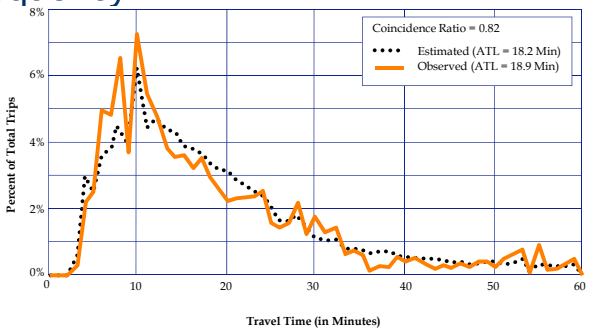
MSA Size	Reported Average Trip Length in Minutes: Home-Based Work	Home-Based Nonwork	Nonhome-Based
MSA population greater than 3 million	31	17	18
MSA population between 1 and 3 million	25	17	19
MSA population between 500,000 and 1 million	24	18	18
MSA population between 250,000 and 500,000	21	18	20
MSA population less than 250,000	20	18	21
Not in MSA	21	19	19

Source: 2001 NHTS

43

Types of Trip Distribution / Destination Choice Checks

- Average trip length
- Trip length frequency distribution
- Coincidence ratio

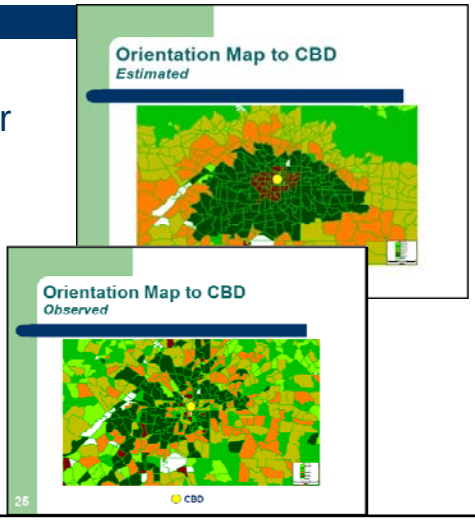


Note: ATL = Average Trip Length.

44

Types of Trip Distribution / Destination Choice Checks (cont.)

- District-to-district interchanges / other visual methods
- Disaggregate prediction-success tables



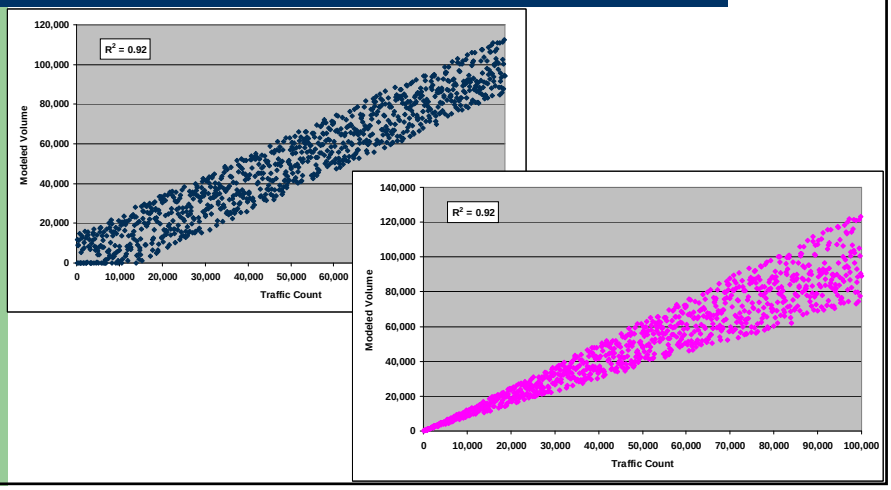
45

Assignment Procedures (Ch. 9)
Highway Assignment

An R^2 value of 0.88 is an arbitrary value – achieving such a standard is neither necessary nor sufficient in establishing the validity of a travel model.

46

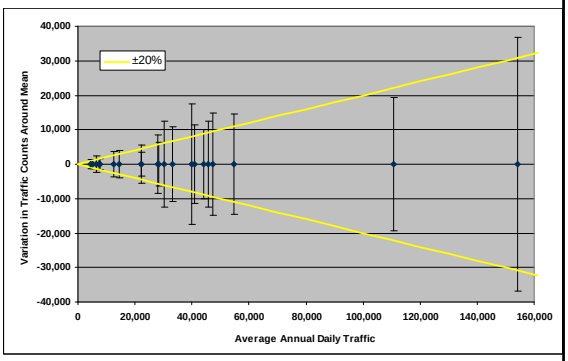
Identical R^2 Values – Different Stories



47

Highway Assignment Validation Data

- Observed volumes are not infallible
  - Source error
  - Sampling error
- Impact on aggregate and disaggregate tests

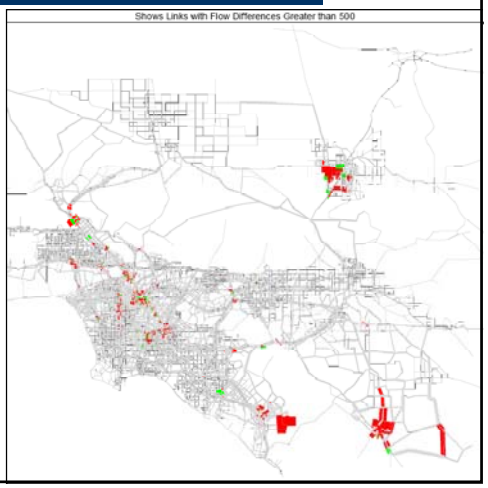


Source: Oak Ridge National Laboratory

48

Highway Assignment Validation Aggregate Checks

- R^2
- RMSE
- VMT
- Screenlines
- Scatterplots
  - Outlier analysis
- Difference Plots ⇨



49

Example VMT Guidelines by Functional Class and Area Type

	Modeled Versus Observed VMT				
Stratification	Ohio[a]	Florida[b]		Michigan[c]	FHWA-1990[c]
Functional Class		Acceptable	Preferable		
Freeways/Expressways	±7%	±7%	±6%	±6%	±7%
Principal Arterials	±10%	±15%	±10%	±7%	±10%
Minor Arterials	±10%	±15%	±10%	±10%	±15%
Collectors	±15%	±25%	±20%	±20%	±20%
All Links		±5%	±2%		
Area Type					
CBD	±10%	±25%	±15%		
Fringe	±10%	±25%	±15%		
Urban	±10%	±25%	±15%		
Suburban	±10%	±25%	±15%		
Rural	±10%	±25%	±15%		

50

Validation Documentation (Ch. 11)

- Purposes of validation documentation
  - Establish confidence in model
  - Identify model capabilities
  - Identify model limitations
  - Limit "misuse" of models…
  - Establish basis for "next round" of improvements

"*Essentially, all models are wrong, but some are useful…the practical question is how wrong do they have to be to not be useful.*" (George Box)

51

Validation Documentation Suggestions

- Stand alone document
  - Establish importance of validation
  - Space to cover necessary information

52

Suggested Validation Documentation Chapters

- Executive summary
  - For the non-modeler (or your executive director)
  - Not the place to mention R^2 or %RMSE
- Model component validation
  - Sensitivity of model parameters
- System validation
  - Mention R^2 or %RMSE here

53

Suggested Validation Documentation Chapters

- Model sensitivities
  - Sensitivity to changes in model inputs
  - Temporal validation
- Model limits (or limitations)
- Recommended model development / enhancement steps
  - Priorities and plans

54