

The logo graphic consists of three overlapping parallelogram shapes pointing to the right. The top shape is light green, the middle is light blue, and the bottom is a darker blue.

CAMBRIDGE  
SYSTEMATICS

Think  Forward

# How to Make the Most of Your Models in a Rapidly Changing World

*presented to*

*All of you fine people*

*presented by*

*Cambridge Systematics, Inc.*

*Thomas Rossi*

April 26, 2023

# DISCLAIMER

---

- *The views and opinions expressed during this webinar are those of the presenters and do not represent the official policy or position of FHWA and **do not constitute an endorsement, recommendation or specification by FHWA.** The webinar is based solely on the professional opinions and experience of the presenters and is made available for information and experience sharing purposes only.*

## Presenter

- **Thomas F. Rossi**, a Principal of Cambridge Systematics, is an internationally recognized expert in travel demand modeling. He has been doing this for more than 40 years but still acts alike he's young. Tom has led or overseen the development of many activity-based and trip-based models throughout the country and has been an expert advisor to federal agencies and metropolitan planning organizations in the development of travel models and survey data collection efforts. For 30 years, he has worked with the U.S. Department of Transportation to conduct research and develop and teach training courses in travel demand modeling. Tom is a past chairman of the Transportation Research Board's Transportation Demand Forecasting Committee.

# A Travel Demand Model ...

---

- ... takes a set of available *input data* ...
- ... and converts it to a set of *output data*, needed for planning analyses ...
- ... using a set of *mathematical formulations*...
- ... which use *parameters* to perform the conversions

# “All models are wrong, but some are useful”

--G. Box

---

- A model is...
  - » An *analytical tool* to provide *important information to planners*
  - » A means to quantitatively estimate the *effects of transportation planning, policy, or investment decisions—or external factors—on transportation demand*
- A model is not...
  - » A crystal ball—it does not predict the future
  - » A way to get “the answer” on a planning decision

# Dealing with Limitations

## (Some of “why all models are wrong...)

---

### ➤ Data limitations

- » Observed data does not cover everything modeled
- » Errors and uncertainties in observed data
- » Inconsistencies among observed data sources

### ➤ Model limitations

- » Limitations in data used for model development
- » Simplifications (in *every* model)
- » Aggregation errors

# How to Create a Useful Model

---

- Define ***what the model needs to provide***
  - » What scenarios do you need to model? Short term? Long term?
- Determine ***how complex the model must be*** to produce the needed information and ***how simple the model can be*** and still provide it
- Consider ***the resources you have*** (data, skills, funding, time)
- Develop ***plans*** (model development, validation) that address these factors

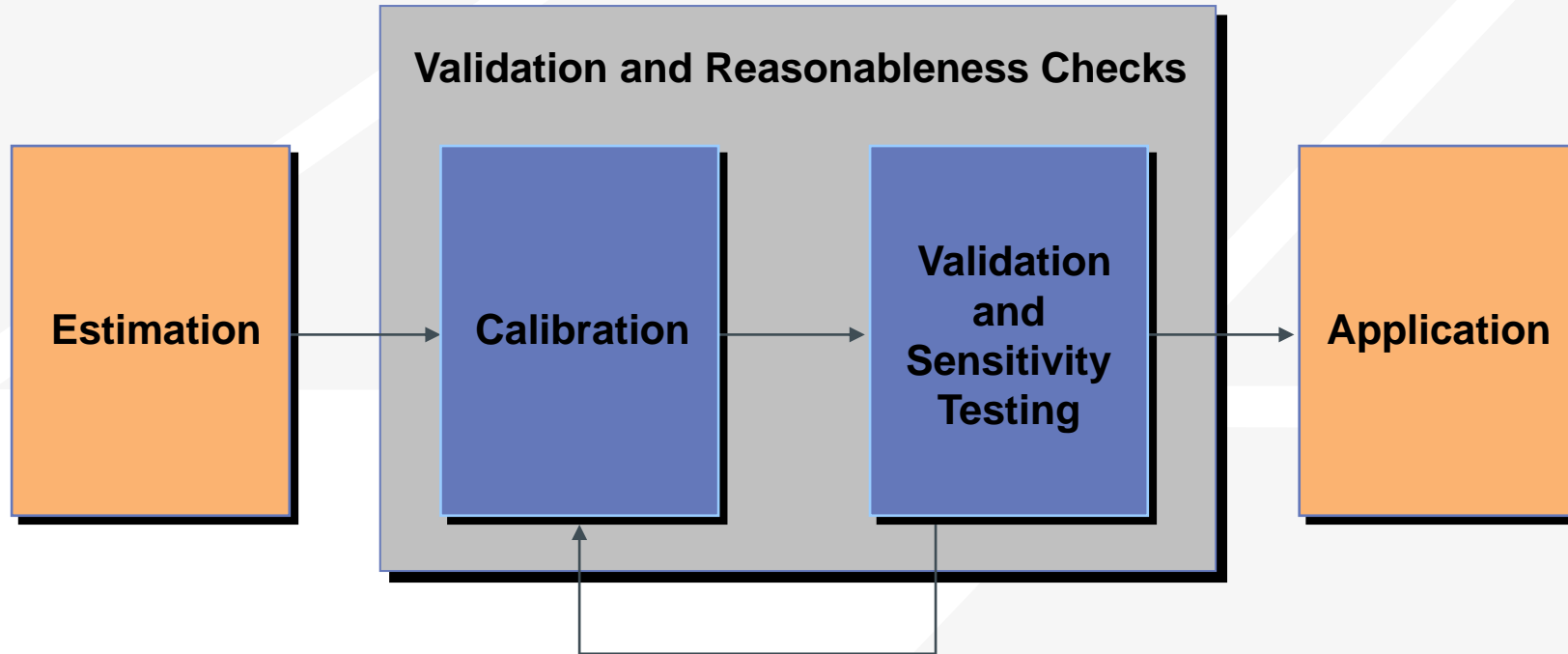
# Key Definitions

---

- **Estimation** – the use of statistical techniques and observed data to develop model parameters
- **Calibration** – the adjustment of model parameters in to make the model results more reasonable
- **Validation** – the application of the calibrated models and comparison of the results to observed data
- **Sensitivity testing** – the application of the model using alternative input data or assumptions to determine reasonableness. **It is a critical part of validation.**



# The Model Development Process



# Model Estimation

---

- We need to get the *parameters* to obtain the model outputs, from the input data, using the mathematical formulations of the model
- Parameters may be
  - » *Estimated* statistically, using data (e.g., from surveys)
  - » *Transferred* from another context
  - » *Asserted*, based on experience

# Model Validation

---

- Confirm that model *accurately reflects travel behavior* in the region, under existing and potential future conditions
  - » Run model for base year, *compare to observed data*
  - » Examine *sensitivity of model results to key variables* (e.g., travel time, cost, demographic changes)
  - » Ensure that *results are reasonable for required types of planning analyses*

# Model Calibration

---

- Model validation checks often show issues that need to be addressed
  - » Not limited to differences between base year results and observed data
  - » Issue could be identified through sensitivity testing
- In some cases changes in parameters are warranted—this is *calibration*
- The goal of calibration is to improve the model's usefulness for planning—*not necessarily to match base year results*

# What We've Been Doing in Validation for Years Is Still Very Important

---

- 2010 FHWA Validation Manual – still the most comprehensive source for validation process/methods
  - » Model validation plan specification
  - » Collection/assessment of validation data
  - » Validation of **all** model components
  - » Sensitivity testing
  - » Documentation of validation results
- But transportation and models are changing...and so must validation
  - » Enriched by greater data availability

The **Travel** Model  
*Improvement*  
Program

**Travel Model Validation and  
Reasonableness Checking  
Manual  
Second Edition**

*Helping Agencies Improve Their Planning Analysis Techniques*

**TMIP**  
Travel Model Improvement Program

# Validation Data Sources

Travel survey data (household, transit, etc.)

“Big Data” – OD flows, time of day, visitor travel, trucks

Traffic count data

Transit ridership data

Census data, including ACS

Other federal sources (e.g., FAF)

NHTS (2017 is the latest)

# Model Validation Plan

---

- Identify validation data sources
- *All components* need to be validated, including demand components, trucks, externals, etc.
- Guidelines for validation tests
  - » Numeric where appropriate
- Single pass validation – all components

# Model Validation Plan (continued)

---

- Highway assignment validation
  - » VMT by segment
  - » Volume checks (key roadways, %RMSE, etc.)
  - » Screenline volumes
- Transit assignment validation
  - » Total ridership by mode/access mode
  - » Route group ridership
- Full feedback validation
- Temporal validation/sensitivity testing
  - » Check sensitivity to recent travel behavior changes (e.g., increased work from home)



# What Has Changed in Recent Years

---

- We are always learning...
  - » Validation guidance
  - » Errors in observed data
- Data sources are changing
  - » Big data sources (e.g., LBS) are expanding and improving
  - » Surveys are changing, too
- Models have changed
- Most importantly – How we travel is changing rapidly
  - » New ways to travel (TNCs, micromobility)
  - » More working from home
  - » Increased ability to do things without traveling
  - » Emphasis on active transportation

# Which Data Sources for Validation Have Errors?

---

- All of them!
  - » Data collection and processing procedures can have errors
  - » Most collected data are *sampled*, and they have *sampling errors*
  - » Some sources have biases
  - » Surveys can have response errors
- Magnitudes and types of errors vary among sources...

# Data Collection and Processing Errors

---

- Collection devices have limitations and do not always function properly
  - » Traffic counting devices
  - » Transit APC
  - » Mobile devices and apps
- Human error in recording/processing data
- Automated error checking algorithms are human-created

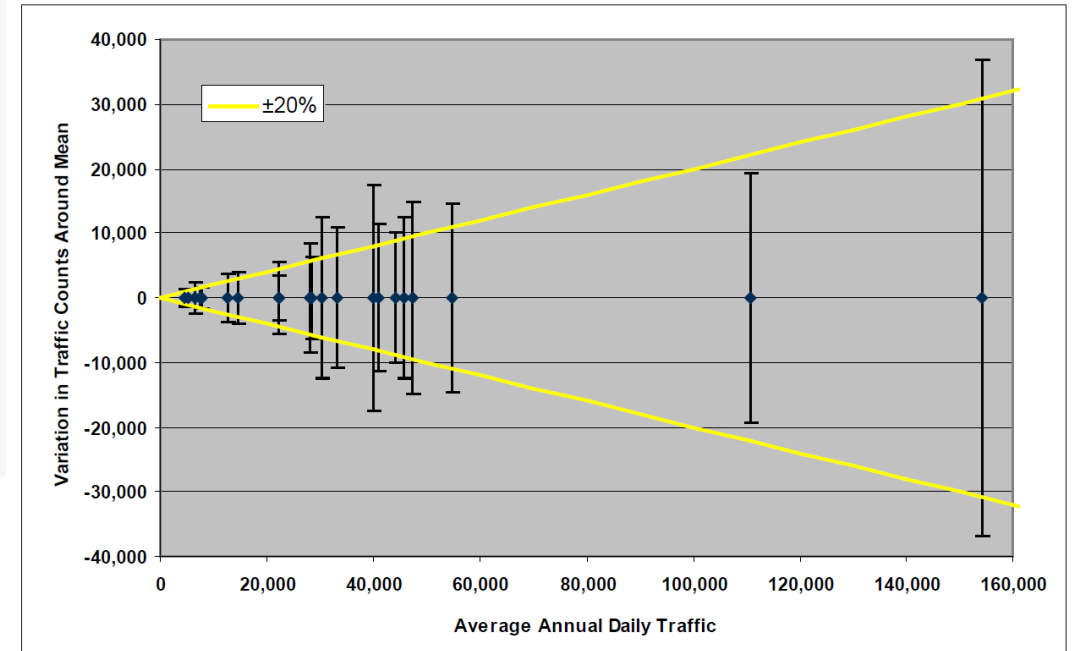
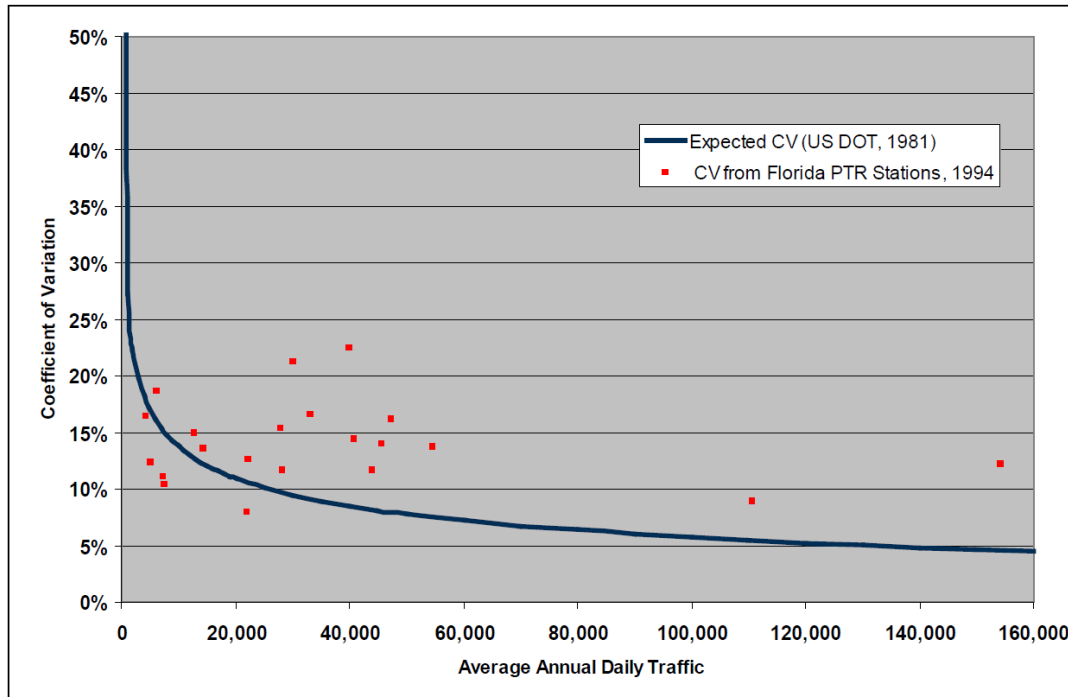
# Guidance on Highway Assignment Validation

---

- Aggregate measures (volume/count ratios, %RMSE, screenline checks) provide a *partial picture* of accuracy
- Individual link checks still have value
- From FHWA model validation manual
  - » Traffic counts for a single facility represent a sample for the link *subject to sampling error*
  - » Substantial variation in daily traffic counts should be assumed
  - » From ORNL study: Many link counts are off by more than 20%

# Traffic Count Error

(from FHWA Validation Manual)



# Impacts of New Mobility Types

---

- TNC mode shares: 2012 – 0%, 2016 – 0.1%, 2019 – 0.4%
  - » When was your survey data collected?
- Reductions in transit and taxi mode shares

# Use of Big Data

---

## ➤ Passively collected location data (LBS)

### » Great for...

- Very recent data (surveys may be many years old)
- O-D information (much larger sample size than surveys)
- Time of day
- Tour information

### » Limitations

- Some users have low device incidence (low income, young children, elderly?)
- No demographic information (e.g., vehicle ownership, income)
- Very limited travel purpose information
- Very limited travel mode information

### » Optimally, use with survey data

# To Conclude...Some Helpful Hints

---

- Recognize that your model has errors
  - » And you can't eliminate them completely through validation
- Don't focus only on matching observed data
  - » Observed data have errors, too
  - » Matching them can make model insensitive to important factors
- Validate model according to what you need it to do
- Don't try to create one super model for all applications!



Sometimes here there is a slide that says, “Questions?,”  
but I’ve found that there are questions  
whether there is a slide or not...